
POČÍTAČOVÉ VIDENIE

DETEKCIA A ROZPOZNÁVANIE OBJEKTOV

ELENA ŠIKUDOVÁ

*Fakulta matematiky, fyziky a informatiky
Univerzita Komenského, Bratislava*

ZUZANA ČERNEKOVÁ

*Fakulta matematiky, fyziky a informatiky
Univerzita Komenského, Bratislava*

WANDA BENEŠOVÁ

*Fakulta informatiky a informačných technológií
Slovenská technická univerzita, Bratislava*

ZUZANA HALADOVÁ

*Fakulta matematiky, fyziky a informatiky
Univerzita Komenského, Bratislava*

JÚLIA KUČEROVÁ

*Fakulta matematiky, fyziky a informatiky
Univerzita Komenského, Bratislava*



WIKINA



Venujem mojim rodičom a neterkám Lucke a Peťke
Elena Šikudová

Venujem mojim rodičom
Zuzana Černeková

Venujem môjmu otcovi in memoriam
Wanda Benešová

Venujem Ome a Opovi
Zuzana Haladová

Venujem mojim rodičom a bratovi
Júlia Kučerová



Obsah

Predslov

ii

I Predspracovanie obrazu

1	Spracovanie obrazu vo frekvenčnej oblasti	2
1.1	Načo slúžia diskrétné lineárne transformácie?	2
1.1.1	Príklad	3
1.1.2	Použitie	4
1.1.3	Lineárna transformácia obrazu.	5
1.2	Fourierova transformácia	5
1.2.1	Fourierov rad	5
1.2.2	Spojité Fourierova transformácia	7
1.2.3	Dvojdimenziálna Fourierova transformácia	8
1.2.4	Vlastnosti Fourierovej transformácie	9
1.2.5	Diskrétna Fourierova transformácia (DFT)	12
1.2.6	DFT ako špeciálny prípad lineárnej transformácie	13
1.2.7	Dvozmerná DFT	15
1.3	Výpočet diskretnej Fourierovej transformácie	15
1.3.1	Výpočet z definície	15
1.3.2	Výpočet pomocou maticových operácií	16
1.3.3	Rýchla Fourierova transformácia (FFT)	17
1.3.4	Doplňovanie nulami	21
1.4	Príklady Fourierovej transformácie obrazu	22

1.5	Kosínusová transformácia	26
1.5.1	Diskrétna kosínusová transformácia	26
1.6	Filtrácia obrazu vo frekvenčnej oblasti	27
1.6.1	Filtrácia vybraných frekvenčných oblastí	35
2	Spracovanie obrazu v priestorovej oblasti	38
2.1	Diskrétna konvolúcia	38
2.1.1	Diskrétna 2D konvolúcia	39
2.1.2	Vlastnosti konvolúcie	39
2.2	Výpočet konvolúcie	39
2.2.1	Znázornenie výpočtu konvolúcie	40
2.3	Najdôležitejšie konvolučné masky	41
2.3.1	Vyhladzovacie filtre	41
2.3.2	Hranové filtre	43
2.4	Použitie konvolúcie	47
2.5	Nelineárna filtrácia	48
2.5.1	Zorad'ovací filter	48
3	Konvolučná teoréma	50
4	Morfologické operácie	52
4.1	Matematická morfológia	52
4.1.1	Morfologické spracovanie obrazu	52
4.2	Základné pojmy a definície	53
4.2.1	Susednosť, súvislosť	53
4.2.2	Minkowského suma	55
4.2.3	Štrukturálny element	55
4.2.4	Morfologická transformácia	55
4.3	Základné morfologické operácie	56
4.4	Binárna dilatácia a erózia	57
4.4.1	Dilatácia	57
4.4.2	Erózia	59
4.4.3	Vlastnosti dilatácie a erózie	61
4.4.4	Použitie dilatácie a erózie	61
4.4.5	Vzájomný vzťah medzi eróziou a dilatáciou	61
4.5	Morfologické otvorenie a uzavretie	62
4.5.1	Morfologické otvorenie	62
4.5.2	Morfologické uzavretie	62
4.6	Šedotónová dilatácia a erózia	64
4.6.1	Nerovný štrukturálny element	64
4.6.2	Rovný štrukturálny element	65

4.7	Morfologický gradient	68
4.8	Morfologická vrchná a spodná časť klobúka	70
4.9	Granulometria	71
4.10	Morfologická operácia Hit-or-Miss	73
4.11	Vzdialenostná transformácia	75
4.12	Kostra	75
4.12.1	Príklady algoritmických postupov výpočtu kostry	76
4.13	Transformácia rozvodím	76
4.13.1	Značková segmentácia rozvodím pre farebný obraz	78
4.14	Geodetické metódy	79
4.14.1	Geodetická vzdialenosť	79
4.15	Morfologická rekonštrukcia	79
4.15.1	Geodetická binárna dilatácia	80
4.15.2	Geodetická binárna erózia	81
4.15.3	Morfologická rekonštrukcia a jej použitie	81
4.16	Morfologická šedotónová rekonštrukcia	82
4.16.1	Geodetická šedotónová dilatácia a erózia	82
4.16.2	Generovanie značiaceho obrazu	83
4.16.3	Príklady použitia	84
4.17	Príklad použitia morfologických operácií pre segmentáciu superpixelov	85
4.17.1	Vyhľadanie morfologickou rekonštrukciou	87
4.17.2	Generovanie značiek – semiačok pre značkovú transformáciu rozvodím	88
4.17.3	Transformácia rozvodím	88
4.17.4	Vyhodnotenie	89
4.17.5	Výsledky	90
	Zoznam použitej literatúry	92

II Príznamy

5	Príznamy v oblasti počítačového videnia	96
6	Príznamy tvaru	98
6.1	Binárny obraz	98
6.2	Opis tvaru založený na hranici oblasti	99
6.2.1	Ret'azcový Freemanov kód	100
6.2.2	Obvod	101
6.2.3	Krivosť krivky	102

6.2.4	Signatúra	104
6.2.5	Fourierov opis	105
6.3	Opis tvaru založený na oblasti	105
6.3.1	Ťažisko a k nemu prislúchajúce príznaky	106
6.3.2	Príznaky odvodené z priemeru a obsahu	107
6.3.3	Osi oblasti	108
6.3.4	Opísaný obdĺžnik	110
6.3.5	Topologické príznaky	111
6.3.6	Morfologické príznaky	112
6.4	Šedotónový obraz	113
6.4.1	Projekcie	113
6.4.2	Opisné štatistiky	113
6.4.3	Energia a entropia	115
6.5	MPEG-7	115
6.5.1	Tvar kontúry	115
6.5.2	Tvar regiónu	117
7	Príznaky farby	118
7.1	Dominantná farba	118
7.2	Histogram farieb	118
7.2.1	Histogram priestorového rozloženia farieb	120
7.2.2	Ďalšie opisy priestorového rozloženia farieb	120
7.3	MPEG-7	122
7.3.1	Deskriptor dominantnej farby	123
7.3.2	Škálovateľný deskriptor farby	123
7.3.3	Deskriptor rozloženia farieb	124
7.3.4	Deskriptor štruktúry farieb	124
8	Príznaky textúry	128
8.1	Autokorelácia	129
8.2	Matica opakovaných výskytov	129
8.3	Diskrétna Fourierova transformácia	130
8.4	Gaborove filtre	131
8.5	MPEG-7	133
8.5.1	Prehľadávanie textúr	133
8.5.2	Homogénne textúry	134
8.5.3	Histogram lokálnych hrán	135
9	Lokálne príznaky	138
9.1	Lokálne vs. globálne príznaky	138
9.2	Lokálne príznaky	140

9.3	Metódy detekcie a deskripcie lokálnych príznakov	141
9.3.1	Harrisov detektor rohov	142
9.3.2	FAST	142
9.3.3	SIFT	143
9.3.4	SURF	147
9.3.5	BRIEF	148
9.3.6	ORB	148
9.3.7	Porovnávanie deskriptorov	150
9.4	Registrácia objektov	150
9.5	Využitie lokálnych príznakov	151
9.5.1	Spájanie panorám	151
9.5.2	Stereo rekonštrukcia	152
9.5.3	Vlastná kalibrácia kamery	152
9.5.4	Sledovanie objektov	152
9.5.5	Vyhľadávanie obrazov	153
9.6	Prípad použitia: Klasifikácia malieb	153
9.7	3D lokálne príznaky	156
10	Predspracovanie príznakov	158
10.1	Príznaky bez rozmeru	160
10.2	Škálovanie	160
11	Výber príznakov	162
11.1	Filter a obálka	162
11.2	Dopredný výber	163
11.3	Spätný výber	164
11.4	Kombinovaný výber	165
11.5	Ohodnotenie príznakov	165
11.5.1	Štatistická závislosť	165
11.5.2	Konzistencia	166
11.5.3	Medzitriedna vzdialenosť	167
11.5.4	Informačno-teoretické miery	167
11.5.5	Hodnotiace miery	169
12	Redukcia príznakov	170
12.1	Metóda hlavných komponentov	170
12.1.1	Použitie PCA	172
12.1.2	Ako môže PCA znížiť dimenziu?	174
12.1.3	Vzťah medzi PCA a SVD	176
12.2	Fisherova lineárna disriminačná analýza	176

Zoznam použitej literatúry	180
----------------------------	-----

III Klasifikácia

13 Klasifikácia príznakov	188
14 Riadené klasifikátory	192
14.1 Bayesovská teória rozhodovania	192
14.1.1 Optimalita Bayesovho klasifikátora	193
14.1.2 Rozhodovacie hranice	195
14.2 Techniky najbližších susedov	195
14.3 Rozhodovacie stromy	196
14.3.1 Rozdeľujúce kritérium	199
14.3.2 Ukončujúce kritérium	200
14.3.3 Orezávanie rozhodovacích stromov	201
14.3.4 Výhody a nevýhody rozhodovacích stromov	203
14.4 Diskriminačná analýza	203
14.4.1 Lineárne diskriminačné funkcie	203
14.4.2 Klasifikácia do dvoch tried	204
14.4.3 Určenie váh	206
14.4.4 Voľba účelovej funkcie	207
14.4.5 Klasifikácia do K tried	210
14.5 Mechanizmy podporných vektorov	211
14.5.1 Neseparovateľné dáta	215
14.6 Neurónové siete	219
14.6.1 Aktívne funkcie	220
14.6.2 Viacvrstvový perceptrón	221
14.6.3 Učenie neurónovej siete	221
14.6.4 Rýchlosť učenia sa neurónovej siete	225
14.6.5 Ukončenie učenia sa neurónovej siete	226
15 Neriadené klasifikátory	230
15.1 Metóda K priemerov	230
15.1.1 Metóda C priemerov	232
15.2 Hierarchické zhlukovanie	233
15.3 Samoorganizujúce sa mapy	235
16 Hodnotenie kvality klasifikácie	240
16.1 Vzájomná validácia	242
16.1.1 Náhodné prevzorkovanie	242

16.1.2	K-násobná vzájomná validácia	242
16.1.3	Vrstvená K-násobná vzájomná validácia	243
16.1.4	Vzájomná validácia s vynechaním jednej vzorky	243
16.2	Bootstrap	243
16.3	Chybová funkcia klasifikátora	244
16.4	Metódy porovnávania klasifikátorov	246
16.4.1	Matica zámen	246
16.4.2	Koeficienty kvality klasifikácie	247
16.4.3	ROC krivky	249
16.4.4	McNemarov test	254
Zoznam použitej literatúry		256

IV Farba a vizuálne vnímanie

17	Svetlo a farba	260
17.1	Veličiny opisujúce svetlo	262
17.1.1	Rádiometrické veličiny	263
17.1.2	Fotometrické veličiny	264
17.2	Funkcie spektrálnej hustoty	265
17.3	Geometrické vlastnosti svetla	265
17.4	Žiarenie čierneho telesa	267
18	Ľudský vizuálny systém	270
18.1	Štruktúra oka	271
18.2	Procesy prebiehajúce na sietnici	273
18.2.1	Trichromatické procesy	273
18.2.2	Oponentné procesy v sietnici	274
18.2.3	Gangliové bunky a receptívne polia	276
18.3	Vizuálny kortex	277
18.4	Poruchy farbecitu	278
19	Kolorimetria	280
19.1	Grassmanov experiment	281
19.2	Grassmanov zákon aditívneho miešania farieb	282
19.3	Kolorimetrický priestor	282
19.4	Štandardný kolorimetrický pozorovateľ	287
19.4.1	CIE 1931 štandardný kolorimetrický pozorovateľ	287
19.4.2	CIE 1964 doplnkový štandardný kolorimetrický pozorovateľ	288

19.5	Kolorimetrická sústava CIE XYZ	289
19.6	CIE diagram chromatickosti	291
19.7	MacAdamove elipsy	295
19.8	Štandardné iluminanty	296
19.8.1	Štandardné zdroje svetla	298
20	Modely a priestory farieb	300
20.1	Homogénne priestory farieb	300
20.1.1	CIE 1960 $\bar{Y}uv$	301
20.1.2	CIE 1964 $U^*V^*W^*$	301
20.1.3	CIE 1976 $\bar{Y}u'v'$	302
20.1.4	CIE 1976 $L^*u^*v^*$	302
20.1.5	CIE 1976 $L^*a^*b^*$	304
20.1.6	CIE CAM02	306
20.2	Hardvérovo orientované modely	306
20.2.1	Lineárny model RGB	307
20.2.2	Model RGBA	308
20.2.3	Nelineárne priestory RGB	308
20.2.4	Model CMY	310
20.2.5	Model CMYK	312
20.3	Modely farieb v televíznej a videotechnike	313
20.3.1	Model YIQ	315
20.3.2	Model YUV	316
20.3.3	Model YPbPr	317
20.3.4	Model YCbCr	318
20.3.5	Model Kodak Photo YCC	319
20.4	Používateľsky orientované modely	321
20.4.1	Model HSV	322
20.4.2	Model HSL	326
21	Vizuálne vnímanie a pozornosť	328
21.1	Vizuálna pozornosť	329
21.2	Určovanie významných oblastí v obraze	331
21.2.1	Sledovanie pohybu očí	332
21.2.2	Modely na detekciu významných oblastí	336
21.2.3	Model pre rýchlu analýzu scény	341
21.3	Využitie detekcie významných oblastí v obraze	346
21.3.1	Hľadanie najlepších pohľadov na objekty	346
21.3.2	Kompresia obrazu	346
21.4	Príklad využitia: Hodnotenie kvality videa	347
21.4.1	Detekcia významných oblastí	347

21.4.2 Hodnotenie kvality videa	348
Zoznam použitej literatúry	350

V Dodatky

A Vybrané témy z pravdepodobnosti	356
A.1 Pravdepodobnosť	356
A.1.1 Nezávislosť javov, podmienená pravdepodobnosť	358
A.2 Náhodná premenná	360
A.3 Číselné charakteristiky náhodnej premennej	361
A.3.1 Momenty	362
A.3.2 Kvantily	364
A.3.3 Modus	365
A.4 Náhodné vektory	365
A.4.1 Marginálne rozdelenie pravdepodobnosti	366
A.5 Číselné charakteristiky náhodného vektora	366
A.5.1 Združené momenty	366
A.5.2 Marginálne charakteristiky	367
A.5.3 Kovariancia	368
A.6 Normálne rozdelenie	369
A.6.1 Viacrozmerné normálne rozdelenie	369
B Vybrané témy z maticového počtu	372
B.1 Singulárny rozklad matice	372
B.1.1 Úsporná forma SVD	373
B.1.2 Vlastnosti SVD	373
C Vybrané kapitoly z lineárnej algebry	376
C.1 Vektorový priestor (Lineárny priestor)	376
C.2 Lineárna kombinácia vektorov	377
C.3 Lineárna transformácia	377
C.4 Báza vektorového priestoru	377
Zoznam použitej literatúry	380

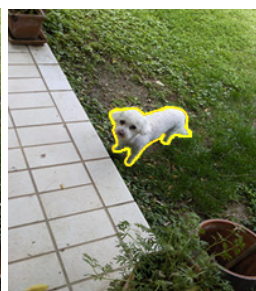
Predslov

Vizuálna detekcia a rozpoznávanie objektov patrí v poslednom desaťročí medzi najväčšie výzvy počítačového videnia. Potenciálne uplatnenie systémov detekcie a rozpoznávania objektov je veľmi široké, od bezpečnostných systémov (napr. identifikácia oprávnených osôb), cez medicínske techniky (napr. detekcia tumorov), priemyselné aplikácie (napr. vizuálna inšpekcia výrobkov), robotiku (napr. navigácia robota v nedostupnom teréne) až po rozšírenú realitu a mnohé ďalšie.

Úloha **automaticky vyhľadať objekt** sa spravidla formuluje ako úloha segmentácie objektu, prípadne iba ako detekcia opísaného štvoruholníka, alebo ako rozpoznanie samotného objektu.



Obrázok 0.1: *
Detekcia objektu
bez segmentácie



Obrázok 0.2: *
Detekcia objektu
so segmentáciou



Obrázok 0.3: *
Rozpoznanie objektu
v obraze

K úlohe detekcie môžeme pristupovať na rôznych stupňoch zovšeobecnenia. Bud' **rozpoznávame konkrétne inštancie** jednotlivých objektov, alebo celú **kategóriu objektov**.

Vizuálna detekcia a rozpoznávanie objektov je veľmi komplexný problém. Musí sa zaoberať mnohými aspektmi, ktoré riešenie rozhodne neľahčujú. Príkladom sú:

- nehomogénne a premenlivé osvetlenie scény,
- meniace sa spektrálne charakteristiky osvetlenia a následne i farba snímaného objektu,
- objekty, ktoré sa prekrývajú,
- rôzne uhly snímania objektov kamerou,
- v prípade rozpoznávania kategórií je samozrejme veľkou výzvou i rôznorodosť objektov v rámci jednej kategórie.

Základný reťazec operácií systému detekcie rozpoznávania môžeme znázorniť nasledovne:



Základné členenie tejto knihy korešponduje s týmto reťazcom. Navyše bude čitateľ oboznámený s problematikou farby a ľudského vnímania vizuálnej informácie, ktoré sú veľmi dôležité pre riešenie detekcie a rozpoznávania objektov a zároveň majú presahy do iných významných odborov, ako sú počítačová grafika alebo interakcia človeka s počítačom.

Na vstupe uvedeného reťazca je digitálna kamera, z ktorej získame **signál obrazu**, či už farebný alebo čiernobiely. Úlohou **pedspracovania obrazu** je odstránenie nežiaducich a pre ďalšie spracovanie rušivých javov ako sú napr. šum v signáli, zvýraznenie informácie (napr. kontúr objektov), ktorá je relevantná a naopak potlačenie informácie, ktorá nie je relevantná z pohľadu danej úlohy.

Dobrým príkladom je filtrácia pre odstránenie šumu, hornopriepustná filtrácia pre odstránenie nehomogenity osvetlenia, prevod do iných farebných priestorov alebo morfológické spracovanie.

Cieľom ďalšieho stupňa spracovania je **získanie príznakov**, ktoré budú dobre reprezentovať obrazovú informáciu dostatočne reprezentatívnym a diskriminatívnym spôsobom.

Ak je našou úlohou napríklad detekcia ľudskej tváre v obraze, celý obraz budeme postupne prechádzať tzv. kľzavým oknom a v každej jeho

pozícii bude úlohou klasifikátora rozhodnúť medzi dvoma triedami – je alebo nie je tvár v tej časti obrazu, ktorú práve vyšetruje? Vstupom pre klasifikátor je práve vektor príznakov, na základe ktorého môže klasifikátor rozhodnúť. V procese návrhu môžeme vygenerovať vektor príznakov podstatne väčší, než ktorý bude použitý vo výslednom klasifikátore. Vhodné (relevantné) príznaky určíme metódami selekcie príznakov.

V ďalšom rozhodovacom stupni **klasifikátor** na základe vektora príznakov rozhodne o príslušnosti k triede.

Každý **systém vizuálneho rozpoznávania objektov** musí byť najskôr naučený na úlohu, pre ktorú je určený. Definovanie rozhodovacej úlohy pomocou súboru deterministických podmienok sa používa skôr zriedkavo. Príkladom môžu byť niektoré systémy priemyselnej inšpekcie, kde v reprodukovateľných podmienkach systém vyhodnotí namerané veličiny, napr. rozmery objektu. Záujem vedcov a výskumníkov sa však koncentruje predovšetkým na metódy na princípe umelej inteligencie.

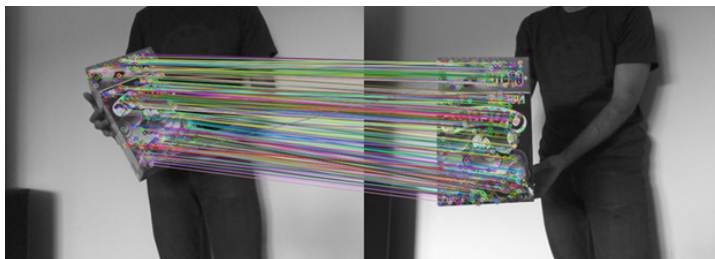
Metódy detekcie a rozpoznávania objektov môžeme podľa základného prístupu k riešeniu rozdeliť na:

- metódy založené na globálnom opise,
- metódy založené na lokálnych príznakoch.

Väčšina metód založených na globálnom opise používa metódu klzaového okna, ktoré vyberie určitý región obrazu a tento je vyhodnotený porovnávaním s danými vzorovými reprezentáciami objektov. Tento prístup sa úspešne používa napr. pri detekcii tváre podľa autorov Viola Jones. Principiálne však má niekoľko nedostatkov, napr. škálovú a rotačnú nezávislosť dosahujeme porovnávaním s otočeným vzorom v rôznej škále, čo je výpočtovo náročné. Prípadné prekryvy objektov alebo zmena tvaru nerigidných objektov veľmi znižujú pravdepodobnosť správnej detekcie.

Druhou alternatívou je využitie lokálnych príznakov. Tieto metódy pracujú smerom zdola nahor, od opisu pomerne malého okolia mnohých kľúčových bodov až po rozhodnutie, či tieto určité zostavy lokálnych deskriptorov patria k detegovanému objektu. Najjednoduchší postup pozostáva z troch základných krokov:

1. Extrakcia lokálnych príznakov. Lokálne príznaky sa počítajú v okolí detegovaných kľúčových bodov a to nezávisle najskôr na vzorovom obraze objektu a potom v obraze, ktorý vyšetrujeme.
2. Porovnanie lokálnych príznakov tak, aby sa našli korešpondencie podobných.
3. Overenie, či spárované príznaky sa nachádzajú v konzistentnej geometrickej konfigurácii.



Obrázok 0.4: *

Príklad párovania lokálnych príznakov

Veľmi perspektívne (v tomto vydaní knihy zatiaľ bez podrobnejšieho opisu) sa javia metódy nazývané ako „vrecia príznakov“. Základná myšlienka sa dá vyjadriť nasledovne. Najskôr sa vygeneruje kódová kniha nazývaná vizuálny slovník. S jednotlivými inštanciami lokálnych príznakov sa ďalej pracuje na základe indexu v kódovej knihe (tzv. vizuálne slovo), ku ktorému patria. Uplatňujú sa metódy tvorby histogramov podľa príslušnosti jednotlivých príznakov k vizuálnym slovám, metódy prieniku týchto histogramov alebo spätného indexovania. Tieto metódy sú veľmi slubné hlavne pre úlohy vyhľadávania objektov vo veľkých databázach.

Cieľom tejto vedeckej monografie je zoznámiť čitateľa s teóriou vizuálnej detekcie a rozpoznávania objektov a uviesť príklady moderných aplikácií, v ktorých je táto teória využívaná.

Členenie knihy je nasledovné.

Prvá kapitola sa zaoberá obrazom, jeho snímaním a prechodom od analógového signálu k digitálnemu.

Ďalej nasleduje časť *Predspracovanie obrazu*, kde sa čitateľ dozvie princípy filtrácie vo frekvenčnej oblasti, v priestorovej oblasti a morfológického spracovania obrazu spolu s príkladmi použitia týchto metód.

V časti nazvanej *Príznyky* sú predstavené početné spôsoby výpočtu globálnych príznakov reprezentujúcich tvar objektu, farbu, textúru ako i lokálne príznaky. Vhodným príkladom využitia globálnych a lokálnych príznakov je klasifikácia malieb na fotografiách. Posledná kapitola tejto časti sa zaoberá metódami vhodného výberu príznakov.

V časti *Klasifikácia* sú opísané najvýznamnejšie klasifikačné metódy rozdelené podľa toho, či sú riadené (s učiteľom) alebo neriadené (bez učiteľa). Samostatnou kapitolou sú spôsoby hodnotenia kvality klasifikácie.

Posledná časť monografie sa nazýva *Farba a vizuálne vnímanie*. Opisuje základné fyzikálne a biologické pojmy používané v počítačovom vi-

dení, metódy kolorimetrie a modely rôznych farebných priestorov. Kapitola o vizuálnej „výraznosti“ (saliency) v obraze a sledovaní pozornosti pozorovateľa uzatvára túto časť. Príkladom súčasného využitia výraznosti je hľadanie najlepších pohľadov na objekty alebo detekcia oblasti záujmu pre zlepšenie kódovania obrazu a videa.

V doplnku čitateľ nájde prehľad vybraných tém z matematiky, ktoré sú nutné pre pochopenie teoretických častí knihy.

Pod'akovanie

Autorky by rady pod'akovali kolegom a priateľom, ktorí pomohli pri realizácii tejto knihy. Obálku navrhol pán Spyridon Giochallas. Fotografie poskytl Dr. Barbora Benešová, Mgr. Peter Berger, Dr. Martina Bátorová a autorky. Obrázky kreslili Dr. Ivana Varhaníková, Spyridon Giochallas a autorky. Skorektúrou textu pomohol doc. Andrej Ferko, a vylepšeniami prispeli recenzenti prof. Ivan Bajla, ÚMER SAV, Bratislava, prof. Jaroslav Polec, ÚT FEI STU, Bratislava a doc. Jarmila Pavlovičová, ÚT FEI STU, Bratislava. Pod'akovanie patrí aj agentúre KEGA, ktorá finančnou podporou projektu 068UK-4/2011 umožnila vydanie publikácie.

ČASŤ I

PREDSPRACOVANIE OBRAZU

Spracovanie obrazu vo frekvenčnej oblasti

Najskôr odporúčame čitateľovi pripomenúť si kapitoly z lineárnej algebry, ktoré sú potrebné pre spracovanie obrazového signálu vo frekvenčnej oblasti. Na pomoc uvádzame v dodatku C najdôležitejšie základné pojmy a definície. Vektorový priestor – Lineárny priestor (viď C.1), Lineárna kombinácia (viď C.2), Lineárna transformácia (viď C.3) a Báza vektorového priestoru (viď C.4).

Teoretické základy spracovania vo frekvenčnej oblasti budú v ďalšom výklade najskôr objasnené pre 1-dimenzionálny signál a vysvetlená teória bude ďalej rozšírená na obrazový, čiže 2-dimenzionálny signál.

1.1 Načo slúžia diskrétné lineárne transformácie?

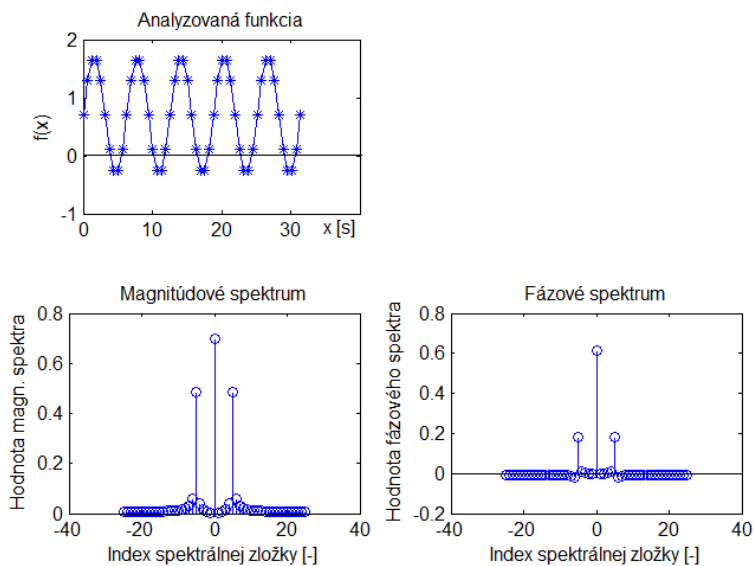
Fourierova transformácia je jedným z najdôležitejších matematických nástrojov používaných pri spracovaní signálov. Popri **Diskrétnej Fourierovej transformácii (DFT)** našla významné uplatnenie v technickej praxi i **Diskrétna kosínusová Transformácia (DCT)**. DFT a DCT sú spolu s vlnkovou transformáciou (Wavelet transform) najpoužívanejšie pri filtrácii obrazu.

DFT a DCT patria medzi lineárne transformácie, ktoré nám umožňujú reprezentovať číslcový vstupný signál ako lineárnu kombináciu diskrétnych báзовých funkcií (viď C.2 a C.4). Lineárne transformácie DFT a DCT teda umožňujú vzájomne jednoznačný prevod signálu (priamo i spätne) medzi reprezentáciou signálu v časovej/priestorovej doméne a frekvenčnej doméne.

Obidve transformácie patria do rodiny tzv. goniometrických transformácií.

1.1.1 Príklad

Pre ilustráciu si najskôr ukážeme jednoduchý príklad.



Obrázok 1.1: Príklad spektrálnej analýzy sínusového signálu.

Na obrázku 1.1 hore je znázornený číslicový vstupný signál, ktorý je generovaný ako jednoduchý sínusový signál. Spektrum je počítané z časového okienka ktoré zodpovedá piatim periódam sínusového signálu. Hodnoty signálu boli nadôvažok ešte zmenené pripočítaním kladnej konštanty, dôsledkom čoho má tento vstupný signál nenulovú strednú hodnotu.

Obrázok 1.1 dole ilustruje výstup frekvenčnej analýzy tohoto signálu, konkrétne po transformovaní do frekvenčnej oblasti prostredníctvom diskkrétnej Fourierovej transformácie. Vľavo dole je magnetúdové spektrum a vpravo dole je fázové spektrum. Očakávame, že magnetúdové spektrum bude mať nenulový nultý koeficient, pretože vstupný signál má nenulovú strednú hodnotu. (Nulový koeficient je nazývaný tiež ako jednosmerná zložka, tzv. DC komponent.) Očakávame tiež, že v spektre uvidíme

jednu výraznú spektrálnu hodnotu, ktorá zodpovedá frekvencii vstupného sinusového signálu.

Na obrázku 1.1 dole vidíme, že vypočítané magnitúdové spektrum je symetrické okolo koeficientu s indexom nula v strede spektra. Spektrálny koeficient s indexom nula reprezentuje jednosmernú zložku signálu, teda v našom prípade strednú hodnotu signálu. Podľa očakávania je nenulový. V spektre tiež vidíme nenulový 5. spektrálny koeficient symetricky rozložený do dvoch častí magnitúdového spektra, ktorý indikuje frekvenciu vstupného sínusového signálu.

Poznámka: **Fourierovské magnitúdové spektrum je pre reálne signály vždy symetrické.**

1.1.2 Použitie

V technickej praxi je výpočet spektra veľmi významný nástroj, pretože umožňuje analýzu a spracovanie signálu selektívne podľa jednotlivých frekvencií obsiahnutých v signáli.

Výhodne môžeme **frekvenčnú analýzu použiť pre opis a tvorbu príznakov**, napr. rôznych textúrových vzorov.

Dôležitým príkladom použitia je aj filtrácia obrazu, a to priamo vo frekvenčnej oblasti. Takto môžeme **odfiltrovať napr. vysokofrekvenčné zložky, alebo naopak odfiltrovať nízko-frekvenčné zložky obrazu**. Po odfiltrovaní vysokých frekvencií z obrazu zmiznú jemné hrany a obraz bude pôsobiť rozmazane. Naopak, ak odfiltrujeme nízke frekvencie, tak získame hranový obraz. Samozrejme záleží od konkrétneho návrhu filtra.

Vo frekvenčnej oblasti môžeme tiež **cielené odstrániť určitú konkrétnu rušivú zložku obrazu**, ktorá má periodický charakter, a ktorú vieme v spektre lokalizovať.

Ďalšou významnou vlastnosťou je **dekorelácia spektrálnych koeficientov**. Z tohoto pohľadu má výborné vlastnosti DCT, čo predurčuje jej použitie pri kompresii dát, napr. v norme JPEG.

Transformovaním obrazu pomocou priamej alebo spätnej lineárnej transformácie takpovediac „prepíname“ medzi

- priestorovou doménou obrazu (alebo časovou doménou, ak sa jedná o jednodimenzionálne signály meniace sa v čase) a
- spektrálnou doménou daného vektorového priestoru,

do ktorého transformujeme (napr. Fourierovho). Ak vypočítané hodnoty nemeníme, tak priamym a spätným transformovaním nedochádza k žiadnej strate informácie (až na malé odchýlky spôsobené limitovanou pres-

nosťou výpočtu). Spektrum daného signálu teda obsahuje kompletnú informáciu o danom signáli.

1.1.3 Lineárna transformácia obrazu.

Diskrétna lineárna transformácia v prípade 1-dimenzionálneho signálu priradí uje postupnosti hodnôt číslícového signálu **postupnosť spektrálnych koeficientov**.

Ako už bolo spomenuté, cieľom diskkrétnej Fourierovej transformácie je **reprezentácia signálu ako lineárnej kombinácie goniometrických funkcií**, ktoré tvoria bázu. Získané súradnice sú vlastne spomínané **váhové koeficienty lineárnej kombinácie** a v technickej praxi ich nazývame **spektrálne koeficienty**.

Fourierova transformácia je definovaná nad oborom komplexných čísiel a spektrálne koeficienty sú komplexné čísla, pozostávajú z reálnej a imaginárnej zložky. V praxi sa častejšie používa prepočet reálnej a imaginárnej zložky na magnitúdu a fázu komplexného čísla, teda máme spektrum magnitúdové a fázové.

1.2 Fourierova transformácia

Ako prvý publikoval francúzsky matematik Jean Baptiste Joseph Fourier v roku 1822 myšlienku, že každú periodickú funkciu je možné vyjadriť ako súčet sínusových a kosínusových funkcií o rôznych frekvenciách vynásobených násobnými koeficientami. Význam Fourierovej vety o reprezentácii periodického signálu je prevratný, keďže každú periodickú funkciu môžeme reprezentovať ako Fourierov rad.

V centre záujmu nášho ďalšieho výkladu bude predovšetkým Diskrétna Fourierova Transformácia dvoch premenných pre spracovanie číslícového signálu obrazu a jej praktické využitie. Pre úplnosť však stručne najskôr uvedieme tiež definíciu Fourierovho radu a definíciu spojitaj Fourierovej transformácie.

1.2.1 Fourierov rad

Máme periodický signál, ktorý matematicky môžeme zapísať ako funkciu $f(x)$ s periódou $T > 0$. Túto periodickú funkciu môžeme interpretovať ako jej rozvoj na nekonečný rad sínusových a kosínusových funkcií.

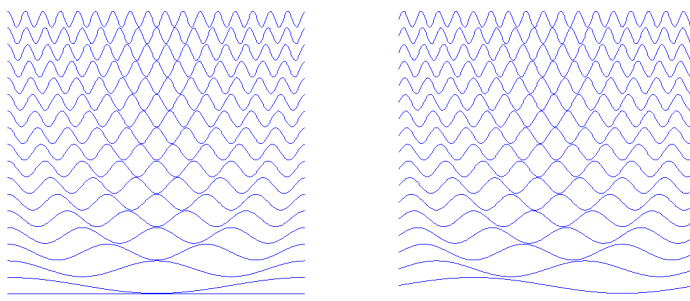
Na intervale $[-\pi, \pi]$ je Fourierov rad funkcie $f(x)$ potom definovaný ako

$$f(x) = a_0 + \sum_{k=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)). \quad (1.1)$$

Fourierov rad môžeme teda rozpísať na jednotlivé členy nasledovne.

$$f(x) = a_0 + a_1 \cos(x) + b_1 \sin(x) + a_2 \cos(2x) + b_2 \sin(2x) + \dots \quad (1.2)$$

Intuitívne si môžeme predstaviť, že akýkoľvek spojitý periodický signál aproximujeme váhovaným súčtom spojitých sínusových a kosínusových funkcií s rôznou uhlovou frekvenciou. Nasledujúci obrázok 1.2 ilustruje prvých 18 týchto funkcií na intervale $[-\pi, \pi]$.



Obrázok 1.2: Ukážka kosínusových C0-C17 a sínusových funkcií S0-S17 s rôznou uhlovou frekvenciou.

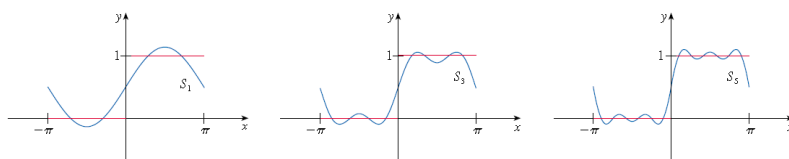
Ak by sme použili **iba kosínusové funkcie**, tak by sme mohli aproximovať **iba párne funkcie**. (Párne funkcie sú funkcie, pre ktoré platí $f(-x) = f(x)$ na intervale $[-\pi, \pi]$, teda sú symetrické okolo y-novej osi podobne ako funkcia kosínus).

Podobne, ak by sme použili **iba sínusové funkcie**, tak by sme mohli aproximovať **iba nepárne funkcie**. (Nepárne funkcie sú funkcie, pre ktoré platí, že: $f(-x) = -f(x)$ na intervale $[-\pi, \pi]$, teda sú symetrické okolo nuly podobne ako sínus.)

Pre aproximáciu ľubovoľného periodického signálu Fourierovým radom teda potrebujeme kombináciu sínusových aj kosínusových funkcií.

Príklad. Obrázok 1.3 ilustruje aproximáciu zvolenej funkcie Fourierovým radom. Konkrétne sme zvolili obdĺžnikovú funkciu (na obr. znázornená červenou). Pre názornosť sú zobrazené iba prvé tri postupné aproximácie pomocou Fourierovho radu. Každá ďalšia postupná aproximácia

by ďalej zmenšila rozdiel medzi obdĺžnikovou funkciou a jej aproximáciou, v limitnom prípade sa chyba aproximácie bude blížiť k nule a počet sínusových funkcií aproximácie sa bude blížiť nekonečnu. Pretože naša obdĺžniková funkcia je nepárna, uplatia sa iba sínusové zložky, kosínusové členy majú nulové váhové koeficienty.



Obrázok 1.3: Aproximácia obdĺžnikovej funkcie $f(x)$ ako Fourierovho radu, čiastočné súčty 1., 3., a 5. člena.

Harmonická analýza. Rozklad periodického signálu na Fourierov rad sa často nazýva aj harmonická analýza. Takéto označenie sa často používa v technickej praxi. Terminológia, ktorú potom používame je nasledovná: nultý koeficient nazývame **nultou harmonickou**. Je to priemerná hodnota signálu, označujeme ju tiež z angličtiny ako **DC komponent**. Prvý koeficient je **základná frekvencia** signálu. Ďalšie koeficienty prislúchajú násobným frekvenciám základnej frekvencie a nazývame ich **druhá harmonická, tretia harmonická** ...

$$f(x) = \underbrace{a_0}_{\text{DC}} + \underbrace{a_1 \cos(x) + b_1 \sin(x)}_{\text{základná harmonická}} + \underbrace{a_2 \cos(x) + b_2 \sin(x)}_{\text{2. harmonická}} + \underbrace{a_3 \cos(x) + b_3 \sin(x)}_{\text{3. harmonická}} \dots \quad (1.3)$$

1.2.2 Spojitá Fourierova transformácia

Spojitá Fourierova transformácia v porovnaní s Fourierovým radom (viď 1.2.1) rozširuje možnosť reprezentácie periodických signálov aj na neperiodické signály spojité v čase, avšak za podmienky, že obsah plochy (v absolútnej hodnote) pod grafom funkcie je konečný

$$\int_{-\infty}^{\infty} |f(x)| dx < \infty. \quad (1.4)$$

Definícia spojitaj Fourierovej transformácie. Fourierovou transformáciou spojitaj funkcie $f(x)$ jednej premennej nazveme funkciu $F(u)$, ktorá je definovaná vzt'ahom

$$F(u) = \int_{-\infty}^{\infty} f(x) e^{-j2\pi ux} dx, \quad (1.5)$$

inverzná spojitá Fourierova transformácia je potom definovaná ako

$$f(x) = \int_{-\infty}^{\infty} F(u) e^{j2\pi ux} du, \quad (1.6)$$

kde

$$j = \sqrt{-1},$$

$$e^{j\theta} = \cos \theta + j \sin \theta,$$

a d'alej platí, že

$$2\pi f = \omega,$$

kde ω je uhlová frekvencia v radiánoch [rad] a f je frekvencia v [Hz].

1.2.3 Dvojdimenziálna Fourierova transformácia

Nech $f(x, y)$ je spojitá integrovateľná funkcia, a nech $F(u, v)$ je integrovateľná. Potom pre výpočet spojitaj dvojdimenziálnej Fourierovej transformácie platí

$$F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-j2\pi(ux+vy)} dx dy, \quad (1.7)$$

a pre výpočet spätnej dvojdimenziálnej Fourierovej transformácie platí vzt'ah

$$f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v) e^{j2\pi(ux+vy)} du dv, \quad (1.8)$$

kde

$$j = \sqrt{-1},$$

$$e^{j\theta} = \cos \theta + j \sin \theta.$$

Ako už bolo spomenuté, $F(u, v)$ je z oboru komplexných čísiel a pre interpretáciu je výhodné použiť nasledovný zápis.

$$F(u) = |F(u)| e^{i\varphi(u)}, \quad (1.9)$$

kde **magnitúdové spektrum** počítame s použitím vzťahu

$$|F(u)| = [R^2(u) + I^2(u)]^{\frac{1}{2}},$$

a **fázové spektrum** počítame podľa nasledovného vzťahu

$$\varphi(u) = \arctan \left[\frac{I(u)}{R(u)} \right].$$

1.2.4 Vlastnosti Fourierovej transformácie

V tomto odstavci si stručne, bez dôkazov, uvedieme najdôležitejšie vlastnosti Fourierovej transformácie:

Linearita. Fourierova transformácia patrí medzi lineárne transformácie, linearita je jej základná vlastnosť.

Majme transformačný pár: funkciu $f_1(t)$ a k nej príslušné spektrum $F_1(u)$. Ďalej majme druhý transformačný pár: funkciu $f_2(t)$ a k nej príslušné spektrum $F_2(u)$.

Takéto páry budeme označovať ako:

$$f_1(t) \Longleftrightarrow F_1(u),$$

$$f_2(t) \Longleftrightarrow F_2(u).$$

Ďalej sú dané konštanty K_1 a K_2 .

Potom platí, že

$$K_1 f_1(t) + K_2 f_2(t) \Longleftrightarrow K_1 F_1(u) + K_2 F_2(u). \quad (1.10)$$

Veta o posunutí v čase. V anglickej literatúre ju nájdeme pod názvom time shifting.

Majme transformačný pár: funkciu $f(t)$ a k nej príslušné spektrum $F(u)$. Keď signál $f(t)$ posunieme v čase o t_0 , tak pre jeho spektrum platí

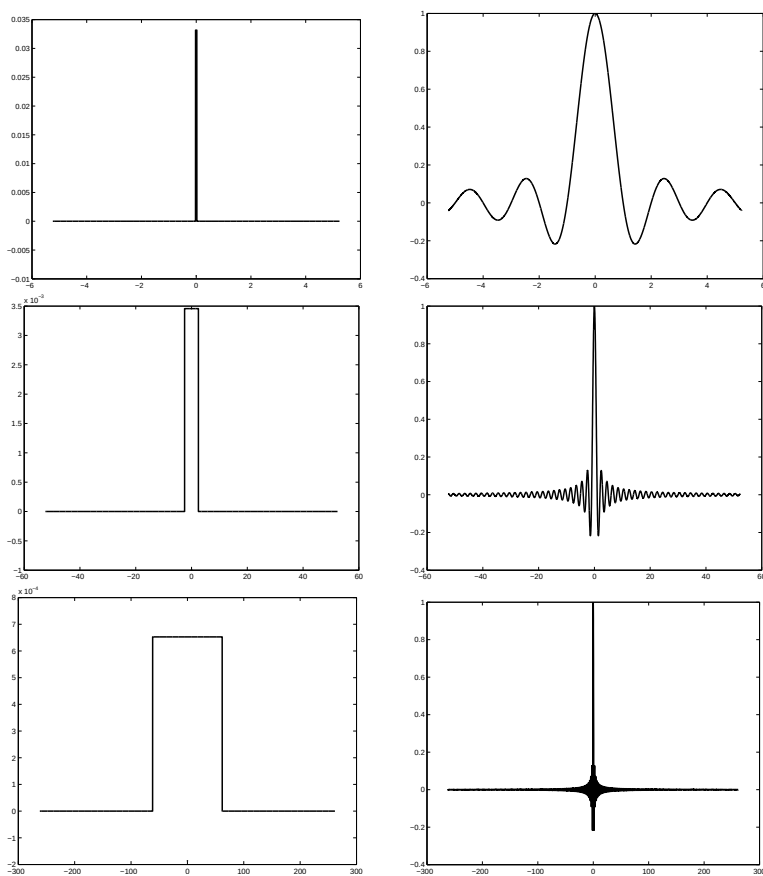
$$f(t - t_0) \Longleftrightarrow e^{-j2\pi u T_0} F(u). \quad (1.11)$$

Na strane spektra po posunutí signálu v čase pribudol multiplikatívny člen, spektrum bude vynásobené komplexnou jednotkou $e^{-j2\pi u T_0}$. Magnitúdové spektrum posunutého signálu zostáva teda nezmenené, zmení sa len fázové spektrum.

Veta o podobnosti. Nazývaná je tiež ako veta o zmene mierky času (time scaling). Majme transformačný pár: funkciu $f(t)$, k nej príslušné spektrum $F(u)$ a konštantu a . Potom platí veta

$$f(at) \Longleftrightarrow \frac{1}{|a|} F(u/a). \quad (1.12)$$

„Roztiahnutie“ signálu v časovej osi sa prejaví ako „zúženie“ signálu jeho spektra, čo je ilustrované na obrázku 1.4



Obrázok 1.4: Ilustrácia vety o podobnosti. Na x-ovej osi je premenná x pre vstupný signál resp. premenná u pre spektrum. Na y-novej osi je amplitúda signálu resp. spektra.

Veta o symetrii. Vyjadruje dualitu signálu a jeho spektrálneho obrazu. Majme transformačný pár: funkciu $f(t)$ a k nej príslušné spektrum $F(u)$. Potom platí

$$f(t) \Longleftrightarrow F(-u). \quad (1.13)$$

Parsevalova teoréma. Parsevalova teoréma hovorí, že energia vstupného signálu $f(t)$ a energia jeho spektra $F(u)$ sa rovnajú

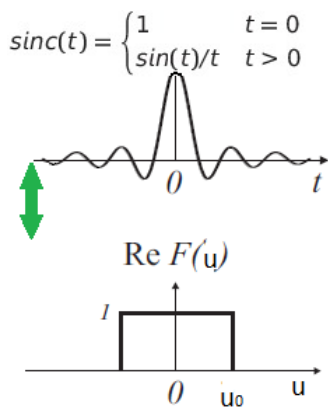
$$\int_{-\infty}^{\infty} |f(t)|^2 dt = \int_{-\infty}^{\infty} |F(u)|^2 du. \quad (1.14)$$

Vybrané transformačné páry funkcií. Niektoré funkcie majú špeciálny význam pre prax, preto si uvedieme tie najdôležitejšie.

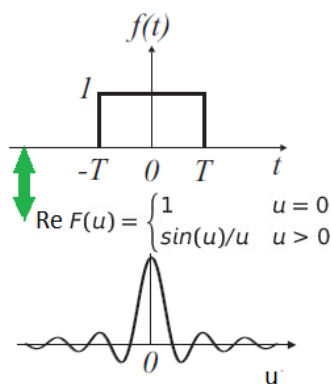
Transformačný pár funkcií $\text{sinc}(t) \Longleftrightarrow \text{obdĺžniková funkcia}$. Definujme funkciu $\text{sinc}(t)$ nasledovne

$$\text{sinc}(t) = \begin{cases} 1 & t = 0 \\ \sin(t)/t & t > 0. \end{cases} \quad (1.15)$$

Spektrum funkcie $\text{sinc}(t)$ bude obdĺžniková funkcia $F(u)$, ako ilustruje obrázok 1.5.



Obrázok 1.5: Funkcia $\text{sinc}(t)$ a jej reálne spektrum.



Obrázok 1.6: Obdĺžniková funkcia a jej reálne spektrum.

Tieto dva páry funkcií a ich spektier sú veľmi dôležité pre objasnenie niektorých javov pri filtrácii obrazu, napr. pri filtrácii tzv. ideálnym filtrom

opísaným v kapitole 1.6. Dôležité je tiež poznamenať, že funkcia $\text{sinc}()$ je definovaná pre definičný obor $[-\infty, \infty]$, a teda nie je fyzikálne realizovateľná.

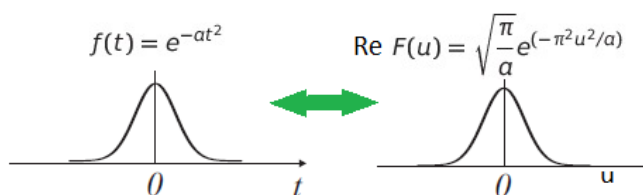
Gaussova funkcia. Pre Gaussovu funkciu definovanú ako

$$f(t) = e^{-a t^2}, \quad (1.16)$$

môžeme vypočítať prostredníctvom Fourierovej transformácie jej Fourierov obraz.

$$F(u) = \sqrt{\frac{\pi}{a}} e^{(-\pi^2 u^2 / a)}. \quad (1.17)$$

Transformačný pár Gaussovej funkcie je ilustrovaný na nasledujúcom obrázku 1.7.



Obrázok 1.7: Gaussova funkcia a jej reálne spektrum.

Fourierovým spektrálnym obrazom Gaussovej funkcie je teda znovu Gaussova funkcia, pričom samozrejme platí veta o podobnosti. „Širšej“ Gaussovej funkcii ako vstupný signál zodpovedá „užšia“ Gaussova funkcia v spektrálnej oblasti.

1.2.5 Diskrétna Fourierova transformácia (DFT)

Ekvivalentom spojitej Fourierovej transformácie pre spojité signály je Diskrétna Fourierova transformácia (DFT) pre diskrétny signály.

Diskrétna Fourierova transformácia (DFT) je na rozdiel od spojitej Fourierovej transformácie definovaná iba pre periodické signály.

Prechod od spojitej k diskkrétnej Fourierovej transformácii si matematicky môžeme odvodiť ak budeme predpokladať, že určitý úsek signálu $x(t)$ signálu bol vzorkovaním prevedený na diskrétne a tento navzorkovaný úsek sa bude periodicky opakovať.

Priama transformácia DFT. DFT je matematicky vyjadrená vzťahom

$$F(u) = \frac{1}{N} \sum_{n=0}^{N-1} f(n) e^{-j2\pi un/N}, \quad (1.18)$$

pre

$$u = 0, 1, 2, \dots, N-1.$$

Spätná transformácia IDFT. Inverzná (spätná) IDFT je definovaná nasledovne.

$$f(n) = \sum_{u=0}^{N-1} F(u) e^{j2\pi un/N}, \quad (1.19)$$

pre

$$n = 0, 1, 2, \dots, N-1.$$

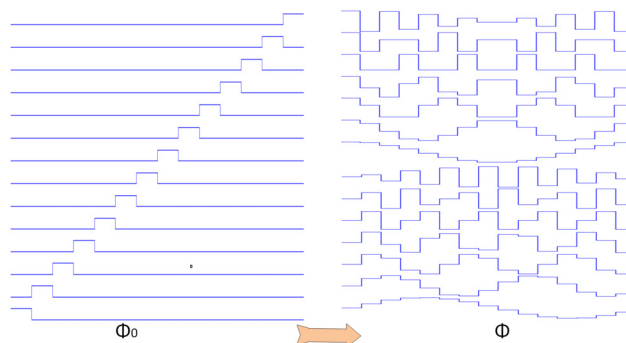
1.2.6 DFT ako špeciálny prípad lineárnej transformácie

Postupnosť vzoriek signálu $Fs(n)$ pre $n = 0, 1, \dots, N-1$, ktorú získame digitalizáciou 1D signálu na konečnom intervale s počtom vzoriek N , si môžeme predstaviť ako jeden vektor v N -rozmernom vektorovom priestore. Princíp lineárnej transformácie môžeme vysvetliť nasledovne.

- Diskrétny signál $Fs(n)$ pre $n = 0, 1, \dots, N-1$, získaný digitalizáciou 1D signálu na konečnom intervale s počtom vzoriek N , môžeme reprezentovať ako jeden vektor v N -rozmernom komplexnom priestore $\mathbb{C}^N$ (viď dodatok C.1) s ortonormálnou bázou Φ_0 (viď dodatok C.4). (U reálnych signálov je imaginárna zložka signálu automaticky nulová.)

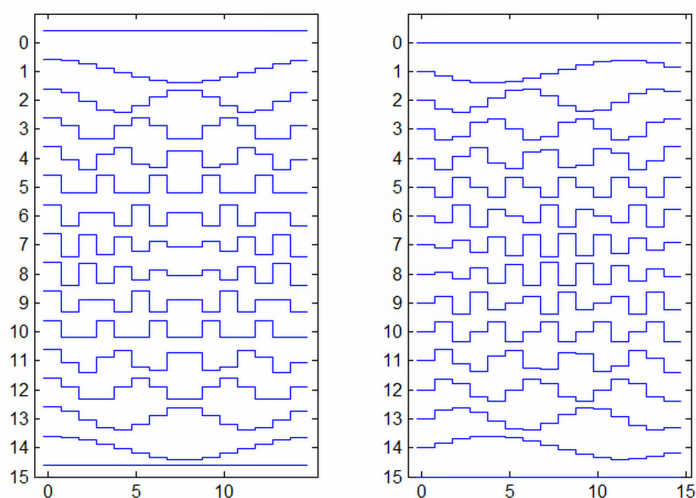
$$Fs(n) \in \mathbb{C}^N. \quad (1.20)$$

- Pomocou lineárnej transformácie (napr. Fourierovej) (viď dodatok C.3) môžeme súradnice každého vektora z N -rozmerného vektorového priestoru $V1$ s ortonormálnou bázou Φ_0 transformovať do iných súradníc v N -rozmernom priestore $V2$ s novou bázou Φ (viď dodatok C.4). V našom prípade $V1 = V2 = \mathbb{C}^N$.
- Späťne môžeme vyjadriť pôvodný číslicový signál ako sumu N vážených básových funkcií bázy Φ (až na násobnú konštantu).
- Tieto váhové koeficienty sú komplexné spektrálne koeficienty získané výpočtom danej lineárnej transformácie.



Obrázok 1.8: Ilustratívny príklad bázových funkcií Φ_0 a Φ .

Diskrétné Fourierove bázové funkcie. Pre názornosť si vizualizujeme 16 kosínusových a 16 sínusových diskretných Fourierových bázových funkcií pre $N = 15$. Definované sú ako súbor sínusových a kosínusových diskretných funkcií s rôznou uhlovou frekvenciou (viď def. v časti 1.2.5).



Obrázok 1.9: Vizualizácia 16 bázových funkcií DFT. Vľavo sú kosínusové funkcie, ktoré tvoria reálnu časť bázových vektorov, vpravo sú sínusové funkcie, ktoré tvoria imaginárnu časť bázových vektorov.

1.2.7 Dvojmerná DFT

Keďže v centre nášho záujmu je **spracovanie digitálneho obrazu**, ktorý je z pohľadu spracovania signálu diskretný dvojdimenzionálny signál, tak potrebujeme definovať **dvojmernú DFT (označovaná tiež ako 2D DFT)**.

Základnú definíciu 2D DFT môžeme zapísať ako

$$F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(ux/M + vy/N)}, \quad (1.21)$$

pre $u = 0, 1, 2, \dots, M-1, v = 0, 1, 2, \dots, N-1$.

Príslušnú inverznú 2D IDFT definujeme ako

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi(ux/M + vy/N)}, \quad (1.22)$$

pre $x = 0, 1, 2, \dots, M-1, y = 0, 1, 2, \dots, N-1$,

kde

- u, v sú súradnice vo frekvenčnej doméne (spektre),
- x, y sú priestorové súradnice v obraze.

1.3 Výpočet diskretnéj Fourierovej transformácie

Pre výpočet DFT máme viaceré možnosti. Uvedieme tie najdôležitejšie.

1.3.1 Výpočet z definície

Najjednoduchší spôsob, ktorý sa nám pre výpočet DFT ponúka, je dosadenie hodnôt vstupného číslicového signálu priamo do definície 1.18 a postupný výpočet jednotlivých spektrálnych koeficientov.

Výpočtová zložitosť. Pri vyhodnotení výpočtovej zložitosti môžeme abstrahovať od určitého počítača a vyhodnotiť iba najhorší prípad výpočtovej zložitosti, či už časovej alebo pamäťovej. Potom používame označenie veľké písmeno O .

Výpočet DFT z definície má **kvadratickú náročnosť výpočtu** $O(n^2)$, čo znamená, že počet krokov algoritmu bude v najhoršom prípade úmerný kvadrátu počtu vzoriek.

Takáto kvadratická zložitosť $O(n^2)$ je samozrejme horšia než napr. lineárna zložitosť $O(n)$ alebo konštantná $O(1)$, ktorá je nezávislá od dĺžky n .

Vzhľadom ku kvadratickej náročnosti výpočtu $O(n^2)$ je výpočet DFT z definície pre aplikácie v rálnom čase prakticky nevhodný. Výpočet DFT z definície používame teda iba výnimočne, prakticky u všetkých softvérových nástrojov nájdeme podporu rýchlejších metód výpočtu.

1.3.2 Výpočet pomocou maticových operácií

Vzťah 1.18 môžeme s pomocou substitúcie $W = e^{2\pi i u/N}$ prepísať do maticového tvaru ako

$$\begin{pmatrix} F_0 \\ F_1 \\ F_2 \\ \vdots \\ F_{(N-1)} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & W^1 & \dots & W^{N-1} \\ 1 & W^2 & \dots & W^{2N-2} \\ \vdots & \vdots & & \vdots \\ 1 & W^{N-1} & \dots & W^{(N-1)^2} \end{pmatrix}}_{\mathbf{W}} \cdot \begin{pmatrix} f_0 \\ f_1 \\ f_2 \\ \vdots \\ f_{(N-1)} \end{pmatrix}. \quad (1.23)$$

Maticový zápis výpočtu DFT skrátene napíšeme ako

$$\mathbf{F} = \mathbf{W} \cdot \mathbf{f}, \quad (1.24)$$

kde

$\mathbf{f}$ je stĺpcový vektor vstupného signálu,

$\mathbf{F}$ je stĺpcový vektor jeho Fourierovho spektra a

$\mathbf{W}$ je Fourierova **transformačná matica**.

Výpočet pomocou uvedeného maticového násobenia má rovnakú výpočtovú náročnosť ako výpočet z definície: $O(n^2)$. Napriek rovnakej teoretickej náročnosti sa dá očakávať, že skutočný potrebný čas výpočtu môže byť kratší, a to v prípade, keď sú v použitom softvérovom nástroji už nachystané optimalizované funkcie pre maticové násobenie.

Maticový zápis pre výpočet 2D DFT. Výpočet dvojrozmernnej 2D DFT zapíšeme pomocou maticových operácií nasledovne.

$$\mathbf{F} = \mathbf{W} \cdot \mathbf{f} \cdot \mathbf{W}^T, \quad (1.25)$$

pričom

$\mathbf{f}$ je maticová reprezentácia dvojdimenzionálneho signálu, v našom prípade obrazu,

$\mathbf{W}^T$ je transponovaná matica $\mathbf{W}$,

$\mathbf{W}$ je transformačná matica Fourierovej transformácie.

1.3.3 Rýchla Fourierova transformácia (FFT)

V anglickej literatúre ju nájdeme pod názvom **Fast Fourier Transform**, skrátené **FFT**.

Dnes existuje značný počet prístupov, ako optimalizovať výpočtovú náročnosť DFT a teda urýchliť jej výpočet. Uvedieme pôvodný Cooleyov-Tukeyov algoritmus nazvaný podľa jeho autorov, publikovaný v roku 1965 [CT65]. Spolu s nástupom špeciálnych čipov pre spracovanie signálu (Digital Signal Processor DSP) v 70. a 80. rokoch 20. storočia priniesol algoritmus FFT hotovú „revolúciu“ v spracovaní signálu a otvoril možnosť spracovania číslicového signálu v spektrálnej oblasti v reálnom čase.

Pokúsime sa vo veľmi skrátenej forme načrtnúť iba jeho podstatu a postup výpočtu, podrobný výklad by bol nad rámec tejto knihy.

Základný koncept FFT. Základný princíp FFT pre vektory s dĺžkou N s ľubovoľným základom (mixed-radix FFT), ktorý predstavili už autori Danielson a Lanczos v roku 1942, je faktorizácia.

Ak N môže byť rozložené na súčin n_f prirodzených čísiel,

$$N = f_1 f_2 \dots f_{n_f}, \quad (1.26)$$

tak výpočet DFT môže byť rozdelený na výpočet viacerých DFT pre každý činiteľ zvlášť. Presnejšie povedané, výpočet DFT s dĺžkou N môže byť rozdelený na

(N/f_1) výpočtov DFT s dĺžkou f_1 ,
 (N/f_2) výpočtov DFT s dĺžkou f_2 ,
 $\dots$
 (N/f_{n_f}) výpočtov DFTs s dĺžkou f_{n_f} .

Celkový počet operácií pre tieto suboperácie potom bude: $O(N(f_1 + f_2 + \dots + f_{n_f}))$. V prípade, že činitele N sú malé celé čísla, tak výpočtová náročnosť bude podstatne menšia ako $O(N^2)$.

Príklad.

Najskôr začneme s výpočtom DFT s celkovou dĺžkou,

$$h_a = \sum_{b=0}^{N-1} g_b W_N^{ab}, \quad (1.27)$$

kde

$$W_N = \exp(-2\pi i / N).$$

Potom rozdelíme túto sumu na párne a nepárne členy podľa ich pozície vo vektore.

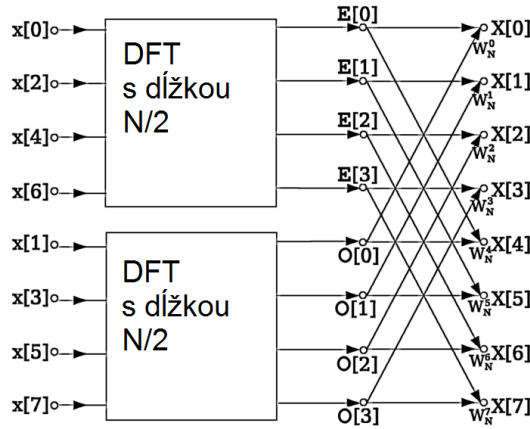
$$h_a = \sum_{b=0}^{N/2-1} g_{2b} W_N^{a(2b)} + \sum_{b=0}^{N/2-1} g_{2b+1} W_N^{a(2b+1)}. \quad (1.28)$$

Týmto konvertujeme DFT s dĺžkou N do dvoch výpočtov DFT s dĺžkou $N/2$,

$$h_a = \sum_{b=0}^{N/2-1} g_{2b} W_{(N/2)}^{ab} + W_N^a \sum_{b=0}^{N/2-1} g_{2b+1} W_{(N/2)}^{ab}. \quad (1.29)$$

Prvý člen predstavuje výpočet DFT elementov na párnych pozíciách g . Druhý člen predstavuje výpočet prvkov na nepárnych pozíciách g , prednásobený exponenciálnym činiteľom W_N^k .

$$\text{DFT}(h) = W_{(N/2)}^{ab} \text{DFT}(g_{\text{párne}}) + W_{(N/2)}^{ab} \text{DFT}(g_{\text{nepárne}}). \quad (1.30)$$



Obrázok 1.10: Rozdelenie DFT s dĺžkou N na dve DFT s dĺžkou $N/2$.

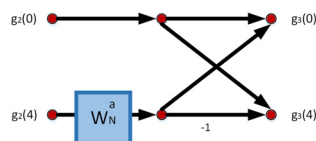
Rozdelením DFT na párnu a nepárnu časť sme zredukovali počet operácií z N^2 (pre výpočet DFT s dĺžkou N) na $2(N/2)^2$ (pre dva výpočty DFT, každý pre vektor s dĺžkou $N/2$).

Môžeme pokračovať v takomto rozdeľovaní ďalej rekurzívne $\log_2 N$ krát (tj. postupujeme k DFT o dĺžkach $N/4, N/8, \dots$ až je dĺžka $= 1$), teda až pokiaľ celý výpočet DFT nie je zredukovaný na výpočet DFT jedného prvku, čo je operácia s nulovou výpočtovou náročnosťou.

Výpočtová náročnosť FFT. Keďže N je mocninou čísla 2, výpočtová náročnosť DFT dĺžky $N = 2^m$ môže byť znížená na $mN/2$ DFT výpočtov s dĺžkou 2, alebo $O(N \log_2 N)$ operácií.

Výpočtová náročnosť v tomto prípade už **nie je kvadratická, ale logaritmická, čiže** $O(N \log_2 N)$.

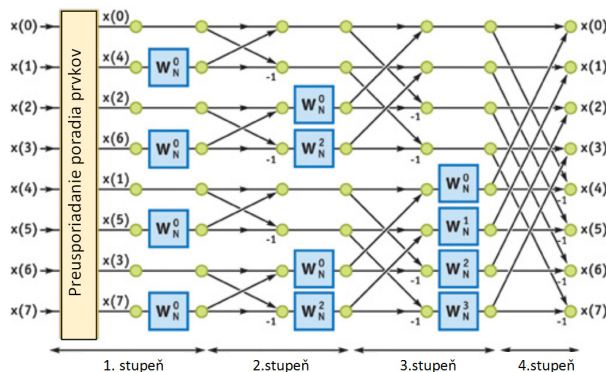
Motýlikový diagram. Je základným prvkom signálového diagramu výpočtu FFT.



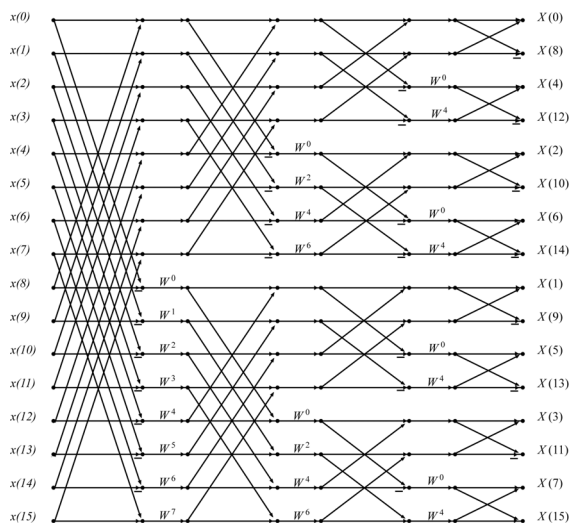
Obrázok 1.11: Motýlikový diagram FFT.

Exponenciálne členy W_N^a sú v anglickej literatúre nazývané tiež ako twiddle factors. Diagram, ktorý znázorňuje výpočet dvoch výstupných prvkov ako kombináciu dvoch vstupných prvkov pomocou sčítania, odčítania a násobenia exponenciálnym členom, sa nazýva motýlikový diagram (vid' obr. 1.11). Je základným prvkom výpočtu FFT.

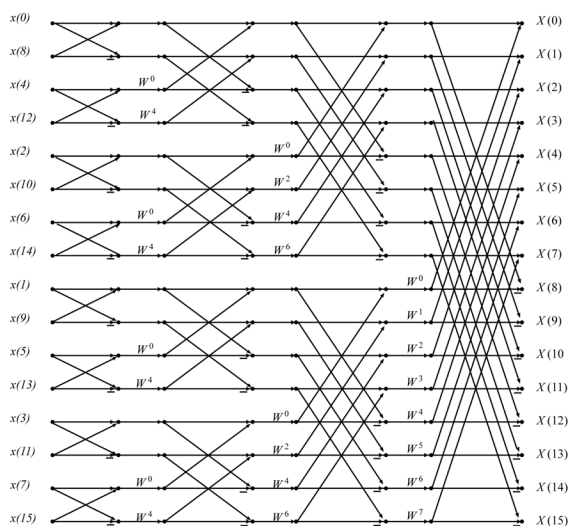
Diagram výpočtu FFT. Nasledujúci diagram (vid' obr. 1.12) znázorňuje signálový diagram: príklad výpočtu FFT pre $N=16$. Výpočet algoritmu prebieha rekurzívne v 4 stupňoch.



Obrázok 1.12: Graf výpočtu FFT pre vektor s dĺžkou 8.



Obrázok 1.13: FFT – decimácia vo frekvencii.



Obrázok 1.14: FFT – decimácia v čase.

Decimácia vo frekvencii a decimácia v čase. V odbornej literatúre nachádzame dva typy algoritmu FFT: decimácia vo frekvencii a decimácia

v čase. Ich zásadný rozdiel je zrejmy z nasledovných znázornených signálových diagramov výpočtu na obrázkoch 1.13 a 1.14.

Bitovo invertované poradie prvkov. Poradie prvkov napíšeme ako bitový reťazec. Jeho bitovo invertované poradové číslo potom získame tak, že tento bitový reťazec interpretujeme opačným smerom zápisu, čiže odzadu smerom dopredu.

V signálovom diagrame FFT je poradie prvkov bitovo invertované, a to na vstupe alebo na výstupe, podľa toho, či sa jedná o FFT typu decimácia v čase alebo decimácia vo frekvencii. V algoritme decimácie vo frekvencii sú prvky preusporiadané bitovou inverziou na vstupe a v algoritme decimácie v čase sú naopak preusporiadané bitovou inverziou na výstupe.

Optimalizácia výpočtu FFT. Algoritmus výpočtu FFT sa nechá paralelizovať, je vhodný na rýchle hardvérové implementácie na špeciálnych čipoch i na výpočet pomocou grafického procesora GPU.

Na záver si pripomeňme, že algoritmus FFT (Cooley–Tukey) je pre vektory s dĺžkou $N = 2^M$.

1.3.4 Doplnenie nulami

Metódu doplnenia nulami nájdeme v anglickej literatúre pod názvom zero padding. Používajú sa dva základné prístupy:

- doplnenie nulami v spektrálnej oblasti,
- doplnenie nulami v priestorovej oblasti.

Doplnenie nulami v spektrálnej oblasti môžeme použiť pre zvýšenie rozlíšenia vstupného obrazu. Novozískané hodnoty sú vlastne interpolované existujúce hodnoty obrazu, neobsahujú novú informáciu.

Doplnenie nulami v priestorovej oblasti síce tiež podobne zvýši rozlíšenie, v tomto prípade rozlíšenie v spektrálnej oblasti, ale najčastejší dôvod použitia je ten, že štandardne používané algoritmy výpočtu FFT (napr. algoritmus Cooley–Tukey popísaný v odstavci 1.3.3) predpokladajú vstupný obraz, ktorého veľkosť je práve mocnina čísla 2. V prípade, že náš vstupný obraz túto podmienku nespĺňa, čo je v praxi veľmi pravdepodobné, tak výhodne môžeme použiť metódu doplnenia nulami. Postupujeme jednoducho.

- Vytvoríme si obraz požadovanej veľkosti (napr. s rozmermi, ktoré môžeme vyjadriť ako mocninu čísla 2). Tento je samozrejme väčší než náš vstupný obraz.

- Skopírujeme doňho pôvodný vstupný obraz a chýbajúce hodnoty doplníme nulami.

Takto vytvorený obraz potom použijeme ako vstup pre FFT.

Doplnením nulami zväčšíme vstupný obraz a tiež zvýšime jeho rozlíšenie v spektrálnej oblasti. Nezískame však žiadnu novú, detailnejšiu informáciu, jemnejšie rozlíšenie v spektre je vlastne výsledkom interpolácie. Musíme si tiež byť vedomí skutočnosti, že týmto postupom vzniknú v signáli obrazu body nespojitosti, ktoré môžu negatívne ovplyvniť očakávané výsledky spektrálneho spracovania.

1.4 Príklady Fourierovej transformácie obrazu

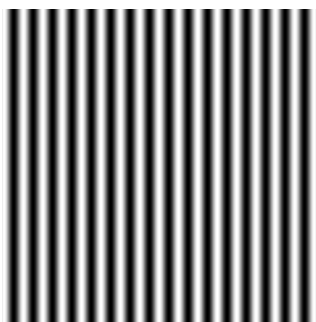
Uvedieme si niekoľko príkladov transformačných párov pozostávajúcich z príkladu testovacieho obrazu a jeho magnitúdového spektra. Cieľom je budovanie prvej intuitívnej predstavy ako správne interpretovať informáciu obsiahnutú v spektre obrazu.

Na obrázku 1.15 je testovací obraz, ktorý je generovaný ako periodická textúra vertikálneho smeru. Jeho magnitúdové spektrum na obrázku 1.16 je i napriek logaritmickému merítku skoro celé čierne. Pozorný čitateľ si všimne v tomto spektre dve maximá horizontálne symetrické poblíž stredu. Táto poloha indikuje, že ide o vertikálne periodické štruktúry s nízkou frekvenciou, čo odpovedá vstupnému obrazu. Pre lepšiu názornosť je na obrázku 1.17 znázornený horizontálny rez týmto obrazom kde sú dve symetrické maximá zreteľne viditeľné. Obrázky 1.18 znázorňujú podobný príklad, ibaže periodická štruktúra obrazu je vodorovná a maximá v spektre sú vertikálne symetrické.

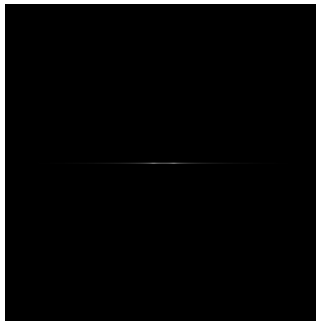
Príklady na obrázkoch 1.20 až 1.24 sú analogické k predchádzajúcim, ibaže frekvencia periodickej štruktúry vo vstupných obrazoch (horizontálnom i vertikálnom) je vyššia. Dôležité je všimnúť si, že poloha maxim v oboch prípadoch sa posunula od stredu smerom ku krajom, keďže ide o vyššiu frekvenciu než v predchádzajúcom prípade.

Na vizualizácii vstupných obrazov na obrázkoch 1.15 a 1.23 je viditeľný jav, ktorý nastal porušením Nyquistovho kritéria pri vizualizácii. Tento jav podvzorkovania nazývame ako tzv. Moire efekt. Pre výpočet však bol použitý vstupný obraz s vysokým rozlíšením, ktorý nebol zaťažený chybou podvzorkovania ako obrazy vizualizované.

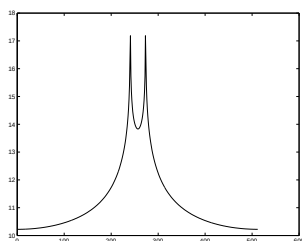
Na obrázkoch 1.27 až 1.30 je v spektre viditeľný vplyv funkcie $\text{sinc}()$.



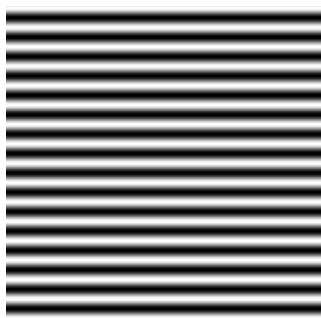
Obrázok 1.15: Nízko-frekvenčnou vertikálnou periodickou textúrou.



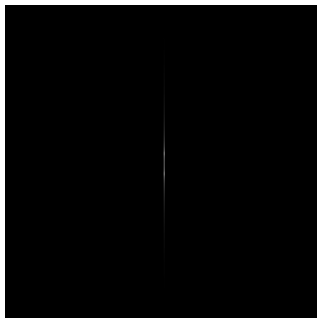
Obrázok 1.16: Magnitúdové spektrum v logaritmickú škále.



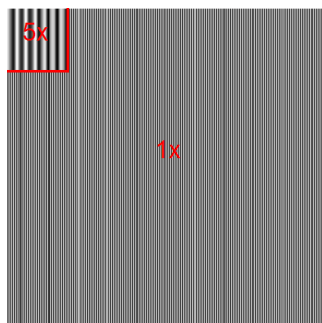
Obrázok 1.17: Rez magnitúdovým spektrom v logaritmickú škále.



Obrázok 1.18: Obraz s nízko-frekvenčnou horizontálnou periodickou textúrou.



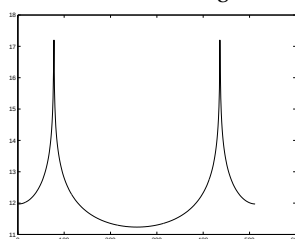
Obrázok 1.19: Magnitúdové spektrum v logaritmickú škále.



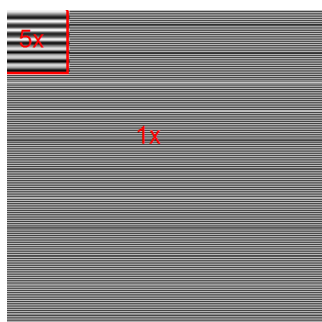
Obrázok 1.20: Príklad vysoko-frekvenčnej vertikálnej periodickej textúry so zväčšeným detailom.



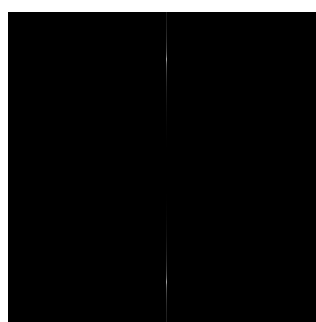
Obrázok 1.21: Magnitúdové spektrum vysokofrekvenčnej vertikálnej periodickej textúry v logaritmickú škále.



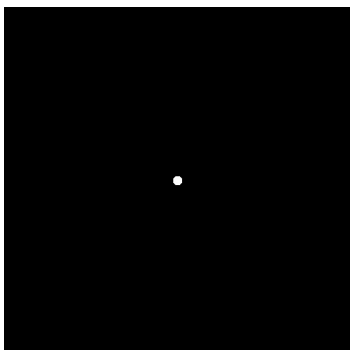
Obrázok 1.22: Rez magnitúdovým spektrom v logaritmickú škále.



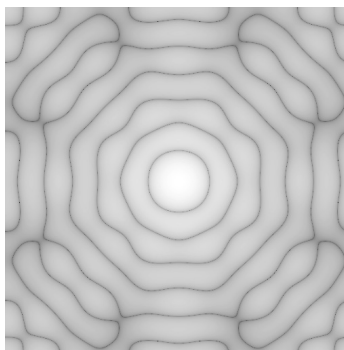
Obrázok 1.23: Príklad vysoko-frekvenčnej horizontálnej periodickej textúry so zväčšeným detailom.



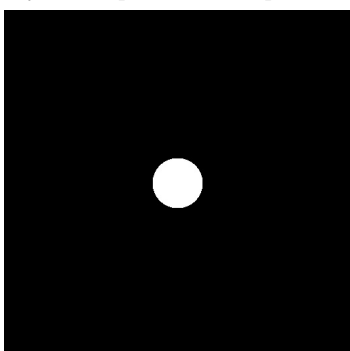
Obrázok 1.24: Magnitúdové spektrum vysokofrekvenčnej horizontálnej periodickej textúry v logaritmickú škále.



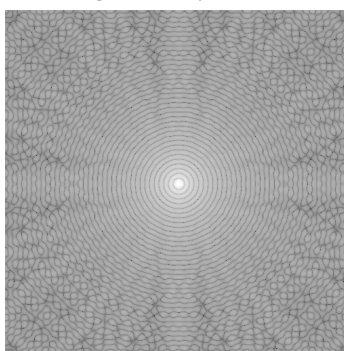
Obrázok 1.25: Obraz s kruhovým objektom s priemerom 15 pixelov.



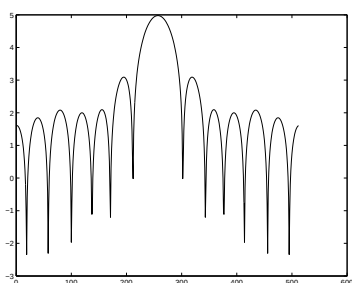
Obrázok 1.26: Magnitúdové spektrum v logaritmickej škále.



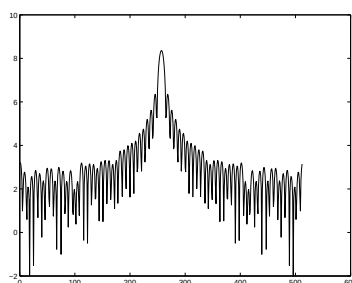
Obrázok 1.27: Obraz s kruhovým objektom s priemerom 75 pixelov.



Obrázok 1.28: Magnitúdové spektrum v logaritmickej škále.



Obrázok 1.29: Rez magnitúdovým spektrom z obr. 1.26 v logaritmickej škále.



Obrázok 1.30: Rez magnitúdovým spektrom z obr. 1.28 v logaritmickej škále.

Na uvedených príkladoch teda 1.15 až 1.24 vidíme, že magnitúdové spektrum je vždy symetrické, odpovedajúcu odozvu v spektre nachádzame na transponovanej pozícii. Frekvencia zobrazená v spektre je v strede nulová a ku kraju obrazu sa frekvencia zvyšuje. Najvyššia frekvencia sa rovná polovici vzorkovacej frekvencie.

1.5 Kosínusová transformácia

1.5.1 Diskrétna kosínusová transformácia

Bázové funkcie príslušné Diskrétnej kosínusovej transformácie (DCT) sú, na rozdiel od Fourierovej transformácie, definované iba pomocou kosínusovej funkcie, a to tak, aby **báza bola úplná**. To znamená, že **každý vektor** reálnych čísel o rovnakej dimenzii ako dimenzia danej bázy DCT, **je možné vyjadriť ako lineárnu kombináciu báзовých vektorov DCT**. Koeficienty, ktoré vypočítame prostredníctvom DCT sú na rozdiel od DFT iba **z oboru reálnych čísel**.

V praxi sa stretávame s viacerými definíciami kosínusovej transformácie, označenými spravidla ako DCT-I, DCT-II, DCT-III, DCT-IV. Všetky definujú úplnú bázu pomocou kosínusových funkcií. V spracovaní obrazu sa používa predovšetkým DCT-II definovaná nasledovne.

$$F(u, v) = \frac{2C(u)C(v)}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos\left(\frac{2x+1}{2N}u\pi\right) \cos\left(\frac{2y+1}{2N}v\pi\right), \quad (1.31)$$

kde

$$C(k) = \begin{cases} \frac{1}{\sqrt{2}} & k = 0, \\ 1 & \text{inak.} \end{cases}$$

DCT-II môžeme definovať pomocou maticového zápisu nasledovne

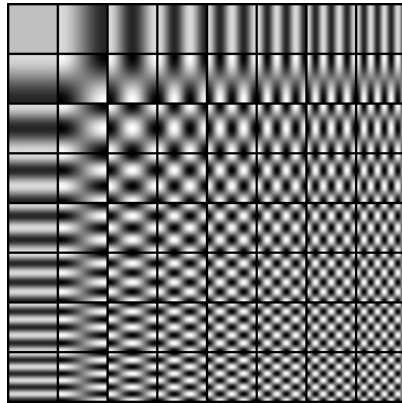
$$F = C_{NN} f C_{NN}^T, \quad (1.32)$$

a inverzná DCT-II je definovaná

$$f = C_{NN}^T F C_{NN}, \quad (1.33)$$

kde

$$C_{NN}(k, l) = \begin{cases} \frac{1}{\sqrt{N}} & l = 0, \\ \sqrt{\frac{2}{N}} \cos\left(a \frac{2k+1}{2N} l \pi\right) & \text{inak.} \end{cases}$$



Obrázok 1.31: Vizualizácia bazových funkcií v tvare matíc veľkosti 8×8 pre 2-dimenzionálnu DCT-II.

Veľmi názorné je zobrazenie dvojdimenzionálnych bazových funkcií diskretnej kosínusovej transformácie (viď obr. 1.31).

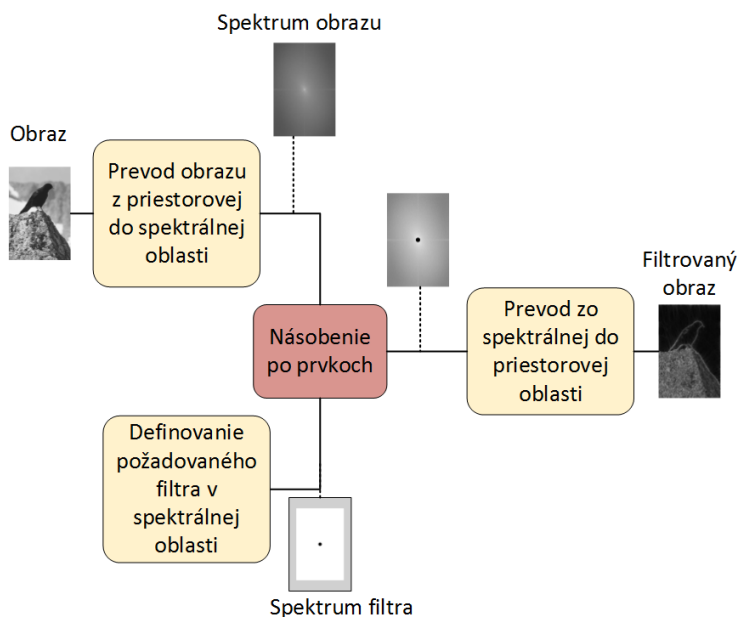
Použitie diskretnej kosínusovej transformácie. Diskrétna kosínusová transformácia (DCT) má pre fotorealistické obrazy veľmi dobré dekorelačné vlastnosti a preto sa používa predovšetkým na kompresiu dát. Konkrétne napr. v norme pre kompresiu obrazu JPEG sa využíva DCT na transformovanie neprekrývajúcich sa regiónov obrazu o veľkosti 8×8 pixelov, pričom získané spektrálne koeficienty sú zaokrúhľované podľa ich relevancie pre ľudské vizuálne vnímanie.

1.6 Filtrácia obrazu vo frekvenčnej oblasti

Filtrácia obrazu patrí medzi základné nástroje pre úpravu a vylepšenie kvality číslicového obrazu.

Základná operácia filtrácie vo frekvenčnej oblasti je operácia **násobenia**, presnejšie povedané, v prípade obrazu použijeme násobenie dvoch matíc po prvkoch.

Na nasledujúcom obrázku 1.32 je schématicky znázornený základný koncept filtrácie vo frekvenčnej doméne.



Obrázok 1.32: Filtrácia v spektrálnej oblasti.

Prvá matica je spektrum vstupného obrazu, ktoré získame výpočtom dvojdimenzionálnej diskretnej Fourierovej transformácie zo vstupného obrazu. Druhá matica je spektrum filtra, ktorý chceme pre filtráciu použiť. Spektrum požadovaného filtra je nazývané tiež ako prenosová funkcia filtra. Spektrum požadovaného filtra môžeme navrhnúť priamo vo frekvenčnej oblasti, alebo si ho môžeme odvodiť z filtra daného v priestorovej oblasti. Obidve matice, spektrum obrazu i spektrum filtra, musia mať samozrejme rovnakú veľkosť.

Filter typu dolná a horná priepust’. Celý postup si ukážeme na príklade. Najskôr si však povieme, aké budú naše intuitívne očakávania. Ktoré zmeny filtrácia prinesie?

Máme vstupný obraz, v našom príklade je na obraze vták na skale. Skala je pomerne texturovaná. Príspevok takejto textúry, jemných hrán ako i veľmi malých prvkov v obraze, bude v spektre predovšetkým vo vyšších frekvenciách. Nízke frekvencie sú spektrálnym obrazom väčších plôch v obraze.

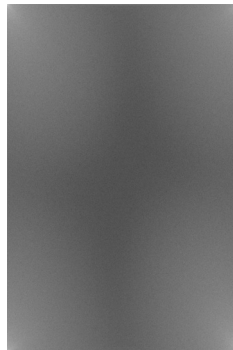
Ak chceme odstrániť drobné detaily z obrazu, ako napr. textúra na skale, tak **odfiltrujeme vyššie frekvencie**. Tým získame vyhladený a rozmazaný obraz bez jemných detailov. Na odfiltrovanie vysokých frekvencií a zachovanie nízkych frekvencií je vhodný filter, ktorý nazývame **dolná priepusť (low pass – LP filter)**.

Naopak, ak chceme získať iba obraz s hranami a textúrou a zachované malé detaily, tak použijeme opačný filter, **odfiltrujeme nízke frekvencie**, použijeme tzv. **hornú priepusť (high pass – HP filter)**.

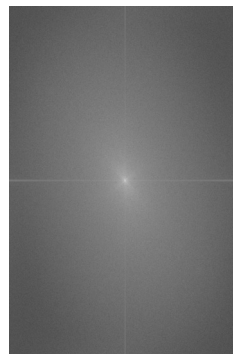
Výpočet Fourierovho spektra vstupného obrazu. Najskôr transformujeme vstupný obraz do spektrálnej oblasti. Aby boli hodnoty spektra lepšie viditeľné, použijeme zaužívanú konvenciu pre vizualizáciu magnitudového spektra v logaritmickú mierku.



Obrázok 1.33: Vstupný šedotónový obraz – vták na skale.



Obrázok 1.34: Magnitudové spektrum (log škálované).



Obrázok 1.35: Posunutú magnitudové spektrum (log škálované).

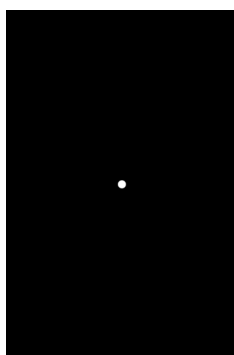
Magnitudové spektrum po výpočte dostávame tak, ako ho vidíme na obrázku 1.34. Nízke frekvencie sú v rohoch obrazu a v strede spektra sú vysoké frekvencie, čo nezodpovedá zvyklostiam zobrazovania spektra. Preto v ďalšom kroku prehodíme jednotlivé kvadranty Fourierovho spektra tak, aby nízke frekvencie boli v strede obrazu spektra. Na centrálnej pozícii v strede spektra je potom DC komponent, ktorý zodpovedá priemernej hodnote jasu obrazu. DC komponent je spravidla dobre viditeľný. Od stredu spektra ku krajom sú zobrazené stále vyššie frekvencie.

Poznámka: Z teoretického pohľadu, uvedené preorganizovanie kvadrantov zodpovedá takému DFT spektru, ktoré by sme vypočítali pre vstup-

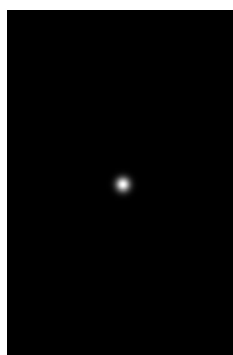
ný signál na intervale $x \in \langle -M/2, M/2 - 1 \rangle$, $y \in \langle -N/2, N/2 - 1 \rangle$, kde x, y sú priestorové súradnice a M, N sú rozmery spracovávaného obrazu, keďže vstupný signál interpretujeme ako periodický signál, ktorý získame opakovaním spracovávaného vstupného signálu.

Dolnopriepustná filtrácia. V nasledujúcom príklade si ukážeme dolnopriepustnú filtráciu, a to s použitím dvoch rôznych LP filtrov.

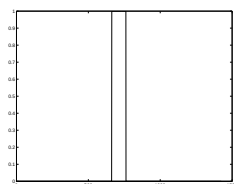
- ideálny LP filter
- Gaussov LP filter



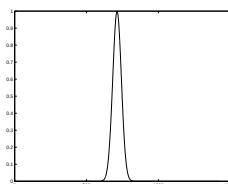
Obrázok 1.36: Ideálny LP filter vo frekvenčnej oblasti.



Obrázok 1.37: Gaussov LP filter vo frekvenčnej oblasti.



Obrázok 1.38: Rez ideálnym LP filtrom.



Obrázok 1.39: Rez Gaussovým LP filtrom.

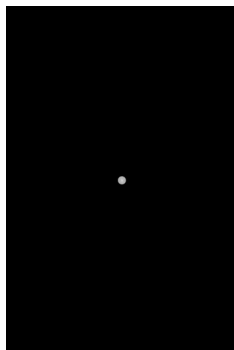
Hodnoty filtra sa pohybujú v rozsahu $\langle 0, 1 \rangle$ reálnych čísiel, operáciu násobenia teda musíme tiež počítať v obore reálnych čísiel. Nulové hodnoty filtra majú za následok úplné odstránenie príslušného spektrálneho

komponentu, jednotková hodnota naopak zachová po vynásobení príslušnú spektrálnu zložku v pôvodnej hodnote. Ostatné hodnoty z intervalu $(0, 1)$ lineárne váhujú príslušnú spektrálnu zložku.

Ideálny filter. Ideálny filter má teoreticky nekonečne strmý prechod medzi priepustnou a nepriepustnou časťou v spektre filtra, čiže nežiadúce frekvenčné komponenty môžu byť selektívne úplne vynulované. Aj keď z pohľadu návrhu sa javí skutočne ako ideálny (odtiaľ názov), a to práve pre možnosť veľmi selektívnej filtrácie bez prechodových oblastí, tak v skutočnosti má jednu nepríjemnú vlastnosť. Táto vlastnosť spôsobuje efekt rozvlnenia (v anglickej literatúre označovaný ako ringing efekt) a jej príčinou je jav, ktorý nazývame Gibbsov jav. Vysvetlíme ho v odstavci 1.6.

Gaussov filter. Na rozdiel od ideálneho filtra, Gaussov filter má plynulý prechod medzi priepustnou a nepriepustnou časťou spektra, pričom hodnoty filtra sú dané dvojrozmernou Gaussovou funkciou.

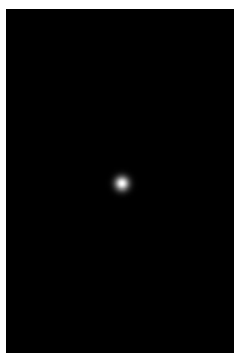
Násobenie v spektre a spätná Fourierova transformácia. Na nasledujúcom obrázku vľavo je vizualizované spektrum po vynásobení ideálnym filtrom, ktorý použijeme pre spätnú 2D DFT a vpravo vidíme už výsledok po spätnej 2D DFT filtrovaného spektra. Filtrovaný obraz ideálnym filtrom (vid' obrázok 1.41) je podľa očakávania rozmazaný, detaily textúr na skale boli odfiltrované. Je tiež viditeľný silne rušivý efekt rozvlnenia. Naproti tomu obraz filtrovaný Gaussovým LP filtrom (vid' obrázok 1.43) je tiež rozmazaný ako sme predpokladali, nadôvažok nie je na ňom viditeľný žiaden rušivý efekt.



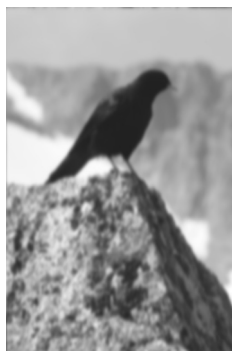
Obrázok 1.40:
Spektrum filtrované
ideálnym LP filtrom.



Obrázok 1.41: Obráz
po filtrácii ideálnym LP
filtrom.



Obrázok 1.42:
Spektrum filtrované
Gaussovým LP filtrom.



Obrázok 1.43: Obráz
po filtrácii Gaussovým
LP filtrom.

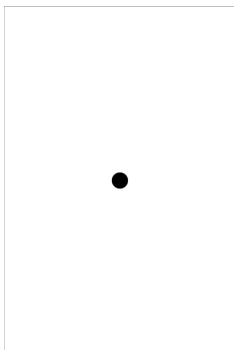
Hornopriepustná filtrácia. Ďalej si ukážeme filtráciu hornopriepustným filtrom.

Pre tvorbu HP filtra vo frekvenčnej oblasti môžeme použiť jednoduchý výpočet

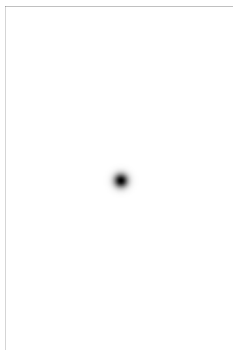
$$f_{HPfilter}(u, v) = 1, 0 - f_{LPfilter}(u, v).$$

Jeho cieľom je prepustiť práve tie frekvenčné zložky, ktoré LP filter zadržal a naopak.

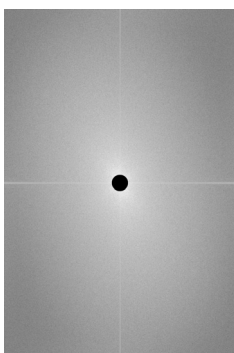
1.6. Filtrácia obrazu vo frekvenčnej oblasti



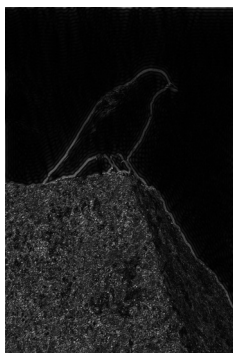
Obrázok 1.44: Ideálny HP filter vo frekvenčnej oblasti.



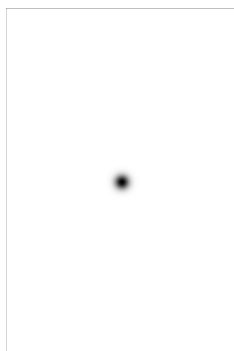
Obrázok 1.45: Gaussov HP filter vo frekvenčnej oblasti.



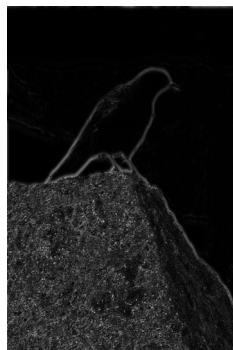
Obrázok 1.46: Spektrum filtrované ideálnym HP filtrom.



Obrázok 1.47: Obraz po filtrácii ideálnym HP filtrom.



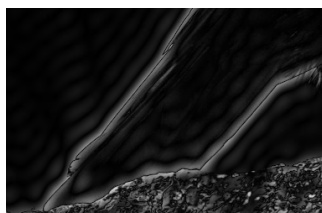
Obrázok 1.48: Spektrum filtrované Gaussovým HP filtrom.



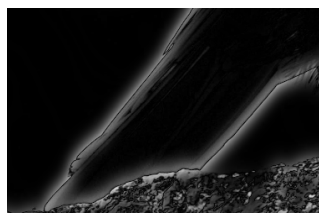
Obrázok 1.49: Obráz po filtrácii Gaussovým HP filtrom.

Obráz, ktorý sme získali po filtrácii hornopriepustným filtrom obsahuje iba vysokofrekvenčné zložky obrazu, je zreteľne vidieť textúra na skale a kontúry vtáka. (Výsledný obraz je zobrazený v absolútnej hodnote pre lepšiu vizualizáciu výsledného hranového obrazu.) Opäť pozorujeme rušivý efekt rozvlnenia v obraz filtrovaný ideálnym filtrom (viď detail na obrázku 1.50) na rozdiel od Gaussovho, kde nevidíme žiaden rušivý efekt.

Efekt rozvlnenia u ideálneho filtra. V odstavci 1.2.4 sme si ukázali, že obdĺžniková funkcia má spektrum dané funkciou $\text{sinc}()$ a naopak. Ideálny filter má charakter obdĺžnikovej funkcie, jeho reprezentácia v časovej doméne alebo v priestorovej doméne (obraz) je teda práve funkcia $\text{sinc}()$. V kapitole 3 sa dozvieme, že podľa konvolučnej teóremy je násobeniu v spektrálnej oblasti ekvivalentná operácia konvolúcie v časovej alebo priestorovej oblasti.



Obrázok 1.50: Detail z obr. 1.47.



Obrázok 1.51: Detail z obr. 1.49.

Vysvetlenie pre efekt rozvlnenia (ringing efekt) je teda nasledovné: násobenie ideálnym filtrom v spektre je ekvivalentné konvolúcii obrazu s funkciou typu $\text{sinc}()$, ktorá zanechá efekt rozvlnenia. Na obrázkoch 1.50 a 1.51 je v detaile tento efekt dobre pozorovateľný.

1.6.1 Filtrácia vybraných frekvenčných oblastí

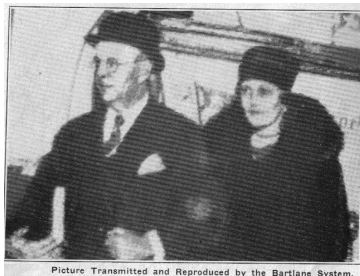
Filtrácia typu dolná alebo horná priepuť filtruje vysoké, resp. nízke frekvencie z obrazového signálu. Môžeme však filtrovať i podľa iných kritérií. Filtrácia vo frekvenčnej oblasti sa nám ponúka ako výborný nástroj, pokiaľ vieme v spektre loklizovať príspevok toho signálu, ktorý chceme odstrániť.

Uvedieme príklad.

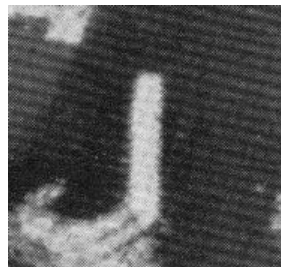
Ako vstupný obraz použijeme jednu z prvých snímok, ktoré boli v digitalizovanej forme prenesené cez Atlantik v 20. rokoch minulého storočia. (Vid' obr. 1.52.) Na ďalšom obrázku 1.54 je znázornené spektrum, v ktorom vidíme niekoľko lokálnych maxim. Tieto selektívne odstránime.

Vybraté lokálne maximá takpovediac „vymažeme“ zo spektra násobením špeciálnym filtrom – viacnásobne použijeme Gaussov filter centrovanej na príslušné maximum. Takto filtrované spektrum vidíme na obrázku 1.55. Po spätnej filtrácii získame obraz znázornený na obrázku 1.56.

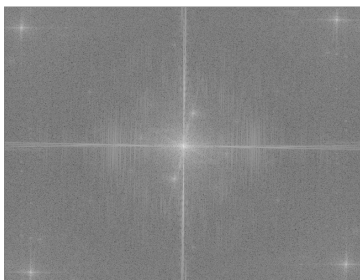
Pri pohľade na vstupný a výstupný obraz vidíme, že periodické rušenie bolo touto filtráciou do značnej miery odstránené.



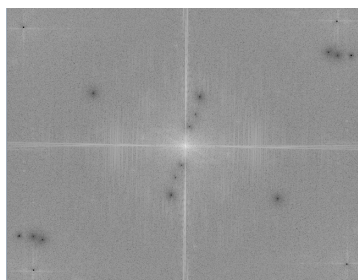
Obrázok 1.52: Vstupný obraz.



Obrázok 1.53: Detail.



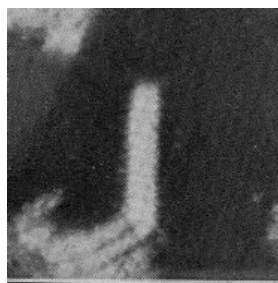
Obrázok 1.54: Spektrum vstupného obrazu.



Obrázok 1.55: Selektívne filtrované spektrum.



Obrázok 1.56: Výstupný obraz.



Obrázok 1.57: Detail.

V tejto kapitole sme si ukázali základný koncept filtrácie v spektrálnej oblasti. Na konkrétnom príklade je filtrácia ilustrovaná s použitím najzákladnejších typov filtrov, s ideálnym filtrom a Gaussovým filtrom. Každý z nich je ukázaný ako LP a HP filter. Taktiež sme ukázali príklad filtrácie vo frekvenčnej oblasti cieľovým odfiltrovaním niekoľkých vybraných oblastí v spektre.

V ďalšej kapitole sa budeme venovať inému modelu filtrácie – filtrácii priamo v priestorovej doméne.



Spracovanie obrazu v priestorovej oblasti

Pre lineárnu filtráciu obrazu môžeme použiť manipuláciu hodnôt obrazovej funkcie priamo v priestorovej doméne. Hovoríme potom o lineárnej filtrácii v priestorovej oblasti. Základný matematický nástroj pre lineárnu filtráciu je operácia **konvolúcia**, ktorej sa budeme venovať v ďalších odstavcoch.

2.1 Diskrétna konvolúcia

Definíciu spojitaj konvolúcie neuvádzame. (V prípade záujmu odkážeme čitateľa na výbornú literatúru [OS75].) Keďže pracujeme s diskretnými funkciami, ktoré používame pre reprezentáciu číslicového signálu, tak pre ich spracovanie potrebujeme konvolúciu v diskretnéj forme.

Diskrétnu konvolúciu dvoch kauzálnych diskretných funkcií definujeme nasledovne.

Sú dané funkcie $f = (f_0, \dots, f_{N-1})^T \in \mathbb{C}^N$, $h = (h_0, \dots, h_{M-1})^T \in \mathbb{C}^M$ potom ich **diskrétna konvolúcia** je vektor $g \in \mathbb{C}^{M+N-1}$ pre členy ktorého platí

$$g(x) = \sum_{m=0}^{N-1} f(m)h(x-m), \quad (2.1)$$

pre $x = 0, \dots, M+N-2$.

Pre spracovanie obrazu je potrebná konvolúcia definovaná pre funkcie dvoch premenných, ktorú si vysvetlíme podrobnejšie v ďalšom odstavci.

2.1.1 Diskrétna 2D konvolúcia

Diskrétnu konvolúciu o dvoch premenných zapíšeme nasledovne

$$g(x, y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n) h(x - m, y - n). \quad (2.2)$$

Pre skrátený zápis operácie konvolúcie používame znak „hviezdička“

$$f(x, y) * h(x, y). \quad (2.3)$$

Funkcia $f(x, y)$ je funkcia obrazu, ktorý ideme filtrovať, $h(x, y)$ je tzv. **filtračné jadro**, nazývané tiež **filtračná maska** alebo **konvolučná maska**.

2.1.2 Vlastnosti konvolúcie

Komutatívnosť. Operácia konvolúcia je komutatívna, čo môžeme vyjadriť nasledovný vzťahom

$$f(x) * h(x) = h(x) * f(x). \quad (2.4)$$

Vzťah medzi konvolúciou a koreláciou. Pojem lineárnej filtrácie obrazu v priestore sa v literatúre používa niekedy tiež pre výpočet filtrácie prostredníctvom korelácie (značená $\circ$). Matematicky môžeme vzťah medzi koreláciou a konvolúciou vyjadriť nasledovne

$$f(x) * h(-x) = f(x) \circ h(x). \quad (2.5)$$

Pri lineárnej filtrácii obrazu to znamená, že použitá filtračná maska pre výpočet konvolúciou je transponovaná matica k matici, ktorú by sme použili na rovnakú filtráciu obrazu prostredníctvom korelácie.

2.2 Výpočet konvolúcie

Pre výpočet konvolúcie používame postup výpočtu daný definíciou.

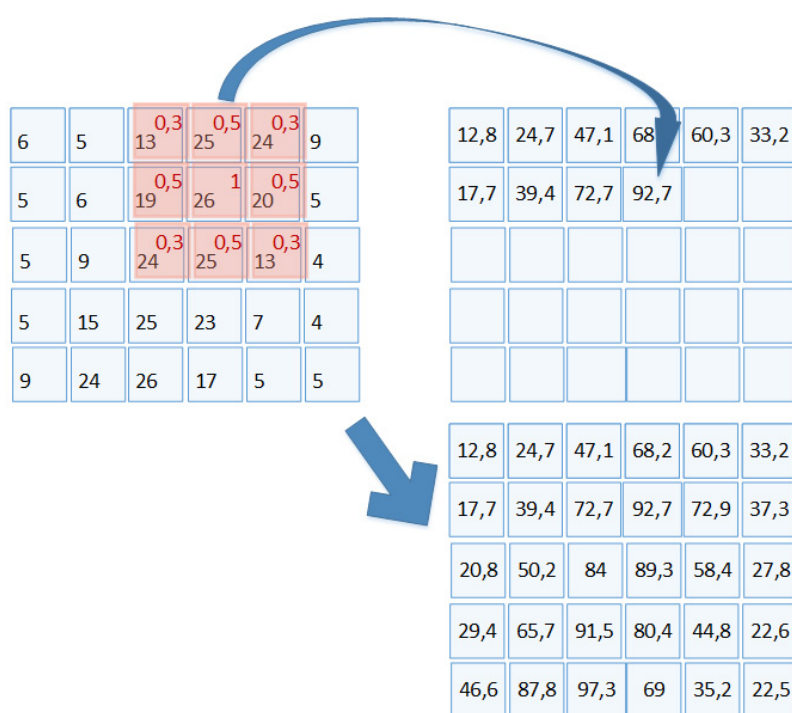
Implementácia operácie konvolúcie, ako kľúčovej funkcie spracovania obrazu, je však už desaťročia v centre záujmu expertov a dnes máme k dispozícii vysoko optimalizované implementácie, napr. pre Intel procesory v knižnici Intel Integrated Performance Primitives (IPP), ktorú používajú Matlab, OpenCV a ďalšie softvérové nástroje pre spracovanie obrazu.

Filtračné jadro spravidla definujeme tak, aby jeho **veľkosť bola nepárna**. Napr. 3×3 , 5×5 ... 15×15 ... V tomto prípade je stredová pozícia

pixela, na ktorú zapíšeme výslednú hodnotu, jednoznačne daná. Ak zvolíme inú pozíciu, tak vstupný a výstupný obraz budú voči sebe, okrem zmeny danej filtráciou, aj vzájomne posunuté. (Môže ísť samozrejme aj o zámer.)

2.2.1 Znázornenie výpočtu konvolúcie

Príklad. Náorne si ukážeme výpočet jedného prvku konvolvovaného obrazu (viď 2.1). Máme príklad matice obrazu znázornený čiernou a konvolučné jadro veľkosti 3×3 znázornené červenou farbou.



Obrázok 2.1: Príklad výpočtu konvolúcie. $f_{(2,4)} = 13 \cdot 0,3 + 25 \cdot 0,5 + 24 \cdot 0,3 + 19 \cdot 0,5 + 26 \cdot 1 + 20 \cdot 0,5 + 24 \cdot 0,3 + 25 \cdot 0,5 + 13 \cdot 0,3 = 92,7$

Samotný postup výpočtu je daný definíciou 2.1. Jednoducho povedané, pre každú posunutú pozíciu konvolučného jadra počítame sumu súčinov hodnoty konvolučného jadra a príslušného pixelu obrazu.

Okraj obrazu. Je zrejmé, že výpočet hodnôt konvolúcie na okraji obrazu musíme riešiť zvláštnym spôsobom. V praxi sa používa niekoľko stratégií.

- Doplnenie chýbajúcich hodnôt pre výpočet konvolúcie nulovými hodnotami, alebo inou konštantou.
- Zrkadlenie. Chýbajúce hodnoty sa doplnia zrkadlovo rovnaké ku hodnotám pixelov na kraji obrazu.
- Zmenšenie výstupného obrazu. Tento postup je z programátorského hľadiska najmenej vhodný, hlavne pri viacnásobnom sekvencnom použití filtrácie. Veľkosť obrazu je potom rôzna v rôznych štádiách spracovania, čo musíme zvlášť zohľadniť pri výpočte.

Spravidla použijeme uvedený prvý alebo druhý spôsob, avšak musíme si byť vedomí, že okraje filtrovaného obrazu (podľa veľkosti konvolučného jadra) sú pre exaktné vyhodnotenie teoreticky nevhodné. V praxi to zväčša problém nespôsobuje, keďže veľkosť filtra sa volí pokiaľ možno malá i z dôvodu výpočtovej náročnosti.

Zisk filtra. Pokiaľ nie je našim zámerom obraz pri filtrovaní konvolúciou zároveň zosvetliť alebo stmaviť, tak zisk filtra (angl. gain) musí byť rovný jednej. Filter v tomto prípade navrhujeme tak, aby súčet všetkých hodnôt konvolučného jadra bol rovný jednej.

2.3 Najdôležitejšie konvolučné masky

Filtráciu konvolúciou používame predovšetkým vtedy, keď potrebujeme obraz vyhladiť, rozmazať hrany, odstrániť šum alebo ak chceme získať hranový obraz. Ako už vieme, v prvom prípade realizujeme filter dolnopásmová priepusť, v druhom prípade naopak filter hornopásmová priepusť. Dôležitý je samozrejme vhodný návrh konvolučnej masky. Konvolučnú masku si môžeme aj sami navrhnuť alebo odvodiť, v tejto kapitole ďalej uvedieme príklady najznámejších, často používaných konvolučných masiek.

2.3.1 Vyhladzovacie filtre

Medzi najznámejšie vyhladzovacie filtre patria priemerovací filter a Gaussov filter.

Priemerovací filter. Priemerovací filter má členy filtračného jadra definované tak, aby výsledkom konvolúcie bol vždy aritmetický priemer okolia bodu daného veľkosťou použitého filtračného jadra.

Príklad filtračného jadra priemerovacieho filtra veľkosti 7×7 je

$$\frac{1}{49} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}. \quad (2.6)$$



Obrázok 2.2: Vstupný obraz – detail nohy vtáka.



Obrázok 2.3: Príklad obrazu filtrovaného priemerovacím filtrom.

Na obrázkoch 2.2 a 2.3 je príklad filtrácie uvedeným priemerovacím filtrom veľkosti 7×7 .

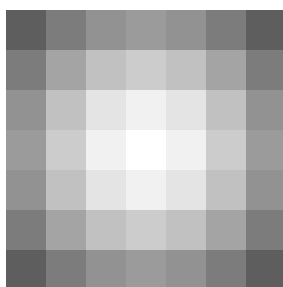
Gaussov filter. Gaussov filter má filtračné jadro odvodené od Gaussovej dvojrozmernej funkcie definované nasledovne

$$h(m, n) = \frac{1}{\sum_m^M \sum_n^N e^{\frac{-(m^2+n^2)}{2\sigma^2}}} e^{\frac{-(m^2+n^2)}{2\sigma^2}}. \quad (2.7)$$

Prvý člen v tomto vzťahu je váhovací faktor tak, aby zisk celého filtra bol rovný jednej.

Ukážme si príklad Gaussovho filtra veľkosti 7 a $\sigma = 3$.

$$h = \begin{bmatrix} 0.0113 & 0.0149 & 0.0176 & 0.0186 & 0.0176 & 0.0149 & 0.0113 \\ 0.0149 & 0.0197 & 0.0233 & 0.0246 & 0.0233 & 0.0197 & 0.0149 \\ 0.0176 & 0.0233 & 0.0275 & 0.0290 & 0.0275 & 0.0233 & 0.0176 \\ 0.0186 & 0.0246 & 0.0290 & 0.0307 & 0.0290 & 0.0246 & 0.0186 \\ 0.0176 & 0.0233 & 0.0275 & 0.0290 & 0.0275 & 0.0233 & 0.0176 \\ 0.0149 & 0.0197 & 0.0233 & 0.0246 & 0.0233 & 0.0197 & 0.0149 \\ 0.0113 & 0.0149 & 0.0176 & 0.0186 & 0.0176 & 0.0149 & 0.0113 \end{bmatrix}.$$



Obrázok 2.4: Gaussov filter h veľkosti 7×7 , $\sigma = 3$.



Obrázok 2.5: Vstupný obraz – detail nohy vtáka.



Obrázok 2.6: Príklad obrazu filtrovaného Gaussovým filtrom.

Na obrázkoch 2.5 a 2.6 je príklad filtrácie uvedeným Gaussovým filtrom.

2.3.2 Hranové filtre

Ďalším, veľmi často používaným typom filtra sú hranové filtre.

Prewittov filter. Prewittov filter pozostáva z dvoch konvolučných jadier: vertikálneho a horizontálneho. Ich použitím pre konvolúciu získame zvlášť vertikálny a zvlášť horizontálny príspevok k hrane v danom bode. Tieto dve informácie môžeme použiť pre výpočet gradientu. Pod gradientom funkcie $f(x, y)$ rozumieme dvojrozmerný vektor, ktorého súradnice sú parciálne diferencie funkcie f podľa x a y . Môžeme tiež vypočítať uhol, ktorý tento vektor zvierá, podľa vzťahu 2.15.

Konvolučná maska vertikálneho Prewittovho filtra je daná v tvare

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}. \quad (2.8)$$

Konvolučná maska horizontálneho Prewittovho filtra je daná ako

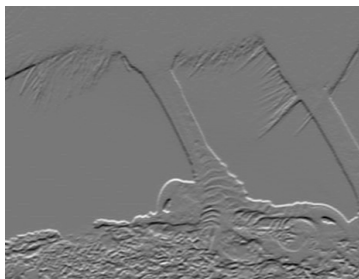
$$\begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}. \quad (2.9)$$

Z vypočítaných vektorov môžeme vyjadriť v každom pixeli obrazu magnitúdu a fázu.

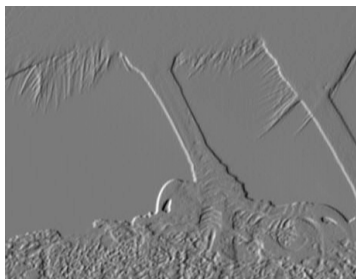
$$\mathbf{G} = \sqrt{\mathbf{G}_x^2 + \mathbf{G}_y^2}, \quad (2.10)$$

$$\Theta = \arctan(\mathbf{G}_y, \mathbf{G}_x). \quad (2.11)$$

Zisk filtra je u tohto filtra nulový. Je však dôležité si uvedomiť, že pri programovaní operácie filtrácie s použitím Prewittovho filtra potrebujeme pre výstupný filtrovaný vertikálny a horizontálny obraz dátový typ, ktorý obsahuje aj záporné čísla (napr. integer alebo float). Vypočítané diferencie môžu byť totižto kladné alebo záporné podľa toho, či odpovedajúca hrana vo vstupnom obraze je nábežná alebo dobežná. Vizualizované sú preškálované obrazy, kde nulová hodnota z hranového obrazu bola posunutá na úroveň 127 z rozsahu 0 – 225. (Z tohoto dôvodu sú vizualizované obrazy šedé.)



Obrázok 2.7: Obráz filterovaný horizontálnym Prewittovým filtrom.



Obrázok 2.8: Obráz filterovaný vertikálnym Prewittovým filtrom.

Prewittov filter je veľmi podobný Sobelovmu filtru uvedenému v ďalšom odstavci, vizuálne je rozdiel medzi obrazom filterovaným Prewittovým alebo Sobelovým filtrom sotva rozpoznateľný. Obidva tieto typy filtrov sa používajú pre výpočet gradientu.

Sobelov filter. Sobelov filter patrí medzi najpopulárnejšie hranové filtre. Konvolučná maska vertikálneho Sobelovho filtra je definovaná v tvare

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}. \quad (2.12)$$

Analogicky môžeme vyjadriť konvolučnú masku horizontálneho Sobelovho filtra ako

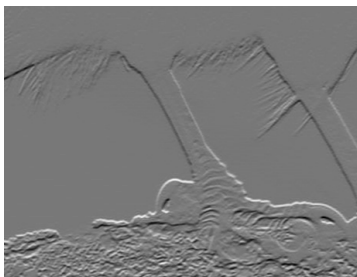
$$\begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}. \quad (2.13)$$

Ak chceme získať kompletný hranový obraz, tak použijeme výsledný vzťah na výpočet magnitúdy z predchádzajúcej x -ovej a y -ovej hodnoty.

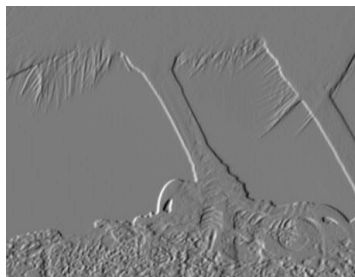
$$\mathbf{G} = \sqrt{\mathbf{G}_x^2 + \mathbf{G}_y^2}, \quad (2.14)$$

$$\Theta = \arctan(\mathbf{G}_y, \mathbf{G}_x). \quad (2.15)$$

Na obrázkoch 2.9 a 2.10 je príklad filtrácie horizontálnym Sobelovým filtrom a vertikálnym Sobelovým filtrom.



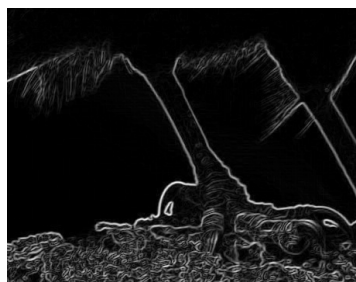
Obrázok 2.9: Obráz filterovaný horizontálnym Sobelovým filtrom.



Obrázok 2.10: Obráz filterovaný vertikálnym Sobelovým filtrom.



Obrázok 2.11: Vstupný obraz – detail nohy vtáka.



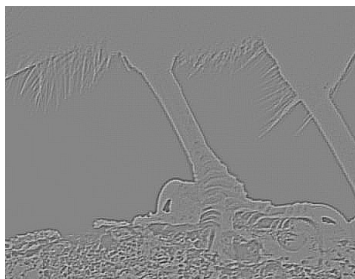
Obrázok 2.12: Magnitúda obrazu filterovaného Sobelovým filtrom.

Laplacian. Diskrétny Laplaceov operátor je často používaný pri spracovaní obrazu a detekcii pohybu. Diskrétny Laplacian je definovaný ako suma druhej diferencie v danom pixeli vzhľadom na okolie pixela. Často používaná definícia Laplaceovho operátora veľkosti 3×3 je

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}. \quad (2.16)$$



Obrázok 2.13: Vstupný obraz – detail nohy vtáka.



Obrázok 2.14: Obráz z obrázku 2.13 filtrovaný Laplaceovým filtrom.

Na obrázku 2.13 je detail obrazu vtáka. Laplaceov operátor, ktorý sme použili pri filtrácii Laplacianom tohoto obrázku bol definovaný ako

$$\begin{bmatrix} 0,5 & 0 & 0,5 \\ 0 & -2 & 0 \\ 0,5 & 0 & 0,5 \end{bmatrix}. \quad (2.17)$$

Na obrázku 2.14 vidíme hranový obraz po filtrácii Laplacianom.



Obrázok 2.15: Absolútna hodnota hranového obrazu (detail nohy vtáka) filtrovaného Laplaceovým filtrom.

Na obrázku 2.15 môžeme pozorovať dvojité hrany, ktoré sú pôvodne nábežnou a dobežnou hranou príslušnej kontúry.

2.4 Použitie konvolúcie

Celkovo môžeme konštatovať, že konvolúcia je veľmi účinný nástroj pre spracovanie obrazu. Podľa cieľa použitia môže odstraňovať alebo zvý-

razňovať detaily, vypočítať hrany, prípadne počítať mieru podobnosti medzi daným filtrom a obrazom.

2.5 Nelineárna filtrácia

Medzi nelineárne filtre patria predovšetkým zorad'ovacie filtre, kde špeciálnym prípadom je filter maxima, minima alebo medián.

2.5.1 Zorad'ovací filter

V anglickej literatúre sa označujú názvom order-statistic filter.

Na rozdiel od lineárnej filtrácie nepočítame konvolúciu. Zorad'ovacie filtre nám výpočtovo pripomínajú skôr morfológickú filtráciu. Obraz postupne prechádzame filtračnou maskou, ktorá obsahuje iba jednotky a nuly. Ich význam je podobný ako ich poznáme u plochého štruktúrneho elementu – jednotky definujú pozície pixelov, ktoré sa aktívne podieľajú na výpočte.

U zorad'ovacieho filtra zoradíme podľa veľkosti všetky prvky dané maskou a vyberieme výstupný prvok, ktorý je v poradí prvý, alebo posledný, alebo stredný – podľa toho, aký zorad'ovací filter počítame – filter minima, maxima alebo mediánový filter. Pozíciu výstupného prvku môžeme tiež definovať podľa potreby na ľubovoľnej inej pozícii v zoradenej postupnosti.

Mediánový filter. Mediánový filter vyberie zo zoradenej postupnosti danej zorad'ovacím filtrom ten prvok, ktorý je na strednej pozícii. Je výborným nástrojom pre odstránenie šumu z obrazu, predovšetkým impulzového šumu alebo šumu typu soľ a korenie (salt and pepper).

Obraz rušený šumom typu soľ a korenie obsahuje jednotlivé pixely šumu, ktoré sú rozmiestnené rôzne v priestore obrazu a ktoré majú bielu alebo čiernu farbu. (Šum typu soľ a korenie môže mať i iné dve úrovne jasů, kombinácia bielej a čiernej je najčastejšia.)



Obrázok 2.16: Vstupný zašumený obraz.



Obrázok 2.17: Mediánovo filtrovaný obraz.

Na obrázku 2.16 je nekvalitný vstupný obraz veľkosti 192×228 pixelov, ktorý obsahuje šum typu soľ a korenie. Na obrázku 2.17 je tento obraz po filtrácii mediánovým filtrom o veľkosti 5×5 .

Konvolučná teoréma

V predchádzajúcich kapitolách sme si najskôr objasnili princíp lineárnej filtrácie obrazu v spektrálnej doméne a potom lineárnej filtrácie v priestorovej doméne. V tejto kapitole, ktorá je síce malá svojím rozsahom strán, ale veľmi významná svojím obsahom, si povieme vzájomný súvis uvedených dvoch prístupov.

Konvolučná teoréma dáva do **priamej vzájomnej relácie lineárnu filtráciu v spektrálnej oblasti a lineárnu filtráciu v priestorovej oblasti** a to nasledovne.

Konvolučná teoréma. Konvolučná teoréma nám hovorí, že **operácia násobenia v spektrálnej doméne je ekvivalentná operácii konvolúcie v priestorovej doméne** a tiež naopak, **operácia konvolúcie v spektrálnej doméne je ekvivalentná operácii násobení v priestorovej doméne**.

Matematicky tento vzťah ekvivalencie zapíšeme nasledovne.

Máme funkciu $f(x, y)$ a jej Fourierove spektrum $F(u, v)$, ďalej máme funkciu $h(x, y)$ a jej Fourierove spektrum $H(u, v)$. Potom platí

$$f(x, y) * h(x, y) \Leftrightarrow F(u, v) H(u, v), \quad (3.1)$$

$$f(x, y) h(x, y) \Leftrightarrow F(u, v) * H(u, v). \quad (3.2)$$

V praxi to znamená, že sa môžeme rozhodnúť pre jednu alebo druhú alternatívu lineárnej filtrácie. Podľa autorov Gonzales a Woods ([GW08]) je filtrácia vo frekvenčnej oblasti intuitívnejšia. Najdôležitejšie faktory,

ktoré môžu ovplyvniť naše rozhodnutie je rýchlosť výpočtu a veľkosť filtračného jadra. Pre efektívny výpočet je spravidla výhodnejšie počítať filtráciu vo frekvenčnej oblasti prostredníctvom FFT. Ďalší dôležitý fakt je, že filtrácia vo frekvenčnej oblasti zachováva veľkosť obrazu.

Konvolučná teoréma nám tiež umožňuje odvodiť postup nazývaný dekonvolúcia.

Morfologické operácie

Za svoj názov vďaka matematickej morfológia pravdepodobne biologickej morfológii, ktorá sa podľa encyklopédie Britannica venuje všeobecným aspektom biologických foriem a usporiadania častí tvarov rastlín a zvierat. Pojmom morfin (morphing) sa označujú animačné techniky počítačovej grafiky, pri ktorých sa zobrazený objekt plynule zmení na iný daný objekt, ktorý môže byť vizuálne i značne odlišný.

4.1 Matematická morfológia

Vo všeobecnosti môžeme o matematickej morfológii v spracovaní obrazu skonštatovať, že sa zaoberá štúdiom objektov z pohľadu ich tvaru, geometrie a topológie, pričom používa kombinácie metód z teórie množín, topológie a diskretnej matematiky. Matematická morfológia teda poskytuje výkonný prístup k spracovaniu obrazu a ďalších diskretných dát.

4.1.1 Morfologické spracovanie obrazu

Morfologické spracovanie obrazu využíva informáciu o susedných pixeloch v topologickom okolí spracovávaného pixela.

Morfologické operácie boli primárne definované a využívané pre binárne obrazy. Nepopierateľný význam však majú i morfologické operácie definované pre šedotónové obrazy a môžu byť aplikované i pre viackanálové (napr. farebné obrazy). Morfologické operácie, ktoré sú defino-

vané pre dvojdimenzionálne dáta, môžeme tiež rozšíriť i na trojdimenzionálne dáta.

4.2 Základné pojmy a definície

Najskôr si zavedieme základné pojmy potrebné pre vysvetlenie samotných morfologických operácií.

4.2.1 Susednosť, súvislosť

Zavedenie niektorých základných pojmov si vyžaduje definície vybraných metrík, ktorými meriame vzdialenosti medzi dvomi bodmi obrazu.

Metrika. Pre obrazové body p, q so súradnicami (x_1, y_1) a (x_2, y_2) definujeme D ako metriku ak

$$D(p, q) \geq 0, \quad (4.1)$$

$$D(p, q) = 0 \text{ vtedy a práve iba vtedy ak } p = q,$$

$$D(p, q) = D(q, p), \quad (4.2)$$

$$D(p, z) \leq D(p, q) + D(q, z). \quad (4.3)$$

Euklidovská vzdialenosť. Euklidovská vzdialenosť medzi bodmi p a q v obraze je definovaná vzťahom

$$D_E(p, q) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}. \quad (4.4)$$

Vzdialenosť v mestských blokoch „City block distance“. Vzdialenosť medzi dvomi bodmi v obraze je v tomto prípade daná počtom krokov (jednotkovej veľkosti) po pravouhlej diskretnéj mriežke

$$D_4(p, q) = |x_1 - x_2| + |y_1 - y_2|. \quad (4.5)$$

Samotný názov „vzdialenosť v mestských blokoch“ evokuje predstavu pohybu v meste.

Šachovnicová vzdialenosť „Chessboard distance“. V prípade, že pravouhlú mriežku rozšírime o diagonálne prepojenia medzi susednými bodmi (viď obr. 4.1), tak môžeme definovať šachovnicovú vzdialenosť nasledovne

$$D_8(p, q) = \max(|x_1 - x_2|, |y_1 - y_2|). \quad (4.6)$$

4-susednosť a 8-susednosť. Ďalším dôležitým pojmom je 4-susednosť a 8-susednosť. 4-susedia daného obrazového elementu sú body s jednotkovou vzdialenosťou v metrike „vzdialenosť v mestských blokoch“. Podobne 8-susedia daného bodu sú body s jednotkovou vzdialenosťou, avšak v šachovnicovej vzdialenosti.

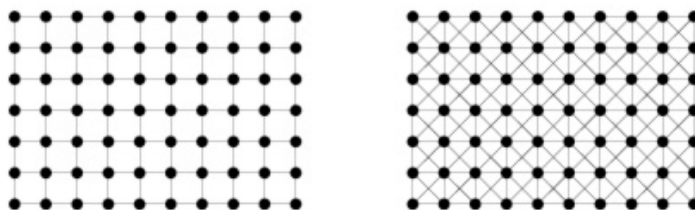
4-súvislosť, 4-spojité komponent, 8-súvislosť, 8-spojité komponent.

4-spojité resp. 8-spojité komponent definujeme nasledovne.

Množina čiernych pixelov F je 4-spojité resp. 8-spojité komponent, ak pre každú dvojicu obrazových bodov p_i a p_j v F existuje postupnosť pixelov $p_i, \dots, p_j$ tak, že

- všetky obrazové body v postupnosti sú v množine F , t. j. sú čierne na obr. 4.1,
- každé 2 pixely, ktoré sú vedľa seba v postupnosti, spĺňajú podmienku 4-susednosti, resp. 8-susednosti.

Čiže daná oblasť je 4-spojité resp. 8-spojité komponent, ak medzi jej ľubovoľnými dvomi vnútornými bodmi existuje cesta zo spojnic spĺňajúcich podmienku 4-susednosti, resp. 8-susednosti.



Obrázok 4.1: Ukážka mriežky 4-súvislosti a 8-súvislosti.

4.2.2 Minkowského suma

Matematickým základom pre binárnu matematickú morfológickú operáciu dilatácie je Minkowského suma.

Minkowského suma bodových množín A a B je bodová množina

$$C = \bigcup_{b \in B} A_b, \quad (4.7)$$

kde

A_b je množina A posunutá o vektor b , teda množina $A_b = a + b \mid a \in A$.

$\bigcup$ označuje zjednotenie (union) množín.

Minkowského sumu množín A a B označujeme

$$C = A \oplus B. \quad (4.8)$$

4.2.3 Štruktúrálly element

V definíciách morfológických operácií je obsiahnutý prvok, ktorý sa nazýva **štruktúrálly element**. V horeuvedenej Minkowského sume je vstupný obraz označený A a štruktúrálly element je tu označený ako B . V ďalšom výklade už budeme označovať štruktúrálly element ako S .

Štruktúrálly element je charakteristický svojím tvarom, ktorý je daný usporiadaním pixelov s hodnotou 1. Typicky sa používajú štruktúrálne elementy štvorcové, elipsové, krížikové alebo ľubovoľne definovaného usporiadania v závislosti od konkrétneho použitia. Dôležitým parametrom je i poloha výsledného pixela. Tento bod je tiež nazývaný počiatok alebo kotevný bod (origin, anchor).

Štruktúrálly element je teda bodová množina – obraz s veľkosťou menšou než vstupný obraz. Jeho typický rozmer býva pomerne malý, napr. 3×3 , 5×5 , ..., 21×21 atď. Rovný (flat) štruktúrálly element obsahuje výlučne iba hodnoty 0 a 1, jednotky v ňom definujú okolie (neighbourhood) kotevného bodu.

Pre intuitívnu predstavu zjednodušene môžeme konštatovať, že hodnota 1 označuje vlastne tie pixely v štruktúrálnom elemente, ktoré sa „aktívne“ podieľajú na stanovení novej hodnoty v počiatku (v kotevnom bode).

4.2.4 Morfológická transformácia

Morfológickú transformáciu môžeme definovať ako matematickú reláciu medzi dvomi bodovými množinami, a to množinou obrazu a množinou štruktúrálneho elementu.

4.3 Základné morfológické operácie

Aplikovanie morfológickej transformácie na obraz **F** zodpovedá výpočtu transformácie pri systematickom posune štruktúrneho elementu **S** po obraze. Výsledok transformácie v každej polohe zodpovedá danej morfológickej relácii.

Pre výpočet základných morfológických operácií sa teda používa mechanizmus podobný ako pri filtrácii, keď každý pixel obrazu je nahradený svojou novou hodnotou vypočítanou daným algoritmom s využitím hodnôt jeho topologických susedov.

Binárna dilatácia zväčšuje plochu bieleho objektu na čiernom pozadí a binárna erózia naopak znižuje plochu bieleho objektu na čiernom pozadí. V prípade čierneho objektu na bielom pozadí by bola situácia analogicky opačná.

Tento výklad vychádza z konvencie, že pixel s hodnotou 1 je biely a pixel s hodnotou 0 je čierny. Predpokladáme tiež, že binárne objekty sú biele objekty na čiernom pozadí.

Pre intuitívnu predstavu, s určitou dávkou nepresnosti, môžeme povedať, že väčší štruktúrny element má za následok väčšie zmeny hraničných oblastí, čiže i plocha objektu sa viac zväčší. Podľa tvaru štruktúrneho elementu môžeme usúdiť, aký tvar budú mať detailné zmeny na okrajoch objektu. Štruktúrny element hranatého tvaru (štvorec, obdĺžnik...) sa prejaví po dilatovaní kontúrami objektu, v ktorých je hranatý tvar štruktúrneho elementu poznateľný. Naopak, u elipsovitého štruktúrneho elementu budeme očakávať nové kontúry skôr zaoblené.

Výhodne môžeme tiež využiť nasledovnú stratégiu výpočtu. Rovnakú morfológickú operáciu s nezmeneným štruktúrnym elementom opakujeme vo viacerých iteráciách za sebou. V softvérových podporách výpočtu morfológických operácií (knižnice pre spracovanie obrazu) spravidla nachádzame tento postup už implementačne pripravený.

Medzi základné morfológické operácie patria

- dilatácia (dilation) 4.4
- erózia (erosion) 4.4.2

a ich základné kombinácie pod názvami

- morfológické otvorenie (opening) 4.5
- morfológické uzavretie (closing) 4.5

4.4 Binárna dilatácia a erózia

4.4.1 Dilatácia

Operáciu binárnej dilatácie spracovávaného obrazu F a štruktúrneho elementu S značíme

$$F \oplus S. \quad (4.9)$$

Binárna dilatácia vyplýva z uvedenej definície Minkowského sumy zjednotenie (union) posunutých bodových množín F a S . Binárnu dilatáciu môžeme teda zapísať ako

$$F \oplus S = \bigcup_{s \in S} F_s, \quad (4.10)$$

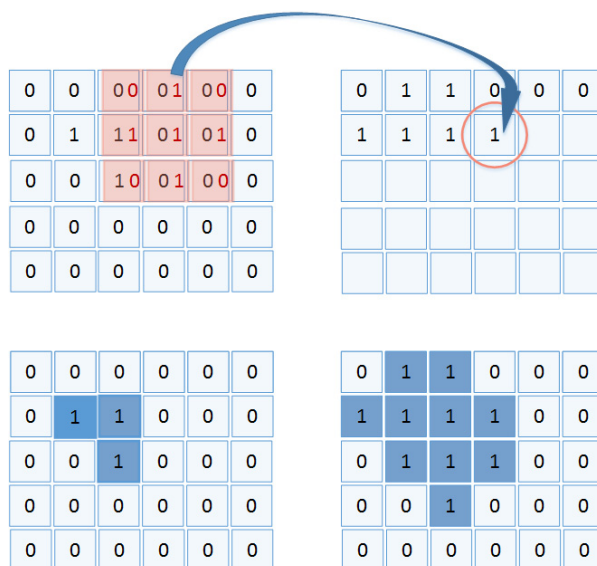
kde F_s je množina F posunutá o vektor s , teda množina

$$F_s = \{f + s \mid f \in F\}. \quad (4.11)$$

Ako bolo už spomenuté, morfológickým dilatovaním zväčšíme plochu bielej binárnej oblasti. Použitý štruktúrny element je dôležitý, pretože determinuje ako presne budú okraje binárnej oblasti menené.

Obrázok 4.2 ilustruje príklad výpočtu výslednej hodnoty dilatácie pre okolie jedného bodu. Okolie bodu dané štruktúrnym elementom je vypočítané ako zjednotenie (v tomto prípade môžeme hovoriť o logickom súčte) tých hodnôt obrazu, ktorým zodpovedajú jednotkové hodnoty pixelov v štruktúrnom elemente. Výsledná hodnota je zapísaná do výstupného obrazu na pozíciu počiatku (typicky stredu) štruktúrneho elementu.

V prípade binárneho obrazu má teda výsledný pixel hodnotu 1, ak aspoň jedna hodnota spomedzi pixelov obrazu selektovaných korešpondujúcimi jednotkovými hodnotami štruktúrneho elementu je tiež nenulová, čiže jednotková.



Obrázok 4.2: Dilatácia – binárny obraz. Vľavo je vstupný, vpravo výstupný obraz. Štrukturálny element je znázornený červenou farbou. V tomto prípade je jednotková výsledná hodnota daná logickým súčtom núl s jedinou jednotkovou hodnotou (na pozícii v druhom riadku, treťom stĺpci).

Nasledujúci obrázok 4.3 ilustruje, ako operácia dilatácie zmenila tvar bielej oblasti (binárneho objektu). Biele objekty boli zväčšené, tým sa spojili do jedného objektu a čierne „diery“ vnútri objektov boli odstránené.



Obrázok 4.3: Vľavo je binárny vstupný obraz, vpravo je dilatovaný výstupný obraz. (Použitý bol štrukturálny element je znázornený na obrázku 4.4).

V tomto i nasledujúcich príkladoch bol použitý štrukturálny element ako je znázornený na obrázku 4.4.

0	0	1	0	0
0	1	1	1	0
1	1	1	1	1
0	1	1	1	0
0	0	1	0	0

Obrázok 4.4: Štruktúrálny element použitý v ďalších príkladoch.

4.4.2 Erózia

Operáciu erózie spracovávaného obrazu F a štruktúrneho elementu S označujeme

$$F \ominus S. \quad (4.12)$$

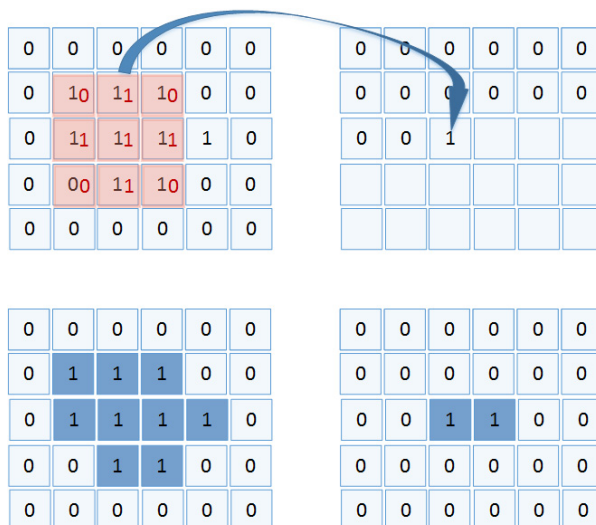
Binárnu eróziu môžeme definovať ako prienik všetkých posunov obrazu F o vektory $-s$, kde $s \in S$

$$F \ominus S = \bigcap_{s \in S} F_s, \quad (4.13)$$

kde

$$F_s = \{f - s \mid f \in F\}. \quad (4.14)$$

Postup výpočtu je znázornený na obrázku 4.5. Tu vidíme príklad výpočtu výslednej hodnoty erózie pre okolie jedného bodu. Pre okolie bodu dané štruktúrnym elementom je vypočítaný prienik (v tomto prípade môžeme hovoriť o logickom súčine) zo všetkých hodnôt obrazu, ktoré zodpovedajú jednotkovým hodnotám v štruktúrnom elemente. Výsledná hodnota je zapísaná do výstupného obrazu na pozíciu počiatku (typicky stredu) štruktúrneho elementu.



Obrázok 4.5: Erózia – binárny obraz. Vľavo je vstupný, vpravo výstupný obraz. Štruktúrálny element je znázornený červenou farbou. V tomto prípade je jednotková výsledná hodnota daná prienikom (logickým súčinom) na pozíciách, kde hodnoty v štruktúrálnom elemente sú jednotkové.



Obrázok 4.6: Vľavo je binárny vstupný obraz, vpravo je erodovaný výstupný obraz. (Použitý bol štruktúrálny element je znázornený na obrázku 4.4).

Obrázok 4.6 ilustruje stenčenie objektov na obrázku následkom morfológického erodovania. Objekty a časti objektov (napr. stred písmena E), ktoré sú menšie ako použitý štruktúrálny element, boli eróziou odstránené.

4.4.3 Vlastnosti dilatácie a erózie

Uved'me si vlastnosti binárnej dilatácie a erózie. Dilatácia, práve tak ako erózia je

- komutatívna, čiže platí $\mathbf{A} \oplus \mathbf{B} = \mathbf{B} \oplus \mathbf{A}$,
- asociatívna, čiže platí $\mathbf{A} \oplus (\mathbf{B} \oplus \mathbf{C}) = (\mathbf{A} \oplus \mathbf{B}) \oplus \mathbf{C}$,
- invariantná voči posunutiu, čiže platí $\mathbf{A} \oplus (\mathbf{B} + \mathbf{x}) = (\mathbf{A} \oplus \mathbf{B}) + \mathbf{x}$.
- Dilatácia a erózia sú vzájomne duálne operácie. Dilatovanie obrazu je ekvivalentné erodovaniu komplementárneho obrazu a naopak.

Z pohľadu použitia erózie a dilatácie je nutné si uvedomiť, že

- veľkosť štruktúrného elementu je vhodné voliť ako nepárne číslo. Potom je poloha stredu štruktúrného elementu jednoznačná. Ak táto podmienka nie je splnená, môžeme doplniť do štruktúrného elementu nuly a tým ho zväčšiť na veľkosť danú nepárnym číslom.
- funkcie na výpočet dilatácie musia ošetriť výpočet na okraji obrazu, napr. doplnením nulových hodnôt, keďže pre výpočet sú potrebné i v skutočnosti neexistujúce hodnoty mimo obraz.

4.4.4 Použitie dilatácie a erózie

Dilatáciu binárnych obrazov úspešne využívame v prípadoch, keď potrebujeme cielene zväčšiť binárne oblasti alebo odstrániť diery v týchto binárnych oblastiach. Podľa požadovaného účinku a požadovanej morfolologickej zmeny tvaru binárnych oblastí v obraze, volíme zodpovedajúcu vhodnú veľkosť a vhodný tvar štruktúrného elementu. Eróziou zase naopak môžeme odstrániť objekty, ktoré sú menšie ako rozmery štruktúrného elementu. Príklad erózie, kde sú odstránené nežiaduce detaily je na obrázku 4.7.

4.4.5 Vzájomný vzťah medzi eróziou a dilatáciou

Ak použijeme postupne operácie erózie a dilatácie s rovnakým štruktúrnym elementom, tak za určitých okolností sa môže výsledný obraz rovnať vstupnému, ale neplatí to všeobecne. **Erózia a dilatácia nemusia byť vždy vzájomne inverzné!**

Ak je napríklad eróziou malý objekt úplne odstránený, tak ho následná dilatácia už nedokáže zrekonštruovať. Táto vlastnosť samozrejme môže



Obrázok 4.7: Vľavo je binárny vstupný obraz, vpravo je erodovaný obraz. Môžeme si všimnúť, že malé biele detaily zo vstupného obrazu erodovaním zmizli na úkor čiernych oblastí.

niekedy ovplyvniť výsledky negatívne, avšak často ju využívame cielene, napr. práve pre odstránenie malých objektov.

4.5 Morfologické otvorenie a uzavretie

4.5.1 Morfologické otvorenie

Uvedieme definíciu morfologického otvorenia.

Operáciu erózie, nasledovanú dilatovaním erodovaného obrazu s použitím rovnakého štrukturálneho elementu pre obidve operácie, nazývame morfologické otvorenie. Zapišeme ako

$$(F \ominus S) \oplus S. \quad (4.15)$$

Príklad morfologického otvorenia je na obrázku 4.8 .

4.5.2 Morfologické uzavretie

Ak sú operácie erózie a dilatácie použité v opačnom poradí, tak hovoríme o morfologickom uzatvorení. Operácia dilatácie, nasledovaná erodovaním dilatovaného obrazu s použitím rovnakého štrukturálneho elementu pre obidve operácie, nazývame morfologické uzavretie. Zapišeme ako

$$(F \oplus S) \ominus S. \quad (4.16)$$



Obrázok 4.8: Príklad morfológického otvorenia. Vľavo je vstupný obraz, vpravo je výstupný obraz. (Použitý bol štruktúrny element znázornený na obrázku 4.4).

Príklad morfológického uzavretia:



Obrázok 4.9: Príklad morfológického uzatvorenia. Vľavo je vstupný obraz, vpravo je výstupný obraz. (Použitý bol štruktúrny element znázornený na obrázku 4.4).

Obidve operácie, morfológické otvorenie i morfológické uzatvorenie, približne zachovávajú veľkosť objektov. Rozdiely pozorujeme predovšetkým ako zmeny v detailoch, ktorých veľkosť je porovnateľná s použitým štruktúrnym elementom. Avšak i kontúry veľkých objektov sa môžu tvarom štruktúrneho elementu zmeniť, napr kruhový štruktúrny element kontúry objektov takpovediac „zaoblí“.

Z príkladu na obrázku 4.8 vidíme, že operácia morfológické otvorenie odstránila malé biele objekty v ľavej dolnej časti obrazu. Objekty v pravej dolnej časti obrazu boli väčšie než použitý štruktúrny element a preto odstránené neboli. Ďalší jav, ktorý pozorujeme, je odstránenie výčnelku v strede písmena E.

Naopak, z príkladu na obrázku 4.9 vidíme, že operácia morfológické uzatvorenie odstránila malé tmavé oblasti vo vnútri bieleho objektu, v na-

šom prípade vo hvieзде, ktoré označujeme ako „diery“ alebo „otvory“ (holes) v objekte. Ďalej vidíme, že objekty, medzi ktorými bola malá medzera (vzhľadom k použitému štruktúrnemu elementu) sa spojili. Toto spojenie blízkych objektov pozorujeme medzi veľkými objektami písmena E a hvieзdy a tiež v dolnej časti obrazu, kde pospájané malé objekty vytvorili jeden objekt krivkovitého tvaru.

4.6 Šedotónová dilatácia a erózia

Koncept binárnej dilatácie a erózie je možno rozšíriť na šedotónové obrazy. Najskôr uvedieme pôvodnú definíciu [HS87], kde obor funkčných hodnôt štruktúrneho elementu môže byť z oboru celých alebo reálnych čísiel a je funkciou, ktorá nemusí byť konštantná na svojom definičnom obore. Takýto štruktúrny element nazývajú autori v článku [VT96] nerovným štruktúrnym elementom (non-flat structuring element), nájdeme ho i označený ako 3D štruktúrny element pre 2D obrazy. Nesmieme si však tento štruktúrny element zamieňať s 3D štruktúrnym elementom pre 3D obrazy!

4.6.1 Nerovný štruktúrny element

Šedotónová dilatácia s použitím nerovného štruktúrneho elementu (po anglicky non-flat structuring element) je definovaná v bode $\mathbf{f}$, ktorý je zároveň kotevným bodom štruktúrneho elementu $\mathbf{S}$, nasledovne

$$\mathbf{F} \oplus \mathbf{S} = \max_{\mathbf{s} \in \mathbf{S}} \{\mathbf{F}(\mathbf{f} - \mathbf{s}) + \mathbf{S}\}, \quad (4.17)$$

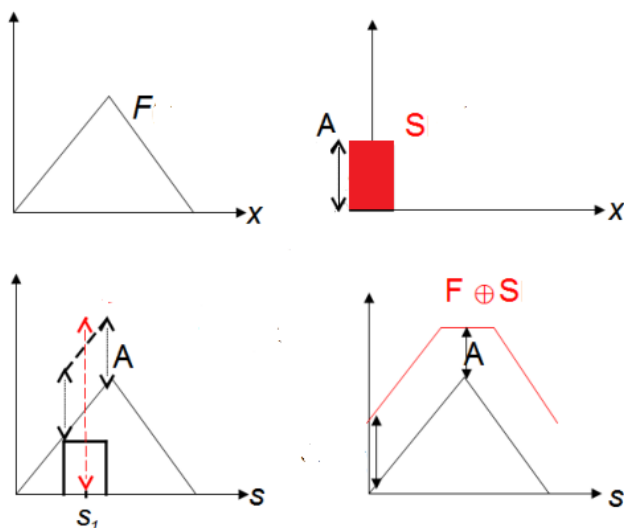
kde $\mathbf{S}$ je nerovný štruktúrny element. Postup výpočtu je ilustrovaný na jednodimenziálnom signále na obrázku 4.10.

K okoliu bodu $\mathbf{f}$, pre ktorý počítame šedotónovú dilatáciu, je najskôr štruktúrny element pripočítaný a potom je výsledná hodnota daná maximálnou hodnotou týchto súčtov.

Šedotónová erózia s nerovným štruktúrnym elementom. Šedotónová erózia s nerovným štruktúrnym elementom je definovaná ako

$$\mathbf{F} \ominus \mathbf{S} = \min_{\mathbf{s} \in \mathbf{S}} \{\mathbf{F}(\mathbf{f} + \mathbf{s}) - \mathbf{S}\}, \quad (4.18)$$

kde $\mathbf{S}$ je nerovný štruktúrny element.



Obrázok 4.10: Šedotónová dilatácia ilustrovaná na 1-dimenzionálnom signále.

Definícia erózie je podobná definícii dilatácie, ibaže operácia sčítania je nahradená odčítaním a namiesto maxima vyhl'adávame minimum.

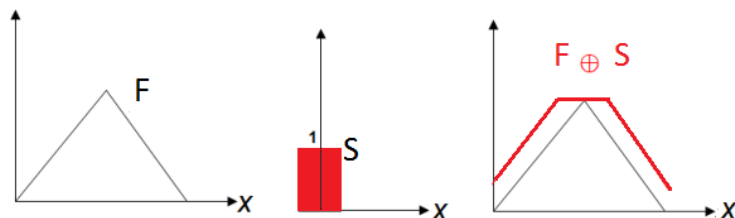
4.6.2 Rovný štruktúrny element

V praxi je bežne používaný prístup k výpočtu šedotónovej dilatácie a erózie, ktorý je založený na zjednodušenej definícii. Používa rovný štruktúrny element (flat structuring element), ktorý obsahuje hodnoty 0 a 1 (podobne ako pre binárne obrazy), a teda definuje okolie (neighborhood) daného bodu. Výstupná hodnota je potom maximum (u dilatácie) alebo minimum (u erózie) hodnôt obrazu v definovanom okolí daného bodu.

Šedotónová dilatácia. Šedotónová dilatácia s rovným štruktúrnym elementom je definovaná ako

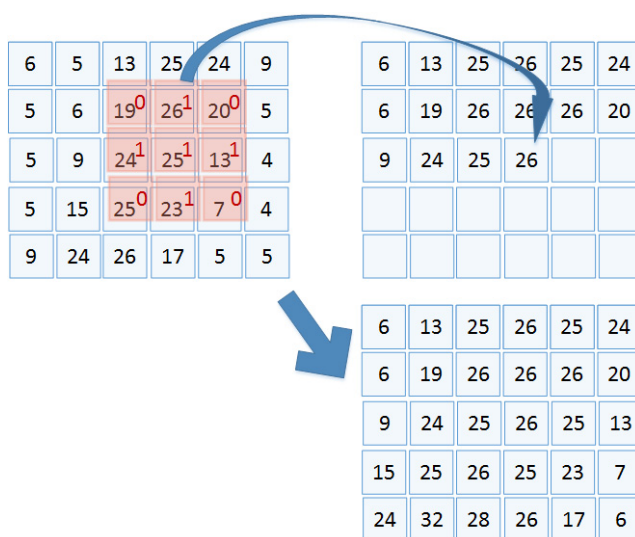
$$F \oplus S = \max_{s \in S} \{F(f - s)\}, \quad (4.19)$$

kde S je rovný štruktúrny element.



Obrázok 4.11: Ukážka šedotónovej dilatácie s rovným štruktúrnym elementom.

Postup výpočtu ilustrovaný na 2D obraze je na obrázku 4.12.



Obrázok 4.12: Výpočet šedotónovej dilatácie. Rovný štruktúrny element je znázornený červenou farbou.

Jednoducho formulované, hodnota výsledného pixela po dilatovaní sa rovná **maximálnej hodnote** zo všetkých pixelov obrazu, ktoré sú selektované korešpondujúcimi jednotkovými hodnotami v štruktúrnom elemente. Ide teda vlastne o ekvivalent zoraďovacieho filtra (order-statistic filter) maxima v prípade dilatácie a zoraďovacieho filtra minima v prípade erózie (viď 2.5.1).



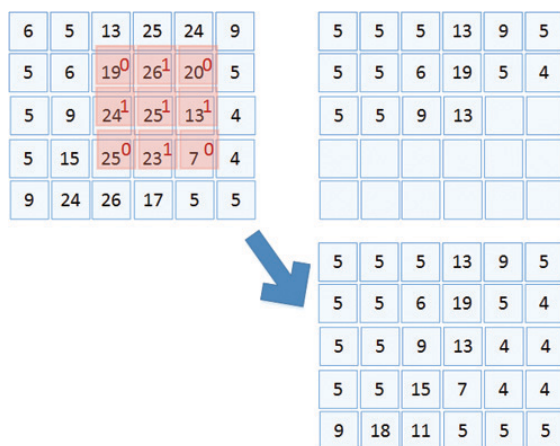
Obrázok 4.13: Príklad šedotónovej dilatácie. Vľavo je vstupný obraz, vpravo je výstupný obraz. (Použitý bol štruktúrny element znázornený na obrázku 4.4).

Šedotónovú dilatáciu použijeme napr. vtedy, keď potrebujeme odstrániť lokálne minimá menšie ako štruktúrny element alebo keď potrebujeme zväčšiť svetlé oblasti v obraze. Samozrejme dôkladne musíme zvážiť veľkosť a tvar štruktúrneho elementu.

Šedotónová erózia s rovným štruktúrnym elementom. Šedotónovú eróziu definujeme vzťahom

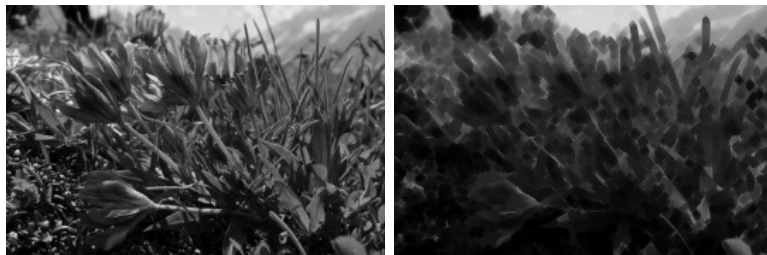
$$F \ominus S = \min_{s \in S} \{F(f + s)\}, \quad (4.20)$$

kde S je rovný štruktúrny element (obr. 4.14).



Obrázok 4.14: Výpočet šedotónovej erózie. Rovný štruktúrny element je znázornený červenou farbou.

Šedotónová erózia nachádza uplatnenie napr. v prípade, keď chceme odstrániť malé lokálne minimá, čiže oblasti v obraze svetlej farby, ktorých veľkosť je menšia ako veľkosť použitého štruktúrneho elementu. Ako ilustráciu uvádzame príklad na obrázku 4.15.



Obrázok 4.15: Príklad šedotónovej erózie. Vľavo je vstupný obraz, vpravo je výstupný obraz. (Použitý bol štruktúrny element znázornený na obrázku 4.4).

Pre zaujímavosť si zdôrazníme rozdiel medzi šedotónovou dilatáciou a filtráciou konvolúciou. Tento rozdiel je úplne principiálny, spočíva v použití rozdielnych matematických operácií. Namiesto súčtu u konvolúcie je u dilatácie výber maxima, namiesto násobenia medzi hodnotami obrazu a filtračným jadrom pri konvolúcii je operácia súčtu u šedotónovej dilatácie. Rozdiel je i v tom, že filtračné jadro je transponované.

4.7 Morfologický gradient

Pre získanie informácie reprezentujúcej hrany v obraze nám základná teória spracovania obrazu ponúka mnohé prístupy. Jednou z možností je využitie morfologických operácií.

Odčítaním pôvodného obrazu od morfologicky dilatovaného, prípadne erodovaného obrazu, počítame morfologický gradient. Samozrejme je potrebné zvážiť absolútnu hodnotu a vhodné preškálovanie vypočítaného rozdielu.

Ďalšou možnosťou je vypočítať rozdiel medzi erodovaným a dilatovaným obrazom, ktoré sme získali morfologickým spracovaním rovnakého vstupného obrazu. Postup výpočtu morfologického gradientu **MG** môžeme zapísať ako

$$\mathbf{MG} = (\mathbf{F} \oplus \mathbf{S}) - (\mathbf{F} \ominus \mathbf{S}). \quad (4.21)$$

Binárny morfológický gradient. Pre extrakciu obrysu objektu (boundary extraction) binárneho obrazu môžeme použiť operáciu morfológický gradient.

Príklad morfológického gradientu pre binárny obraz je na obrázku 4.16.



Obrázok 4.16: Príklad binárneho morfológického gradientu. Vľavo je vstupný obraz, vpravo je výstupný obraz.

Ako vidíme, morfológický gradient nám pre binárne obrazy umožňuje získať kontúry binarizovaných oblastí. Morfológický gradient však môžeme využiť i pre šedotónový obraz.

Šedotónový morfológický gradient. Príklad morfológického gradientu pre šedotónový obraz je na nasledujúcom obrázku 4.17.



Obrázok 4.17: Príklad šedotónového morfológického gradientu. Vľavo je vstupný obraz, vpravo je výstupný obraz.

Pre uvedený príklad šedotónového morfológického gradientu sme použili štruktúrny element veľkosti 47×47 kruhového tvaru, pričom veľkosť vstupného obrazu bola 610×536 pixelov.

4.8 Morfológická vrchná a spodná časť klobúka

Kombinovaním základných morfológických a aritmetických operácií vznikajú ďalšie typy morfológických operácií, ako sú napr. morfológická vrchná a spodná časť klobúka. V anglickej literatúre nachádzame tieto morfológické operácie pod názvami Top-hat a Bottom-hat.

Morfológická operácia vrchná časť klobúka (Top-hat). Morfológická operácia vrchná časť klobúka vznikne odčítaním morfológického otvorenia od pôvodného obrazu. Môžeme ju zapísať ako

$$F - ((F \ominus S) \oplus S). \quad (4.22)$$

Morfológická operácia spodná časť klobúka (Bottom-hat). Duálna morfológická operácia spodná časť klobúka naopak vznikne odčítaním pôvodného obrazu od morfológického uzavretia pôvodného obrazu. Zapíšeme ju nasledovne.

$$((F \oplus S) \ominus S) - F. \quad (4.23)$$

Príklad použitia morfológickej operácie vrchná časť klobúka (Top-hat) pre šedotónový obraz je na obrázku 4.18. Použitý štruktúrny element je kruhového tvaru o veľkosti 47×47 a veľkosť vstupného obrazu je 584×515 pixelov. Vo výstupnom obraze si všimneme výrazné oblasti, ktoré vo vstupnom obraze predstavujú svetlý objekt na tmavom pozadí, pričom jeden rozmer objektu je menší ako veľkosť štruktúrneho elementu.



Obrázok 4.18: Príklad operácie vrchná časť klobúka. Vľavo je vstupný obraz, vpravo je výstupný obraz. Môžeme pozorovať zvýraznené oblasti, napr. tam kam ukazuje červená šípka, ktoré odpovedajú oblastiam s rozmerom menším ako štruktúrny element, pričom sú tieto oblasti svetlé na tmavom pozadí.



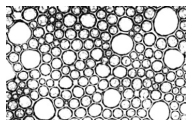
Obrázok 4.19: Príklad operácie spodná časť klobúka. Vľavo je vstupný obraz, vpravo je výstupný obraz. Podobne ako v predchádzajúcom príklade môžeme pozorovať zvýraznené oblasti, napr. tam kam ukazujú červené šípky, ktoré odpovedajú oblastiam s rozmerom menším ako štrukturálny element. Na rozdiel od predchádzajúceho prípadu sú tieto oblasti tmavé na svetlom pozadí.

Príklad použitia morfolologickej operácie spodná časť klobúka (Bottom-hat) pre šedotónový obraz vidíme na obrázku 4.19. Použitý štrukturálny element je rovnaký ako v predchádzajúcom príklade vrchnej časti klobúka, ale na výstupnom obraze pozorujeme opačný jav než v predchádzajúcom príklade. Dominantné oblasti, pozorovateľné vo výstupnom obraze, odpovedajú tým oblastiam vstupného obrazu, ktoré sú tmavé a sú na svetlom pozadí, pričom jeden rozmer takéhoto objektu je menší ako veľkosť štrukturálneho elementu.

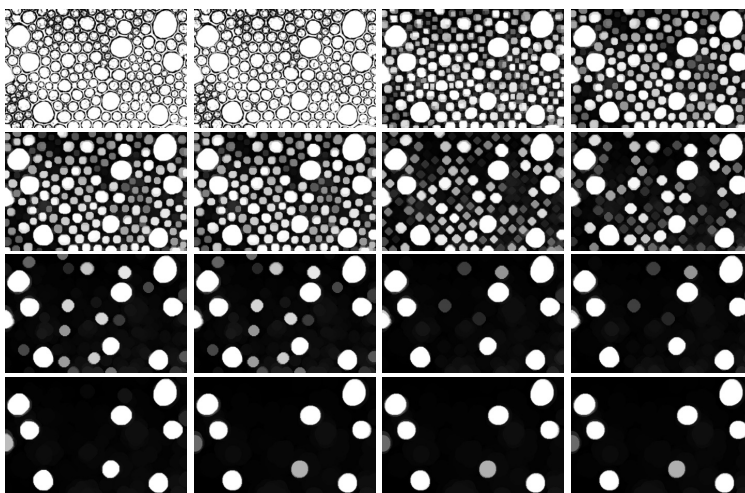
4.9 Granulometria

Morfologické operácie môžeme výhodne použiť v úlohách, ktoré si kladú za cieľ odhadnúť počet, veľkosť, prípadne distribúciu určitých častíc (particles) v obraze.

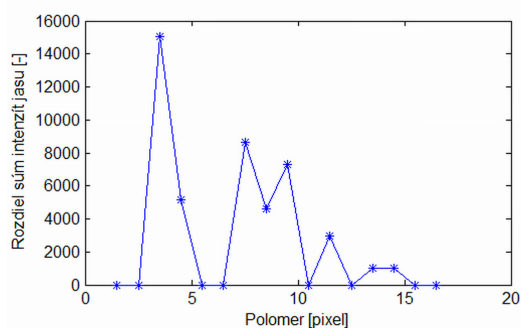
Postup je jednoduchý. V prípade, že detegované častice sú svetlejšie než pozadie a sú pravidelných tvarov, tak použijeme operáciu morfolologickeho otvorenia (opening) so štrukturálnym elementom, ktorého veľkosť sa bude meniť. (Je vhodné, aby tvar štrukturálneho elementu kopíroval typický tvar detegovaných častíc.) Operáciu otvorenia aplikujeme opakovane, pričom veľkosť štrukturálneho elementu sa každý raz zväčší. Pri každom kroku vyhodnotíme rozdiel sumy všetkých pixelov (surface area) v obraze pred a po operácii otvorenia. Získame graf na obrázku 4.22.



Obrázok 4.20: Vstupný obraz pre príklad morfológického vyhodnotenia granulometrie.



Obrázok 4.21: Morfológické otvorenie s postupne sa zväčšujúcim štruktúrnym elementom.



Obrázok 4.22: Závislosť zmeny sumy hodnoty jasu pred a po operácii morfológického otvorenia pre štruktúrne elementy kruhového tvaru s rôznymi polormi.

Na x -ovej osi je vynesný polomer štruktúrneho elementu, na y -ovej osi je príslušný rozdiel súm intenzity obrazu (surface area) pred a po operácii morfológického otvorenia.

Z grafu vidíme dominantné vrcholy, ktoré zodpovedajú veľkostiam častíc, ktoré sú najvýznamnejšie zastúpené.

4.10 Morfológická operácia Hit-or-Miss

„Hit-or-Miss“ morfológická operácia (známa i pod názvami „Hit-or-Miss Transform“ alebo „Hit-and-Miss Transform“) patrí medzi všeobecne známe binárne morfológické operácie, ktoré môžu byť použité na hľadanie konkrétnych objektov na pozadí. „Hit-or-Miss“ je definovaná pre dva komplementárne štruktúrne elementy a pracuje s binárnym obrazom na vstupe i na výstupe.

Dva komplementárne štruktúrne elementy **Shm1** a **Shm2** môžeme zapísať ako

$$\mathbf{Shm} = (\mathbf{Shm1}, \mathbf{Shm2}), \quad (4.24)$$

kde **Shm1** je štruktúrny element asociovaný s detegovaným objektom a **Shm2** je štruktúrny element asociovaný s pozadím.

Morfológická operácia „Hit-or-Miss“ je potom

$$\mathbf{HM} = (\mathbf{F} \ominus \mathbf{Shm1}) \cap (\mathbf{F}^c \ominus \mathbf{Shm2}), \quad (4.25)$$

kde $\mathbf{F}$ je pôvodný binárny obraz, $\mathbf{F}^c$ je jeho komplementárny obraz.

Morfológická operácia „Hit-or-Miss“ je vhodná na detekciu všetkých binárnych objektov, ktoré zodpovedajú danému štruktúrnemu elementu **Shm1** („hit“) a zároveň štruktúrny element **Shm2** („miss“) definuje ich nutné pozadie. Inými slovami **Shm1** definuje pixely, ktoré musia byť jednotkové a zároveň **Shm2** definuje pixely okolia, ktoré nesmú byť jednotkové, čiže musia byť nulové.

Je zrejmé, že základná definícia morfológickej operácie „Hit-or-Miss“ deteguje iba objekty s tvarom presne zodpovedajúcim daným štruktúrnym elementom. Ak ju chceme využiť i na menej reštriktívnu detekciu, tak je vhodná implementácia s tzv. kompozitným štruktúrnym elementom (viď obrázok 4.23), ktorá akceptuje tri rôzne hodnoty v štruktúrnym elemente

$$\mathbf{Shm} = (\mathbf{Shm1}, \mathbf{Shm2}, \mathbf{Shm3}). \quad (4.26)$$

Príklad na obrázku 4.24 ilustruje použitie kompozitného štruktúrneho elementu pri detekcii objektu, v tomto prípade číslu 6.



Obrázok 4.23: Príklad kompozitného štruktúrneho elementu. Pixely zobrazené bielou farbou zodpovedajú štruktúrnemu elementu **Shm1** („hit“), pixely znázornené čiernou zodpovedajú štruktúrnemu elementu **Shm2** („miss“), šedé pixely znamenajú ignorovať.

```

1 CEMENTO DAK63660   dlažba 60x60 I.A
  DAK636601/IK/8   /
2 CEMENTO DAK63661   dlažba 60x60 I.A
  DAK636611/IK/8   /D006
3 CEMENTO DAKSE660   dlažba 30x60 I.A
  DAKSE6601/IK/8   /
  
```



Obrázok 4.24: Príklad použitia „Hit-or-Miss“ transformácie s kompozitným štruktúrnym elementom.

Z uvedeného príkladu je zrejmé, že pomocou „Hit-or-Miss“ transformácie sa nám podarilo detegovať číslicu „6“ pre ktorú sme definovali štruktúrne elementy.

K uvedenej metóde je potrebné ešte podotknúť, že nie je ani škálovo ani rotačne invariantná.

4.11 Vzdialenostná transformácia

V anglickej literatúre nájdeme vzdialenostnú transformáciu pod názvom „distance transform“.

Máme binárny obraz, v ktorom nulové hodnoty označujú pozadie. Vzdialenostná transformácia nahradí každý nenulový bod obrazu (pixel) jeho vzdialenosťou k najbližšiemu pixelu pozadia.

Najčastejšie používaná metrika pre výpočet vzdialenostnej transformácie je Euklidovská metrika.

Príklad vzdialenostnej transformácie je na ďalšom obrázku 4.25.



Obrázok 4.25: Príklad vzdialenostnej transformácie s Euklidovou metrikou. Vľavo je vstupný obraz, vpravo je výstupný obraz.

4.12 Kostra

Morfologická operácia kostra (anglicky Skeleton) je definovaná pre binárne obrazy. Samotná definícia kostry vychádza z definície maximálneho kruhu.

- Maximálny kruh $B(p, r)$ so stredom p a polomerom r , $r \geq 0$ je množina bodov, pre vzdialenosť ktorých od stredu p platí $d \leq r$.
- Maximálny kruh B vpísaný do množiny X sa dotýka hranice množiny X v dvoch a viacerých bodoch.
- Kostra je potom zjednotenie stredov maximálnych kruhov.

4.12.1 Príklady algoritmických postupov výpočtu kostry

Výpočet podľa definície vpisovania maximálnych kruhov. Tento postup je výpočtovo veľmi náročný a prakticky sa nepoužíva.

Sekvenčné stenčovanie. Oblasť sa eroduje vhodným štruktúrnym elementom tak, aby nebola porušená súvislosť kostry. Príklad kostry počítačnej sekvenčným stenčovaním je na obrázku 4.26.



Obrázok 4.26: Príklad výpočtu kostry (skeleton). Vľavo je vstupný binárny obraz, vpravo je obraz kostry.

Pomocou vzdialenostnej transformácie. Tento postup patrí pravdepodobne medzi najpoužívanejšie. Umožňuje rýchly výpočet.

4.13 Transformácia rozvodím

Originálny anglický názov je Watershed transformation. Táto morfológická operácia je pre šedotónové obrazy. Koncept transformácie rozvodím je založený na predstave obrazu ako topografického reliéfu, kde vysoké intenzity zodpovedajú pohoriam a nízke zase údoliam. Existujú dva základné algoritmické prístupy – *Zhora nadol* a *Zdola nahor*.

Zhora nadol. Pri tomto prístupe je potrebné nájsť pre každý pixel cestu vedúcu do minima v oblasti zodpovedajúcej zbernej nádrži. Pixel sa usporiadať podľa toho, v ktorom minime sa skončí ich cesta. Každý bod topologického priestoru, čiže každý pixel v obraze, patrí do jednej z týchto kategórií:

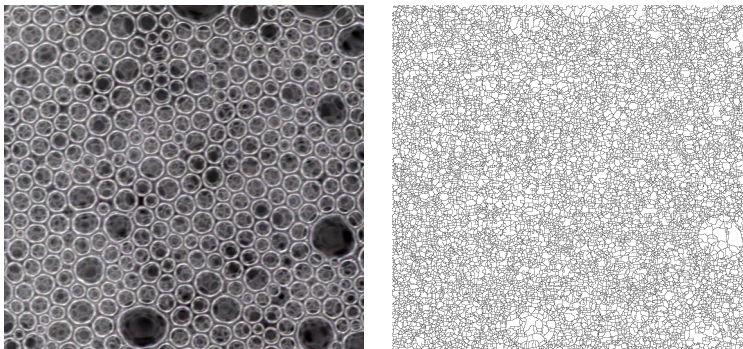
- je regionálnym minimom,
- alebo existuje z tohto bodu práve jedna cesta do regionálneho minima,
- alebo s rovnakou pravdepodobnosťou existuje cesta do viacerých miním.

V poslednom prípade je pixel označený ako súčasť vysegmentovanej hranice rozvodí.

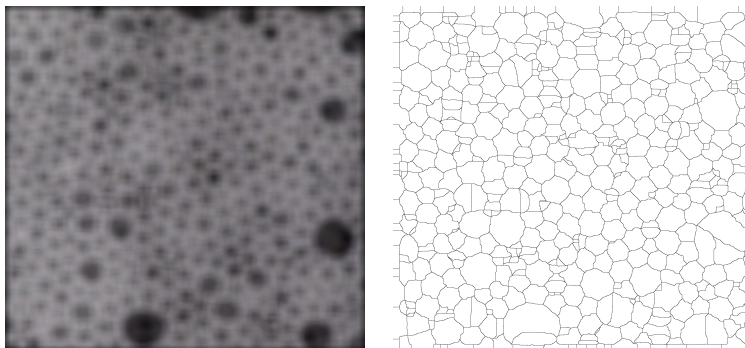
Zdola nahor. Pri druhom prístupe každé minimum reprezentuje jednu zbernú nádrž a stratégia spočíva v plnení týchto nádrží smerom nahor. Ak by pri naplňaní malo dôjsť k spojeniu dvoch nádrží, tak sa vytvorí umelá hranica (watershed), ktorá tomu zabráni. Práve vytvorené hranice sú výstupom segmentácie rozvodím.

Značková segmentácia rozvodím. Pre niektoré aplikácie je vhodné použitie vopred daných značiek pre polohu miním. Tento algoritmický prístup vyžaduje využiť aj iný zdroj informácie, či už získaný automaticky alebo manuálne, pre riadenie transformácie rozvodím.

V praxi sa teda stretávame s implementáciami segmentácie rozvodím s použitím algoritmov, ktoré využívajú značky – semiačka pre definovanie polohy minima. Na druhej strane sú v praxi používané i implementácie segmentácie rozvodím, ktoré pracujú bez značiaceho obrazu a výpočet polohy lokálnych miním je ich súčasťou.



Obrázok 4.27: Príklad transformácie rozvodím. Vľavo je vstupný obraz, vpravo presegmentovaný výstupný obraz.



Obrázok 4.28: Príklad transformácie rozvodia. Vľavo je vstupný obraz z obr. 4.27 vyhladený gaussovým filtrom. Vpravo je výstup transformácie rozvodia vyhladeného vstupu.

Typickým problémom pri použití transformácie rozvodím je presegmentovanie (oversegmentation), keď sa uplatnia príliš mnohé lokálne minimá (často na úrovni šumu v obraze) a výsledná segmentácia rozvodím nesleduje iba dominantné oblasti, ale vygeneruje veľké množstvo malých segmentov.

V prípade, že nechceme presegmentovanie generovať zámerne (napr. pri tvorbe tzv. „superpixelov“), tak použijeme metódy pre vyhladenie lokálnych miním – či už pomocou jednoduchej dolnopriepustnej filtrácie, ako je demonštrované na príklade 4.28, kde bol obraz vyhladený konvolyčne Gaussovským filtrom o veľkosti 51×51 , alebo sofistikovanejšími metódami, ako napr. morfológická rekonštrukcia. (Vid' ďalej 4.15).

4.13.1 Značková segmentácia rozvodím pre farebný obraz

Meyer navrhol efektívny algoritmus výpočtu značkovej segmentácie rozvodím metódou rastúcich regiónov. Tento algoritmus bol publikovaný v [Mey92].

- Inicializácia. Značkám (semiačkam) je priradený jednoznačný identifikátor (číslo). Každý región si zachová hodnotu identifikátora tej značky, ktorá bola jeho zdrojom. Potom je vygenerovaná usporiadaná postupnosť s takým počtom úrovní priority, koľko je úrovní šedej v obraze. Všetky body susediace so značkou (semiačkou) sú vložené do tejto usporiadanej postupnosti.

- Rast značiek do regiónov. Ak má bod vo svojom okolí jeden a práve len jeden identifikátor regiónu, tak je priradený do tohoto regiónu. Jeho susedia bez identifikátorov, ktorí sú stále ešte nezaradení do usporiadanej postupnosti, sú vložené do usporiadanej postupnosti s rovnakým príslušnou prioritou ako vo fáze inicializácie.
- Pokiaľ jeden bod susedí s dvomi regiónmi s rôznymi identifikátormi, tento bod sa stane hraničným bodom a dostane špeciálne označenie charakterizujúce hranice.

Použitá metrika podobnosti d medzi bodom p a regiónom r je daná rozdielom vo farbe nasledovne:

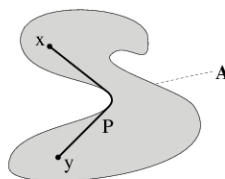
$$d = \max\{|p_r - r_r|, |p_g - r_g|, |p_b - r_b|\}. \quad (4.27)$$

4.14 Geodetické metódy

Geodetické metódy riadia morfológické operácie tak, aby operovali len na časti obrazu. Ak sa má napríklad rekonštruovať objekt a my máme značku, ktorá nám daný objekt ukazuje, tak je vhodné obmedziť rekonštrukciu iba na hranice objektu.

4.14.1 Geodetická vzdialenosť

Ústredným pojmom geodetických metód je geodetická vzdialenosť. Je to cesta medzi dvoma bodmi, ktorá je obmedzená tým spôsobom, že musí ležať vo vnútri danej množiny.



Obrázok 4.29: Príklad geodetickej vzdialenosti.

4.15 Morfológická rekonštrukcia

Morfológická rekonštrukcia je založená na iteratívnom procese, pri ktorom sa opakuje výpočet geodetickej binárnej dilatácie, alebo sa opakuje

výpočet geodetickej binárnej erózie. V prvom prípade hovoríme o rekonštrukcii dilatáciou, v druhom prípade hovoríme o rekonštrukcii eróziou.

4.15.1 Geodetická binárna dilatácia

Okrem spracovávaného obrazu, označovaného tiež ako maskovací obraz (mask image) **F** a štrukturálneho elementu **S**, potrebujeme ďalší, tzv. značiaci obraz (marker image) **G**. Maskovací obraz je vlastne vstupný spracovávaný obraz.

Vstupom pre prvý krok iteratívneho procesu je značiaci obraz **G**. Na tento sa najskôr aplikuje operácia dilatácie s daným štrukturálnym elementom **S** a následne prienik s maskovacím obrazom **F**. Výstup týchto operácií je vstupom pre ďalšiu iteráciu. Postup sa opakuje až po dosiahnutie nemenného stavu v maskovacom obraze. Túto vlastnosť nazývame **idempotencia**.

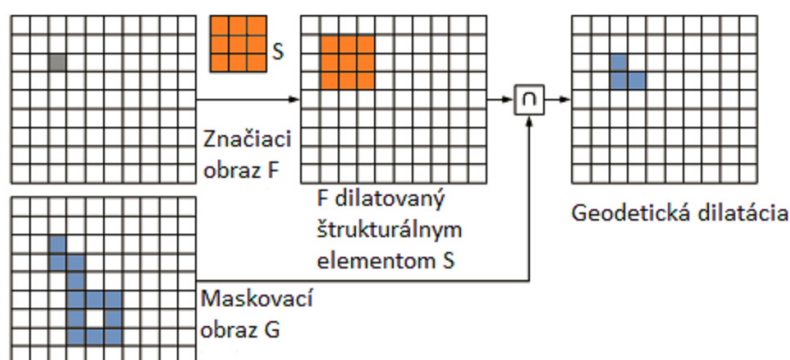
Iteratívny proces geodetickej dilatácie maskovacieho obrazu **F** podľa značiaceho obrazu **G** s použitím štrukturálneho elementu **S** zapíšeme nasledovne

$$\mathbf{D}_G^{(0)} = \mathbf{F}, \quad (4.28)$$

$$\mathbf{D}_G^{(1)} = (\mathbf{F} \oplus \mathbf{S}) \cap \mathbf{G}, \quad (4.29)$$

$$\mathbf{D}_G^{(n)} = \mathbf{D}_G^{(1)}(\mathbf{D}_G^{(n-1)}(\mathbf{F})), \quad (4.30)$$

$\mathbf{D}_G^{(n)}$ je výstup geodetickej dilatácie po n -tej iterácii podľa značiaceho obrazu **G**.



Obrázok 4.30: Príklad geodetickej dilatácie.

Postupnými iteráciami dochádza k vyplňaniu tých súvislých oblastí, ktoré boli označené v značiacom obraze. Selektované oblasti, ktoré neboli označené v značiacom obraze sú v priebehu výpočtu vymazané, až po ukončení výpočtu úplne zmiznú. Je vhodné zvoliť malý štruktúrálny element, napr veľkosti 3×3 . Väčšie štruktúrne elementy by takpovediac „preklenuli“ medzeru zodpovedajúcej veľkosti medzi dvomi samostatnými súvislými oblast'ami. Pokiaľ túto vlastnosť nepoužívame cielene, napr. na spojenie blízkych súvislých oblastí, tak príliš veľký štruktúrálny element môže spôsobiť nežiadúci efekt.

4.15.2 Geodetická binárna erózia

Podobne ako geodetickú dilatáciu, definujeme i geodetickú eróziu

$$\mathbf{D}_G^{(0)} = \mathbf{F}, \quad (4.31)$$

$$\mathbf{D}_G^{(1)} = (\mathbf{F} \ominus \mathbf{S}) \bigcup \mathbf{G}, \quad (4.32)$$

$$\mathbf{D}_G^{(n)} = \mathbf{D}_G^{(1)} (\mathbf{D}_G^{(n-1)}(\mathbf{F})), \quad (4.33)$$

$\mathbf{D}_G^{(n)}$ je výstup geodetickej erózie po n -tej iterácii podľa značiaceho obrazu $\mathbf{G}$.

Geodetická erózia je duálna operácia ku geodetickej dilatácii.

Iteratívny výpočet prebieha rovnako, avšak s tým rozdielom, že operácia dilatácie je nahradená eróziou a namiesto prieniku s maskovacím obrazom, počítame zjednotenie s maskovacím obrazom. Čiže pri geodetickej erózii je na obraz $\mathbf{F}$ postupne aplikovaná erózia, pričom po každej erózii je na erodovaný obraz ešte aplikujeme zjednotenie s maskou $\mathbf{G}$, čím sa zabráni prílišnej erózii $\mathbf{F}$ pod zvolenú minimálnu hranicu, definovanú maskou.

4.15.3 Morfológická rekonštrukcia a jej použitie

Rekonštrukcia obrazu eróziou $\mathbf{F}$ vzhľadom na značiaci obraz $\mathbf{G}$ pomocou erózie je definovaná ako geodetická erózia $\mathbf{F}$ vzhľadom na $\mathbf{G}$ až do okamihu dosiahnutia nemenného stavu. Duálna definícia je pre rekonštrukcia obrazu dilatáciou.

Diery v binárnom obraze môžeme interpretovať ako regionálne minimum, ktoré nie sú priamo spojené s okrajom obrazu. Rekonštrukciu obrazu môžeme použiť na vyplňovanie dier, keďže odstraňuje všetky minimum, ktoré nie sú spojené s okrajom obrazu $\mathbf{F}$.

Opakovanie geodetickej dilatácie umožňuje rekonštrukciu súvislých regiónov v obraze F , ktoré boli pôvodne označené pomocou značky v značiacom obraze G . Naopak, súvislé regióny v obraze F , ktoré neboli označené v značiacom obraze G , aplikovaním geodetickej dilatácie zmiznú.

4.16 Morfológická šedotónová rekonštrukcia

Metódu pod názvom „Morfológická šedotónová rekonštrukcia“ publikoval Luc Vincent v článku [Vin93a]. Je to veľmi účinná metóda na spracovanie obrazu s využitím regionálnych maxím.

Najskôr definujeme geodetickú dilatáciu a geodetickú eróziu pre šedotónové obrazy.

4.16.1 Geodetická šedotónová dilatácia a erózia

Definícia šedotónovej geodetickej dilatácie je odvodená od binárnej geodetickej dilatácie (viď odstavce 4.15.1) s tým rozdielom, že operácia dilatácie je v tomto prípade šedotónová dilatácia a prienik množín je nahradený výberom minimálnej hodnoty po bodoch medzi množinami. Iterovaním takto definovanej šedotónovej geodetickej dilatácie maskovacieho obrazu F s použitím značiaceho obrazu G získame šedotónovú morfológickú rekonštrukciu. Zapišeme ako

$$D_G^{(0)} = F, \quad (4.34)$$

$$D_G^{(1)} = (F \oplus S) \bigwedge G, \quad (4.35)$$

$$D_G^{(n)} = D_G^{(1)} (D_G^{(n-1)}(F)), \quad (4.36)$$

pričom

$\bigwedge$ značí výber minimálnej hodnoty po bodoch,

$D_G^{(n)}$ je výstup šedotónovej geodetickej dilatácie po n -tej iterácii podľa značiaceho obrazu G .

$$D_G^{(0)} = F, \quad (4.37)$$

$$D_G^{(1)} = (F \ominus S) \bigvee G, \quad (4.38)$$

$$D_G^{(n)} = D_G^{(1)} (D_G^{(n-1)}(F)), \quad (4.39)$$

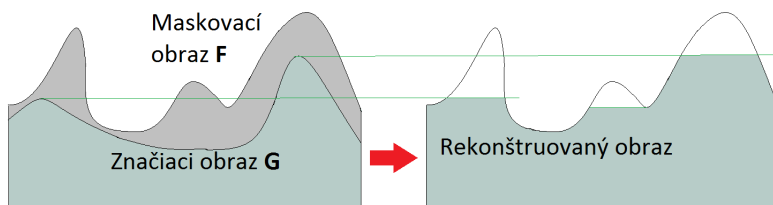
pričom

$\bigvee$ značí výber maximálnej hodnoty po bodoch,

$D_G^{(n)}$ je výstup šedotónovej geodetickej erózie po n -tej iterácii podľa značiaceho obrazu G .

Šedotónovú rekonštrukciu ilustruje nasledovný príklad. Pre vykonanie operácie šedotónovej rekonštrukcie potrebujeme najskôr vytvoriť značiaci obraz G , ktorý riadi spracovanie. Značiaci obraz $G(x,y)$ musí spĺňať podmienku

$$G(x, y) \leq F(x, y). \quad (4.40)$$



Obrázok 4.31: Morfológická šedotónová rekonštrukcia v 1D.

Na obr. 4.31 je ilustrovaný jednodimenzionálny prípad morfolologickej šedotónovej rekonštrukcie. Funkcia F tu reprezentuje vstupné hodnoty (masku) v šedotónovej škále, G reprezentuje značiaci obraz. Morfológická šedotónová rekonštrukcia odstráni regionálne maximá. Ktoré maximá budú odstránené, aká bude miera ich odstránenia – toto je určené použitým značiacim obrazom G .

4.16.2 Generovanie značiaceho obrazu

Spôsob a konkrétny algoritmus generovania značiaceho obrazu je priamo závislý od aplikácie pre ktorú túto operáciu spracovania chceme použiť. Uvedieme niektoré najpoužívanejšie prístupy

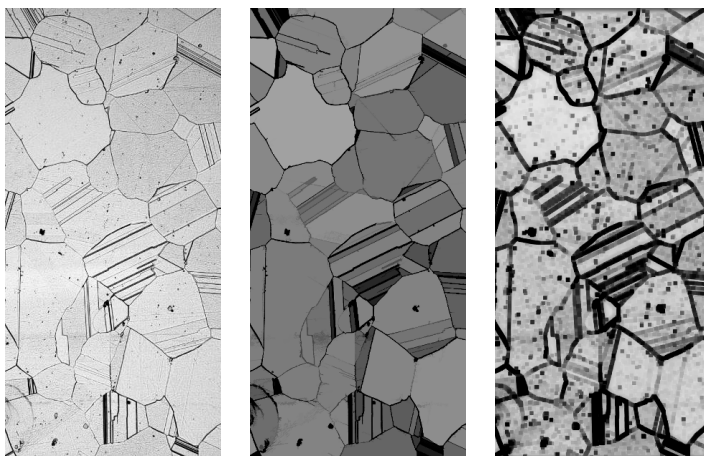
- Morfológická erózia
- Odčítanie konštantnej hodnoty od vstupného obrazu
- Sofistikované prístupy

Prvé dva menované prístupy zaručujú splnenie podmienky (4.40). Ak vytvárame značiaci obraz iným algoritmom, musíme splnenie tejto podmienky garantovať.

4.16.3 Príklady použitia

Vyhľadanie a odstránenie lokálnych maxím z obrazu. Zlepšenie kvality obrazu (image enhancement) odstránením tej časti signálu, ktorá má pre danú aplikáciu rušivý charakter, je často používaný prístup pri spracovaní obrazu.

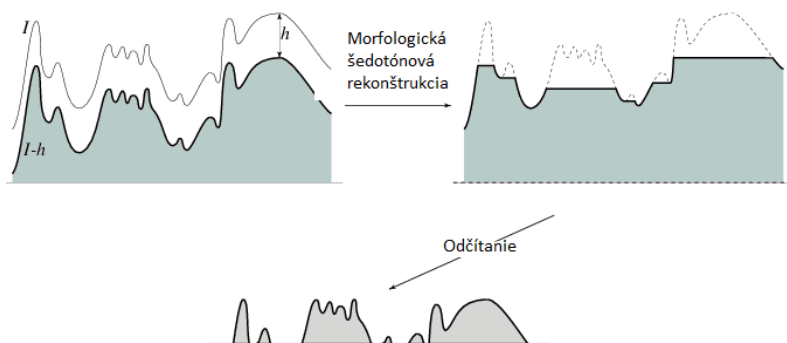
Morfológická šedotónová rekonštrukcia nám umožňuje odstrániť rušivé lokálne extrémny bez porušenia hrán. Ak chceme odstrániť lokálne maximá so známymi požiadavkami na ich maximálnu veľkosť, prípadne so známymi požiadavkami na ich tvar, tak je vhodné vytvoriť značiaci obraz s použitím morfológickej erózie. Na obrázku 4.32 je ukážka úpravy mikroskopických obrazov zliatiny Inconel ako bola publikovaná v článku [BRJ06]. Šedotónová rekonštrukcia bola použitá dvakrát, a to najskôr na mikroskopický obraz a potom na invertovaný obraz po rekonštrukcii. Značiaci obraz bol generovaný sofistikovaným algoritmom so zámerom definovať oblasti, ktoré majú byť chránené pred vyhladením, ako sú napríklad hrany regiónov odpovedajúcich zrnám v oceli.



Obrázok 4.32: Príklad vylepšenie kvality obrazu šedotónovou morfológickou rekonštrukciou. Vľavo je vstupný obraz, v strede je obraz po morfológickej rekonštrukcii, vpravo ukážka jedného z dvoch použitých značiacich obrazov.

Extrahcia lokálnych maxím. V niektorých aplikáciách je dôležitým krokom extrahovanie lokálnych extrémov. V tomto prípade môžeme získať tieto dáta odčítaním výstupného obrazu od vstupného ako je ukázané na obrázku 4.33.

4.17. Príklad použitia morfológických operácií pre segmentáciu superpixelov



Obrázok 4.33: Odčítanie rekonštruovaného obrazu od vstupného.

4.17 Príklad použitia morfológických operácií pre segmentáciu superpixelov

Superpixely sú segmenty obrazu, ktoré môžu slúžiť ako základné jednotky ďalšieho spracovania obrazu. Ich cieľom je znížiť redundanciu v obraze a zvýšiť účinnosť z hľadiska ďalšieho spracovania úlohy. Metódy na výpočet superpixelov využívajú rôzne prístupy, ako napr. metódy posunu stredov (Mean-shift) [CMM02], priestorovo koherentné zhľukovanie s použitím rezov grafov (Graph Cuts) [ZK04], entropickú segmentáciu superpixelov (Entropy Rate Superpixel Segmentation) [Liu+11], jednoduché lineárne iteratívne zhľukovanie (Simple Linear Iterative Clustering (SLIC)), prípadne iteratívne algoritmy ako TurboPixels [Lev+09].

Na obrázku 4.34 je ukážka segmentovania na superpixely.

Postup, ktorý je ďalej popísaný, demonštruje výpočet superpixelov práve s použitím morfológických operácií, ktoré sme uvádzali v tejto kapitole. Nejedná sa už o jednoduchú morfológickú operáciu, ale sofistikovaný prístup, ktorý kombinuje viaceré morfológické metódy. Výsledkom je rýchly spôsob segmentácie obrazu do superpixelov pomocou morfológického spracovania vhodný pre aplikácie v reálnom čase alebo takmer v reálnom čase. Algoritmus v tejto forme doteraz nebol publikovaný. Ďalej ho budeme skrátene nazývať MSS (Morphological Superpixel Segmentation). **Požiadavky na navrhovanú metódu** môžeme deklarovať nasledovne

- Výpočet musí byť dostatočne rýchly pre využitie v aplikáciách bežiacich v reálnom čase.



Obrázok 4.34: Ukážka segmentovania na superpixely.

- Vytvorené superpixely musia mať dostatočne malú variáciu v rámci jedného superixela a naopak môžu mať veľkú variáciu medzi rôznymi superpixelmi.
- Musí byť garantovaná priestorová koherencia v obraze.
- Hranice segmentovaných superpixelov musia sledovať hranice pravidelných obdĺžnikových oblastí, pokiaľ nie je v blízkosti významná hrana (saliency edge).
- Lokálne priestorové maximá a lokálne minimá v obraze, ktorých plocha je malá v porovnaní s požadovanou veľkosťou segmentovaného superixela (bez ohľadu na zodpovedajúcu hodnotu intenzity tohoto lokálneho extrému, budú považované za irelevantné a nebudú prispievať k segmentácii.

Nový prístup na báze morfológického spracovania. Navrhovaný prístup na báze morfológického spracovania pracuje s 8-súvislosťou, čím je garantovaná priestorová koherencia. Kľúčová otázka je, ktoré hrany sú výrazné hrany a majú prispievať k tvorbe hrán superpixelov a ktoré sú nepodstatné, a teda majú byť ignorované? Pre riešenie tohoto problému sú v prvom kroku vyhladené malé lokálne extrémny v obraze a tým odstránené hrany s lokálnou signifikanciou. Algoritmus potom pozostáva z ďalej uvedených základných krokov.

4.17. Príklad použitia morfológických operácií pre segmentáciu superpixelov

- Najprv využijeme morfológickú rekonštrukciu obrazu na odstránenie irelevantných lokálnych extrémov v obraze s cieľom odstrániť nepodstatné hrany.
- Potom generujeme značky pre značkovú transformáciu rozvodím (Watershed transform). Za týmto účelom použijeme sofistikovaný postup zaplavenia vo vstupnom obraze vytvorenom ako rozdiel medzi pôvodným obrazom a obrazom filtrovaným Gaussovým filtrom.
- V ďalšom kroku prebehne výpočet samotnej morfológickej značkovej transformácie rozvodí.

4.17.1 Vyhľadanie morfológickou rekonštrukciou

Použitá metóda bola uvedená v [Vin93a] a je popísaná v kapitole 4.15.

Výber vhodnej metódy generovania značiaceho obrazu silne závisí na požadovanej aplikácii. V tejto navrhovanej metóde je značiaci obraz generovaný morfológickou eróziou. Parameter s , ktorý udáva veľkosť štruktúrneho elementu, môže byť priamo odvodený z maximálneho priemeru oblastí lokálnych extrémov, ktoré majú byť odstránené. Experimentálne bol stanovený tento parameter na hodnotu $s = 7$, pričom samotný výpočet prebehol 6-krát: 3-krát pre jednotlivé R,G,B kanály (odstránenie lokálnych maxím) a 3-krát pre jednotlivé R,G,B kanály invertovaného obrazu (odstránenie lokálnych miním). Príklad výstupu morfológickej rekonštrukcie je na nasledujúcom obrázku 4.35.



Obrázok 4.35: Vľavo je vstupný obraz. Vpravo je obraz vyhladený morfológickou rekonštrukciou.

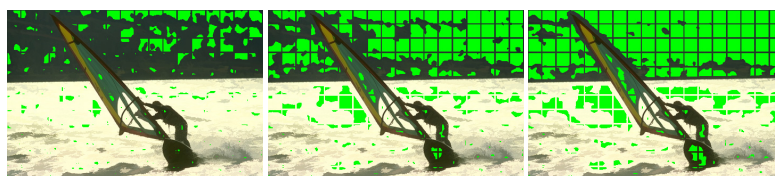
4.17.2 Generovanie značiek – semiačok pre značkovú transformáciu rozvodím

Semiačko je značka počiatku rastúcich regiónov. V anglickej literatúre sa semiačka označujú ako termínom *seeds*. Superpixely musia na jednej strane rešpektovať požiadavku na ich pravidelné rozmiestnenie a zároveň sledovať významné hrany objektov na strane druhej. Jednou z možností je generovanie značiek v pravidelnej mriežke. Túto štandardnú metódu zároveň využijeme ako referenčnú metódu pre vyhodnotenie tejto navrhutej metódy.

Vstupom je vyhladený obraz, ktorý sme získali v predchádzajúcom kroku. Hranový obraz v ďalšom kroku vytvoríme ako rozdiel obrazu po rozmazaní Gaussovým filtrom a pôvodným obrazom. Výstup ďalej slúži na tvorbu značiek nasledovným postupom

- Najskôr je každé semiačko umiestnené do stredu pravidelne vytvorenej mriežky.
- Metódou rastúcich regiónov s použitím 8-súvislosti v obraze F sú tieto oblasti riadené parametrom p , ktorý určuje kedy sa zastaví rast regiónu. Zároveň je rast obmedzený pravidelne vygenerovanou mriežkou. Parameter p je vlastne maximálna tolerovaná diferenciacia vo farbe medzi pixelom a jeho okolím patriacim do jeho 8-susedného okolia vo funkcii pre výpočet rastúcich regiónov.

Vplyv rôznych hodnôt parametra p je demonštrovaný na nasledujúcich obrázkoch 4.36.

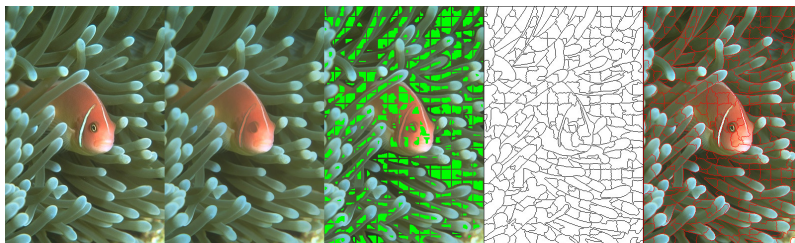


Obrázok 4.36: Vplyv rôznych hodnôt parametra p . Hodnota parametra na obrázkoch zľava doprava: 0, 1, 2.

4.17.3 Transformácia rozvodím

V ďalšom kroku bola počítaná transformácia rozvodí pre farebný vstupný obraz tak, ako bola popísaná v odstavci 4.13.1. Sekvencia obrázkov 4.37 demonštruje jednotlivé kroky algoritmu.

4.17. Príklad použitia morfológických operácií pre segmentáciu superpixelov



Obrázok 4.37: Zľava doprava: vstupný obraz, obraz po morfológickej rekonštrukcii, značky pre transformáciu rozvodia (zelené), hranice superpixelov po transformácii rozvodia, hranice superpixelov znázornené v obraze.

4.17.4 Vyhodnotenie

Vyhodnotenie obnovenia hraníc (Boundary Recall)

Pre vyhodnotenie obnovenia hraníc (Boundary Recall) je potrebná referencia ručne segmentovaných hraníc objektov (saliency edges), ktoré majú byť sledované hranicami superpixelov. Použitý súbor testovacích obrazov „Berkeley Segmentation Dataset and Benchmark“ [Mar+01] tieto referenčné dáta poskytuje. Pretože aj pri manuálnej segmentácii objektov sa opakované pokusy neprekrývajú s presnosťou na jeden pixel, je vhodné zaviesť určitú toleranciu i pre vyhodnocovanie. Štandardne sa pri vyhodnocovaní tohoto testovacieho súboru používa tolerančná odchýlka $k = 2$ pixely.

Potom počítame vyhodnotenie pre obnovenie hranice (Boundary recall) BR podľa vzťahu

$$BR = \frac{TP}{TP + FN} \quad (4.41)$$

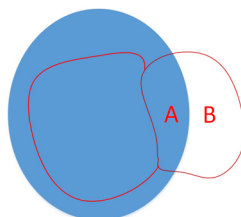
Pričom hodnoty TP (skratka z angl. true positive) a FN (skratka z angl. false negative) získame nasledovne.

TP sú tou časťou referenčných dát (ground truth), ktoré sa nachádzajú s uvedenou toleranciou $k = 2$ pixely v Euklidovskej vzdialenosti od najbližšej hranice vytvorených superpixelov.

FN sú tou časťou referenčných dát (ground truth), ktoré nemajú s uvedenou toleranciou $k = 2$ pixely v Euklidovskej vzdialenosti žiadne hranice vytvorených superpixelov.

Chyba podsegmentovania (Undersegmentation Error)

Chyba podsegmentovania (Undersegmentation Error) je vysvetlená pomocou obrázku 4.38.



Obrázok 4.38: Chyba podsegmentovania. Modrou farbou je znázornený objekt, červenou farbou sú znázornené hranice superpixelov.

Modrý objekt z referenčného súboru je chybné rozdelený do znázorneného superpixelu, pričom vzniknú dve časti: vnútorná A a vonkajšia B. Ako chybový údaj je vyhodnotená tá plocha vnútornej alebo vonkajšej časti superpixela, ktorá generuje menšiu chybu. Celkovo sa potom vyhodnocuje podiel plôch v pomere: súčet chybových plôch ku celkovej ploche segmentu z referenčného súboru.

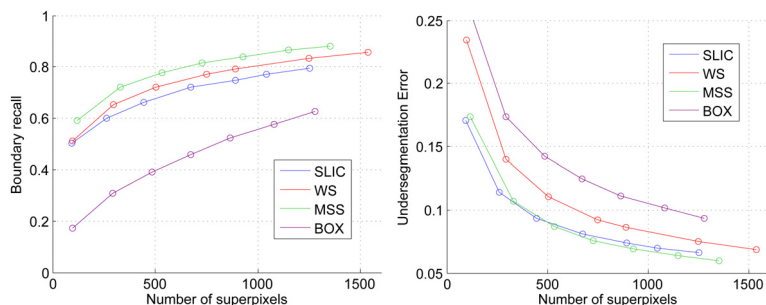
4.17.5 Výsledky

Prezentovaná morfológická metóda na výpočet superpixelov (MSS) bola vyhodnotená s použitím nástroja Superpixel Benchmark Toolbox [NP12], ktorý umožní porovnávanie úspešnosti akejkoľvek novej metódy s najvýznamnejšími, doteraz publikovanými, metódami (vid' rešerš v kapitole 4.17). Metóda bola vyhodnotená s použitím uvedených metrík uvedených v odstavci 4.17.4 a to v závislosti na počte superpixelov v obraze ako premenlivého parametra.

Na grafoch v obrázku 4.39 sú znázornené výsledky porovnania popísanej metódy (MSS) s metódou SLIC, transformáciou rozvodia s rovnomernou mriežkou (WS – regular grid watershed segmentation) a s jednoduchým generovaním superpixelov v tvare pravidelnej mriežky (BOX).

Z grafov na obrázku 4.39 vidíme, že výsledky pomocou popísanej metódy sú konzistentne lepšie v porovnaní s ostatnými metódami vrátane metódy SLIC, ktorá patrí v súčasnosti medzi najprominentnejšie. Po vyhodnotení časovej výpočtovej náročnosti vidíme, že metóda MSS je tiež

4.17. Príklad použitia morfológických operácií pre segmentáciu superpixelov



Obrázok 4.39: Vľavo sú výsledky vyhodnotenia obnovenia hraníc (Boundary recall). Vpravo je vyhodnotená chyba podsegmentovania (Undersegmentation error).

konzistentne rýchlejšia než metóda SLIC, ktorá patrí v súčasnosti medzi najrýchlejšie.

Rozlíšenie obrazu	SLIC	MSS
192 × 108	33 ms	26 ms
384 × 216	127 ms	33 ms
960 × 540	766 ms	118 ms
1344 × 756	1511 ms	228 ms
1920 × 1080	3132 ms	470 ms

Tabuľka 4.1: Priemerný čas potrebný pre výpočet superpixelov (Intel Core i3 processor clocked at 2.27 GHz).

Popísaná metóda demonštruje možnosti uplatnenia morfológických operácií a predovšetkým ich kombinovanie pre riešenie zložitejších úloh.



Zoznam použitej literatúry

- [Ber30] D. Bernoulli. „Fourier Series“. (1830), str. 1–10.
- [BRJ06] W. Benesova, A. Rinnhofer a G. Jakob. „Determining the Average Grain Size of Super-Alloy Micrographs“. *IEEE International Conference on Image Processing*. IEEE 1-4244-0480-0, 2006, str. 2749–2752.
- [CMM02] D. Comaniciu, P. Meer a S. Member. „Mean shift: A robust approach toward feature space analysis“. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24 (2002), str. 603–619.
- [CT65] W. J. Cooley a J. W. Tukey. „An algorithm for the machine calculation of complex Fourier series“. *Math. Comp* (1965), str. 297–301.
- [GW01] R. Gonzalez a R. Woods. *Digital Image Processing*. Druhé vyd. Prentice Hall, 2001.
- [GW08] R. C. Gonzalez a R. E. Woods. *Digital Image Processing*. Tretie vyd. Prentice Hall, 2008.
- [Hla12] V. Hlaváč. „Fourierova transformace v 1D a 2D“. *Prednáška* (2012).
- [HS05] V. Hlaváč a M. Sedláček. *Zpracování signálů a obrazů, skriptum*. Vydavatelství ČVUT, 2005.

- [HS87] R. M. Haralick a S. R. Sternberg. „Image Analysis Morphology“. *IEEE transactions on pattern analysis and machine intelligence* July (1987), str. 532–550.
- [JB89] L. Jaroslavskij a I. Bajla. *Metódy a systémy číslicového spracovania obrazov*. Alfa, Bratislava, 1989.
- [KR96] G. Kotuliaková a J. Rozinaj. *Číslicové spracovanie signálov*. Bratislava: FABER, 1996.
- [Lev+09] A. Levinshtein a kol. „TurboPixels: Fast Superpixels Using Geometric Flows“. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31.12 (2009), str. 2290–2297.
- [Lih02] J. Lihov. „Algebra a jej vety II.“ (2002), str. 1–29.
- [Liu+11] M.-Y. Liu a kol. „Entropy rate superpixel segmentation“. *Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition*. CVPR '11. Washington, DC, USA: IEEE Computer Society, 2011, str. 2097–2104.
- [Mar+01] D. Martin a kol. „A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics“. *Proc. 8th Int'l Conf. Computer Vision*. Zv. 2. Júl 2001, str. 416–423.
- [Mey92] F. Meyer. „Color image segmentation“. *Image Processing and its Applications, 1992., International Conference on*. 1992, str. 303–306.
- [NP12] P. Neubert a P. Protzel. „Superpixel Benchmark and Comparison“. *Proc. of Forum Bildverarbeitung*. Regensburg, Germany. 2012.
- [NT02] R. Neumann a G. Teisseron. „Extraction of dominant points by estimation of the contour fluctuations“. *Pattern Recognition* 35 (2002), str. 1447–1462.
- [Ond+07] Ondráček O. a kol. *Biosignály*. Bratislava: Vydavateľstvo STU v Bratislave, 2007.
- [OS75] A. V. Oppenheim a R. W. Schaffer. *Digital signal processing*. Prentice-Hall, 1975.
- [PPP05] J. Pavlovičová, M. Partyk a J. Polec. *Číslicové spracovanie obrazu*. Bratislava, Vydavateľstvo STU v Bratislave, 2005.
- [RY00] K. Rao a P. Yip. *The transform and data compression handbook*. 2000.

- [SHB07] M. Sonka, V. Hlavac a R. Boyle. *Image Processing, Analysis, and Machine Vision*. Tretie vyd. Cengage Learning, 2007.
- [Vin93a] L. Vincent. „Morphological grayscale reconstruction in image analysis: applications and efficient algorithms“. *Image Processing, IEEE Transactions on* 2.2 (1993), str. 176–201.
- [Vin93b] L. Vincent. „Morphological Grayscale Reconstruction in Image Analysis: Applications and Efficient Algorithms“. *IEEE Transactions on Image Processing*, 2.2 (1993), str. 176–201.
- [VT96] M. Van Droogenbroeck a H. Talbot. „Fast computation of morphological operations with arbitrary structuring elements“. *Pattern Recognition Letters* 17.14 (dec. 1996), str. 1451–1460.
- [Vyr06] R. Výrut. „Minkowského suma“. *Geometry and Computer Graphics*. Nové Město na Moravě, 2006.
- [ZK04] R. Zabih a V. Kolmogorov. „Spatially Coherent Clustering Using Graph Cuts“. *CVPR (2)*. 2004, str. 437–444.

ČASŤ II

PRÍZNAKY

Príznamy v oblasti počítačového videnia

V tejto časti knihy sa budeme zaoberať príznakmi a ich prípravou na proces klasifikácie. Objekt v oblasti počítačového videnia je útvar reprezentovaný oblasťou v segmentovanom obraze. Príznakom sa rozumie výsledok merania, ktoré kvantifikuje nejakú vlastnosť objektu. Viacero jednotlivých príznakov sa spája do príznakového vektora (vektora príznakov). Práca s príznakmi uľahčuje následné úlohy klasifikácie, lebo znižuje objem spracovávaných dát. Ak by sme obraz s rozmermi 1920×1080 pixelov reprezentovali ako jeden vektor intenzít, jeho rozmer by bol 2073600. Ak by sme tento obraz opísali napríklad 256-triednym histogramom intenzít, znížime objem spracovávaných údajov 8100-krát.

Typy príznakov môžeme určovať podľa niekoľkých kritérií. Na základe úrovne abstrakcie definujeme nasledovné triedy príznakov.

Nízkoúrovňové príznaky opisujú základné vlastnosti objektov, ako tvar, farbu, textúru, hrany, rohy atď.

Strednoúrovňové príznaky vzniknú spojením nízkoúrovňových príznakov, ako napríklad kontrast, použitá paleta farieb alebo súpis základných tvarov v obraze. Spojením informácie o farbe objektu, textúre, umiestnení a pod. vieme určiť sémantické kategórie (voda, sneh, skala, vegetácia, ...). Takisto využitím nízkoúrovňových príznakov vieme detegovať napríklad ľudské tváre.

Vysokoúrovňové príznaky sú príznaky najvyššej úrovne, ktoré opisujú sémantiku zobrazenej scény – napríklad osoba sediaci na pláži (obraz obsahuje vodu, piesok, oblohu a osobu v určitej vzájomnej konštelácii).



Obrázok 5.1: Príklad segmentácie obrazu. Zľava pravidelná segmentácia bez prekryvu, pravidelná segmentácia s prekryvom v strede, nepravidelná segmentácia s prekryvom v strede a segmentácia podľa obsahu obrazu.

Nízkoúrovňové príznaky môžeme deliť na bodové, lokálne a globálne. Bodové príznaky opisujú každý bod pozíciou, intenzitou alebo farbou. Globálne príznaky opisujú celý obraz alebo časti (objekty) vzniknuté segmentáciou obrazu. Segmentácia môže byť pravidelná s prípadným prekryvom, alebo založená na obsahu obrazu (obr. 5.1). Medzi globálne príznaky patria napríklad dominantná farba, textúra, tvar oblasti atď., s ktorými sa bližšie oboznámime v kapitolách 6, 7 a 8. Lokálne príznaky, ktorým je venovaná kapitola 9, opisujú body a ich najbližšie okolie.

Ďalej si ukážeme, ako príznaky pripraviť na proces klasifikácie (kap. 10) a ako určiť vhodné príznaky (kap. 11 a kap. 12).

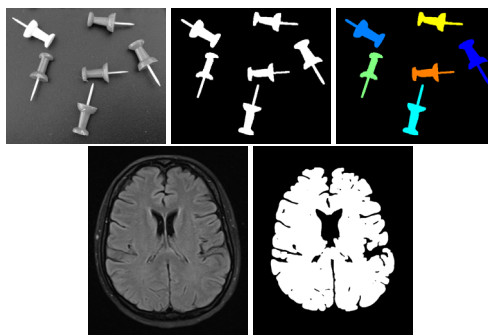
Vektor príznakov ako deskriptor objektu by mal spĺňať určité podmienky:

- dva rovnaké objekty musia mať rovnaké deskriptory a naopak, dva rovnaké deskriptory musia opisovať rovnaké objekty,
- podobné objekty musia byť klasifikovateľné do jednej triedy, ak majú podobné (blízke) deskriptory,
- príznaky by mali byť invariantné. Najmä voči rotácii, posunutiu, škálovaniu. Invariantnosť voči afinným a perspektívnym transformáciám je vhodná najmä, ak máme rozpoznať objekty snímané z rôznych zorných uhlov,
- deskriptor by mal byť kompaktný. Mal by obsahovať najmenšiu množinu informácií, ktorá je potrebná na identifikáciu daného objektu a jeho odlíšenie od objektov iných tried.

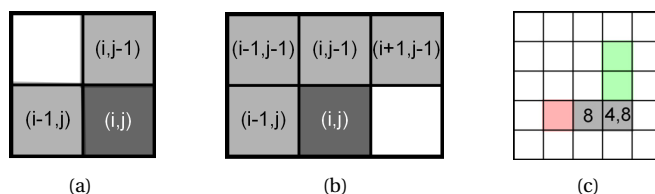
Príznamy tvaru

6.1 Binárny obraz

Príznamy tvaru sa najľahšie hľadajú v binárnych obrazoch, ktoré dostaneme po segmentácii obrazu na objekty (obr. 6.1). Jednotlivé vysegmentované oblasti sú oddelené od pozadia, ktoré je reprezentované nulou. V prípade, že tieto oblasti netvorí jeden objekt, môžeme identifikovať jednotlivé objekty pomocou algoritmu 6.1. Tento algoritmus pracuje so 4- alebo 8-susednosťou (pozri kap. 4.2.1) a podľa toho určí použitú masku (obr. 6.2). Príklad jednotlivých krokov algoritmu je na obrázku 6.3.



Obrázok 6.1: Segmentácia obrazu na oblasti. V prvom prípade oblasti netvorí jeden objekt.



Obrázok 6.2: Masky používané v algoritme 6.1. (a) Maska pre 4-susednosť. (b) Maska pre 8-susednosť. (c) Kolízia pre masku 4- a 8-susednosti.

```

1. prechod
   skenuj obraz, nájdi nenulový bod obrazu P
   ak sú susedné pixely určenej masky nulové
       bodu P priradiť nové označenie
   inak
       ak všetky nenulové pixely majú rovnaké označenie
           bodu P priradiť toto označenie
       inak
           bodu P priradiť ľubovoľné označenie
           z hodnôt pod maskou
           zaznamenať kolíziu do tabuľky ekvivalencie

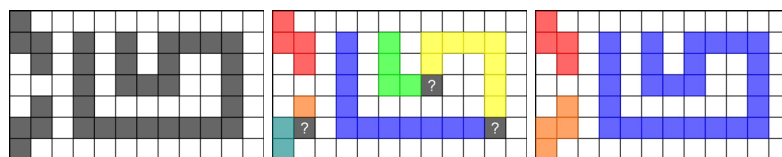
2. prechod
   skenuj obraz, zjednot' označenie podľa
   tabuľky ekvivalencie
  
```

Algoritmus 6.1: Identifikácia oblastí v binárnom obraze.

6.2 Opis tvaru založený na hranici oblasti

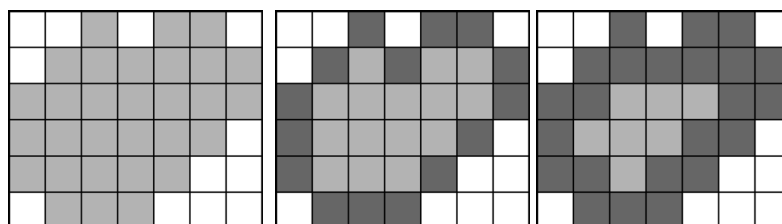
Oblasť je definovaná svojím vnútrom a hranicou. Bod oblasti leží na hranici, ak aspoň jeden z jeho susedov nepatrí oblasti. Hranicu väčšinou hľadáme algoritmom sledovania hranice, kde po nájdení prvého hraničného bodu postupujeme v smere (alebo proti smeru) hodinových ručičiek a označujeme ďalšie body hranice.

Na to, aby sme určili, či bod je hraničný, musíme vedieť, či je oblasť 4- alebo 8-susedná. Rôzne typy susednosti oblastí vedú k rôznym hraniciam. Susednosť hranice a oblasti je vždy komplementárna. To znamená, že hranica 4-susednej oblasti je 8-susedná a naopak, ako vidíme na obrázku 6.4.



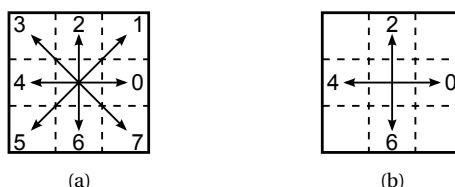
(a) Vstup do algoritmu – segmentované oblasti. (b) 1. prechod algoritmu s vyznačenými kolíziami. (c) 2. prechod algoritmu.

Obrázok 6.3: Ilustrácia vstupu a výstupov algoritmu 6.1 pri použití masky 6.2(a).



(a) Vysegmentovaná oblasť. (b) 8-susedná hranica 4-susednej oblasti. (c) 4-susedná hranica 8-susednej oblasti.

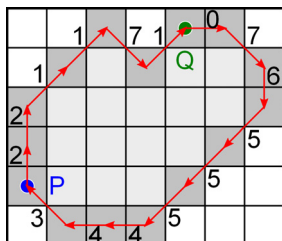
Obrázok 6.4: Oblasť a jej hranica pre rôznu susednosť oblasti.



Obrázok 6.5: Kódovanie použité vo Freemanovom kóde. (a) Kódovanie 8-susedného objektu. (b) Kódovanie 4-susedného objektu.

6.2.1 Reťazcový Freemanov kód

Pre 8-susednú hranicu kódujeme 8-okolie bodu číslami $0, \dots, 7$ ako na obrázku 6.5. Reťazcový (Freemanov) kód hranice objektu je potom postupnosť čísel, ktorá obsahuje informáciu, ktorým smerom hranica z daného bodu pokračuje. Reťazcový kód začína vo zvolenom začiatočnom bode. Príklad reťazcového kódu objektu na obrázku 6.6 začínajúci v bode P je 221171076555443.



Obrázok 6.6: Freemanov reťazcový kód hranice z obrázka 6.4(b).

Problémom reťazcového kódu je voľba začiatočného bodu a neinvariantnosť voči rotácii.

Prvý problém spôsobí, že dva rovnaké objekty budú mať rôzne reťazcové kódy, ak nezvolíme rovnaký začiatočný bod. Riešením je cyklický posun kódu tak, aby vzniklo najmenšie prirodzené číslo, čo v podstate znamená nájdenie referenčného začiatočného bodu hranice. Keď sa takto upravené reťazcové kódy dvoch objektov zhodujú, majú objekty rovnaký tvar. Pre oblasť na obrázku 6.6 je referenčným bodom bod Q a cyklicky posunutý kód je 076555443221171.

Invariantnosť voči rotácii rieši diferenčný reťazcový kód, ktorý nekóduje smer ako Freemanov kód, ale zmenu smeru. Nech $\{c_i\}_{i=1}^N$ je pôvodný Freemanov kód oblasti. Diferenčný kód $\{c'_i\}_{i=1}^N$ získame nasledovne

$$c'_i = \begin{cases} (c_{i+1} - c_i) \bmod 8 & \text{pre } 1 \leq i < N, \\ (c_1 - c_N) \bmod 8 & \text{pre } i = N. \end{cases} \quad (6.1)$$

Diferenčný reťazcový kód objektu na obrázku 6.6 začínajúci v bode P je 070627777007077.

Vhodnosť použitia Freemanovho alebo diferenčného kódu závisí od kvality segmentácie.

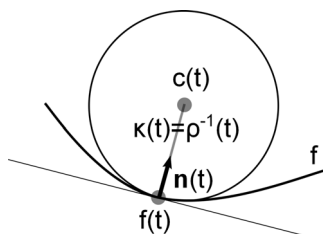
6.2.2 Obvod

Z Freemanovho kódu ľahko vypočítame obvod (perimeter) objektu

$$O = N_p + N_n * \sqrt{2}, \quad (6.2)$$

pričom N_p , resp. N_n je počet párnych resp. nepárnych číslíc v kóde. Predpokladáme štvorcový obrazový bod so stranou dĺžky 1.

Takto definovaná dĺžka hranice nie je invariantná voči otočeniu. Horizontálne, vertikálne a diagonálne hrany (so sklonom $\pm 45^\circ$) sú odme-



Obrázok 6.7: Krivosť v bode spojitely parametricky zadanej krivky sa rovná prevrátenej hodnote polomeru vpísanej kružnice v danom bode.

rané presne, ale dĺžka ostatných hrán vykazuje určitú chybu. Táto chyba je najväčšia pre hrany so sklonom $\pm 22,5^\circ$.

Tento problém sa dá riešiť úpravou váh a miernym zvýšením chyby dĺžky horizontálnych, vertikálnych a diagonálnych hrán

$$O = N_p * 0,948 + N_n * 1,340, \quad (6.3)$$

alebo pridaním počtu rohov N_r (miest, kde sa mení kód) do výpočtu

$$O = N_p * 0,980 + N_n * 1,406 - N_r * 0,091. \quad (6.4)$$

6.2.3 Krivosť krivky

Krivosť v bode spojitely parametricky zadanej krivky sa rovná prevrátenej hodnote polomeru vpísanej kružnice v danom bode (obr. 6.7). Rovnako krivosť vyjadruje zmenu v smere dotyčnice.

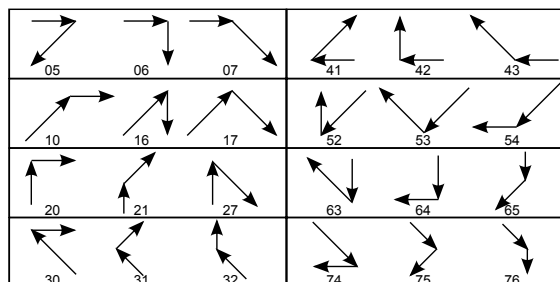
Algoritmy na počítanie krivosti kriviek v digitalizovaných obrazoch sú rôzne. Jednou možnosťou je určenie uhlovej zmeny pozdĺž hranice oblasti. Krivosť je potom daná vzťahom

$$\kappa(t) = \phi_{t+1} - \phi_{t-1}, \quad (6.5)$$

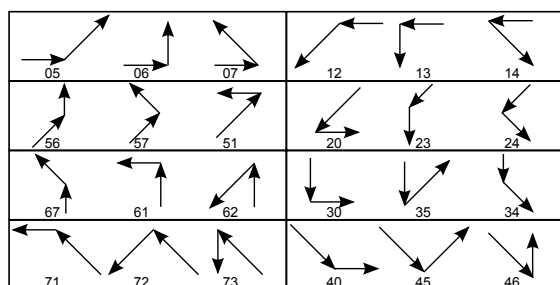
kde ϕ_t je smer gradientu množiny bodov definujúcich segment krivky. Smer gradientu určíme pomocou detektora hrán alebo priamo vypočítame pomocou vzťahu

$$\phi_t = \arctan\left(\frac{y_{t-1} - y_{t+1}}{x_{t-1} - x_{t+1}}\right), \quad (6.6)$$

kde (x_t, y_t) určuje bod na krivke.



(a)



(b)

Obrázok 6.8: Príznaky krivosti pre (a) konvexné a (b) konkávne uhly.**Príznaky krivosti**

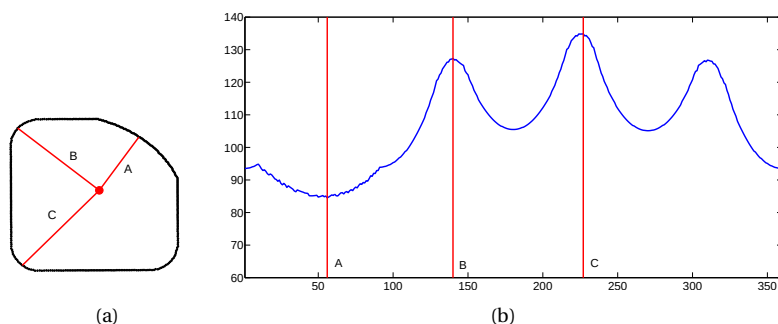
Tieto príznaky sa definujú pre konkávne a konvexné uhly medzi za sebou nasledujúcimi segmentami hranice. Na obrázku 6.8 sú uvedené možnosti. Máme teda 16 príznakov (8 pre konvexné a 8 pre konkávne uhly).

Priemerná energia zakrivenia

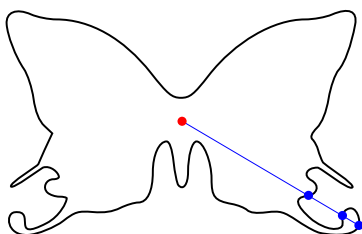
Tento pojem z teórie elasticity vyjadruje množstvo energie potrebné pri transformácii uzavretej krivky na kruh s rovnakým obvodom. Počíta sa ako

$$EZ = \frac{1}{L} \sum_{t=1}^L \kappa^2(t), \quad (6.7)$$

kde $\kappa(t)$ je krivosť hranice v bode t a L je dĺžka hranice.



Obrázok 6.9: Signatúra ako vzdialenosť jednotlivých bodov krivky od ťažiska. V útvere (a) sú vyznačené body, ktoré sú zvýraznené aj vo funkcii vzdialenosti (b).

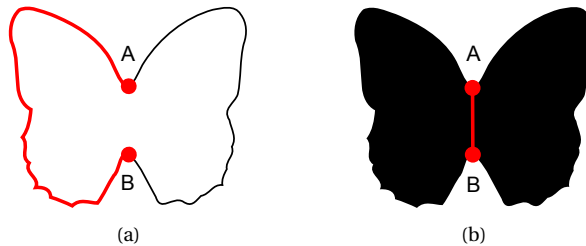


Obrázok 6.10: Mnohouholník nie je hviezdicovitý, preto niektorým uhlom zodpovedá viacero bodov krivky.

6.2.4 Signatúra

Signatúra je 1D signál opisujúci 2D krivku. Najpoužívanejšou signatúrou krivky je vzdialenosť jednotlivých bodov krivky od referenčného bodu ako funkcia uhla (obrázok 6.9). Týmto referenčným bodom býva napríklad ťažisko.

Takto definovaná signatúra je korektne zadaná, ak opisujeme hviezdicovitý mnohouholník a referenčný bod umiestnime do jeho jadra. Jadro mnohouholníka $M = V_1 \dots V_N$ tvoria body P , pre ktoré všetky úsečky PV_i , ležia vnútri mnohouholníka M . V prípade, že mnohouholník nie je hviezdicovitý, a teda jeho jadro je prázdne alebo, ak referenčný bod umiestnime mimo jadra, nastáva situácia nakreslená na obrázku 6.10. Aj vtedy môžeme použiť signatúru definovanú vzdialenosťou (maximálnou alebo minimálnou pre daný uhol), ale výsledná funkcia nebude spojitá.



Obrázok 6.11: Pri reprezentácii objektu pomocou hranice (a) je vzdialenosť bodov A a B polovica obvodu. Pri reprezentácii oblasťou (b) sú body blízko seba.

Medzi signatúry patria napríklad postupnosť x -ových a y -ových súradníc $x(n)$, $y(n)$, postupnosť komplexných čísel $z(n) = x(n) + i \cdot y(n)$, reťazcový kód, krivosť krivky.

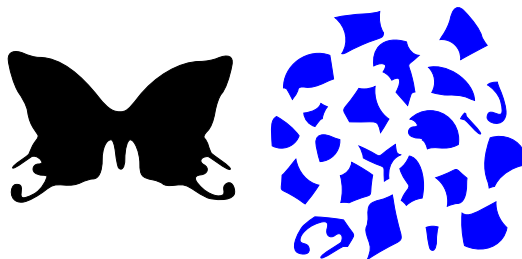
6.2.5 Fourierov opis

Názov Fourierov opis označuje množinu príznakov založených na Fourierovej transformácii (pozri kap. 1.2), ktoré sú jednými z najpopulárnejších spôsobov opisu tvaru. Spoločným prvkom týchto príznakov je, že skúmaný tvar je reprezentovaný pomocou 1D alebo 2D signatúry (pozri kap. 6.2.4), na ktorú sa aplikuje Fourierova transformácia. Z tejto sa potom vypočítajú Fourierove deskriptory. Deskriptor môžu tvoriť priamo koeficienty DFT, prípadne normalizované koeficienty DFT.

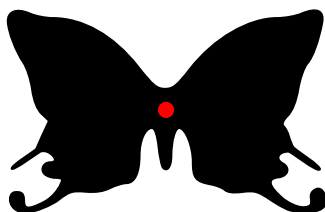
6.3 Opis tvaru založený na oblasti

Na obrázku 6.11 vidíme situáciu, ktorá ilustruje rozdiel medzi reprezentáciou oblasti pomocou hranice a celkovej oblasti (hranica a vnútro). Ak uvažujeme iba o hranici objektu, body A a B sú od seba vzdialené o polovicu obvodu, ale v rámci oblasti sú tieto body blízko seba. Niektoré z príznakov uvedených v tejto kapitole sa dajú počítať z hraničných bodov (napr. ťažisko, priemer), ale zvyčajne sa počítajú z celej oblasti. Hľadanie hranice len kvôli týmto príznakom by bolo krokom navyše.

Medzi jednoduché opisy tvaru patria príznaky odvodené z priemeru a obsahu, príznaky súvisiace s osami objektu alebo opísané obdĺžniky. Tieto príznaky môžu jednoducho identifikovať chybu produktu na výrobnnej linke, ale na komplexnú klasifikáciu objektov nie sú postačujúce, pretože viaceré objekty rôzneho tvaru môžu mať rovnaký priemer a ob-



Obrázok 6.12: Oba objekty (čierny aj modrý) majú rovnaký obsah.



Obrázok 6.13: Ťažisko objektu.

sah (pozri obr. 6.12). Vhodnejšími príznakmi sú kostra alebo štatistické momenty.

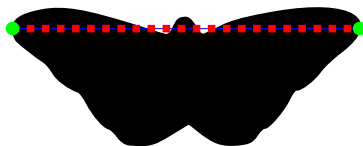
6.3.1 Ťažisko a k nemu prislúchajúce príznaky

Ťažisko

Ťažisko objektu (nazývané aj centroid) je bod, ktorého súradnice sa vypočítajú ako priemer x -ových a y -ových súradníc pixelov objektu. Súradnice ťažiska nemusia byť nutne celočíselné. Ak potrebujeme identifikovať pixel zodpovedajúci geometrickému ťažisku, musíme jeho súradnice zaokrúhliť.

Vzdialenosť od ťažiska k hranici

Určité vlastnosti objektov opisujú aj vzdialenosti od ťažiska k hranici – minimálna D_{min} , maximálna D_{max} , priemerná D_{mean} , prípadne ich pomery alebo histogram vzdialeností. Tieto príznaky sa dajú použiť napríklad na rozlíšenie kruhových ($D_{max} \approx D_{min}$) a elipsových útvarov. Pomery $\frac{D_{max}}{D_{min}}$, $\frac{D_{max}}{D_{mean}}$ a $\frac{D_{min}}{D_{mean}}$ sú bez rozmeru (pozri kap. 10.1), takže sú škálovo invariantné. Príznaky zahŕňajúce D_{mean} sú menej náchylné na šum.



Obrázok 6.14: Priemer objektu.

6.3.2 Príznyaky odvodené z priemeru a obsahu

Obsah

Obsah objektu vyjadruje počet bodov patriacich objektu

$$A = \sum_{(x,y) \in \text{Img}} I(x, y), \quad (6.8)$$

kde $I(x, y) = 1$, ak bod patrí objektu a $I(x, y) = 0$, ak bod patrí pozadiu.

V prípade, že máme objekt reprezentovaný pomocou N hraničných bodov $\{x_i, y_i\}_{i=1}^N$, vypočítame obsah nasledovne

$$A = \frac{1}{2} \left| \sum_{i=0}^{N-1} (x_{i+1} + x_i)(y_{i+1} - y_i) \right|, \quad (6.9)$$

kde $(x_0, y_0) = (x_N, y_N)$. Iný vzorec na výpočet obsahu je

$$A = \sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 + \sum_{i=1}^N (y_i - \bar{y})^2}, \quad (6.10)$$

kde $(\bar{x}, \bar{y})$ sú súradnice ťažiska.

Pri nekonvexnom objekte a objekte s dierami sa musíme rozhodnúť, či do obsahu budeme započítavať aj diery a zálivy na okrajoch. Tento obsah tvorí rozdiel medzi obsahom konvexného obalu a čistým obsahom objektu.

Priemer

Priemer objektu (obr. 6.14) je maximálna vzdialenosť medzi bodmi objektu

$$P = \max_{Q, R \in \text{Obj}} d(Q, R). \quad (6.11)$$



Obrázok 6.15: Komplexnosť kruhu je 1. Komplexnosť druhého objektu je vyššia.

Efektívny priemer

Efektívny priemer je priemer kruhu, ktorý má rovnaký obsah A ako skúmaný objekt

$$d = 2\sqrt{\frac{A}{\pi}}. \quad (6.12)$$

Kruhovitosť

Kruhovitosť (alebo kompaktnosť) udáva, ako veľmi je objekt podobný kruhu a je daná pomerom obsahu objektu A a obsahu kruhu A_K s rovnakým obvodom O

$$C = \frac{A}{A_K}. \quad (6.13)$$

Keďže

$$A_K = \frac{O^2}{4\pi}, \quad (6.14)$$

tak

$$C = \frac{4\pi A}{O^2}. \quad (6.15)$$

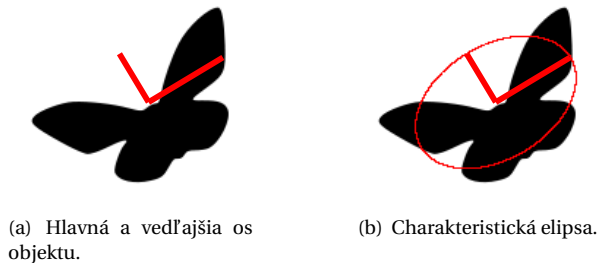
Kruhovitosť je maximálna pre kruh ($C = 1$) a blíži sa k nule pre dlhé tenké objekty.

Komplexnosť

Miera komplexnosti objektu je často používaným príznakom, ktorého definícia však nie je jednotná. Viacerí autori používajú prevrátenú hodnotu kompaktnosti, ktorá stúpa, čím viac sa odlišuje objekt od kruhu (pozri obr. 6.15). Ďalším používaným príznakom je pomer obsahu a obvodu objektu, prípadne iné vzťahy funkcií obsahu a obvodu.

6.3.3 Osi oblasti

Dôležité vlastnosti objektu sa dajú opísať príznakmi založenými na vlastných číslach a vektoroch.



Obrázok 6.16: Súvislosť charakteristickej elipsy s hlavnou a vedľajšou osou objektu. Charakteristická elipsa má rovnaký centrálny moment druhého rádu (varianciu) ako opisovaný objekt.

Hlavná a vedľajšia os

Smer, v ktorom objekt dosahuje najväčšiu dĺžku, sa nazýva hlavná os. Na ňu kolmá je vedľajšia os (obr. 6.16(a)). Tieto smery sa dajú určiť vypočítaním vlastných vektorov kovariančnej matice, pričom za prvky náhodného vektora berieme body objektu. Podrobnejšie o výpočte vlastných vektorov píšeme v kapitole 12.1.1.

Charakteristická elipsa

Charakteristická elipsa má rovnaký centrálny moment druhého rádu (varianciu) ako opisovaný objekt (obr. 6.16(b)).

Z charakteristickej elipsy vieme odvodiť niekoľko príznakov, napríklad percento obsahu elipsy pokryté objektom

$$PO = \frac{A_{\text{Objekt} \cap \text{Elipsa}}}{A_{\text{Elipsa}}} \cdot 100 \%, \quad (6.16)$$

alebo percento nepokrytia objektu elipsou

$$NO = \frac{A_{\text{Objekt} - \text{Elipsa}}}{A_{\text{Objekt}}} \cdot 100 \%, \quad (6.17)$$

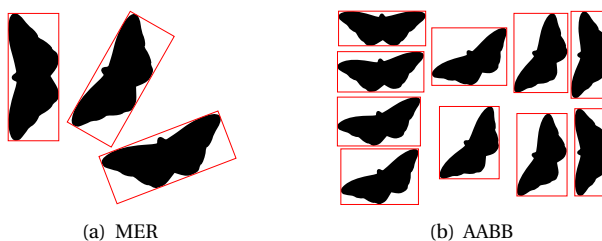
použiteľné napr. pri detekcii tvárí.

Pomer hlavnej a vedľajšej dĺžky

Hlavná a vedľajšia dĺžka sú dĺžky objektu v smere hlavnej a vedľajšej osi. Pomer týchto dĺžok je bezrozmerný príznak (zvaný aj elongácia alebo



Obrázok 6.17: Pomer hlavnej a vedľajšej dĺžky sa dá využiť napríklad na klasifikáciu druhov motýľov.



Obrázok 6.18: Opísaný obdĺžnik MER sa nemení s rotáciou objektu. Veľkosť osovo zarovnaného obdĺžnika AABB sa mení.

excentricita, po anglicky aspect ratio) opisujúci tvar objektu (pozri obrázok 6.17).

6.3.4 Opísaný obdĺžnik

Opísaný obdĺžnik je najmenší obdĺžnik, ktorý obsahuje celý objekt. Príznakom je veľkosť (šírka a dĺžka) obdĺžnika alebo ich kombinácia. Poznáme dva typy opísaných obdĺžnikov.

MER

Obdĺžnik na obrázku 6.18(a) sa nazýva minimálny opísaný obdĺžnik (minimal enclosing rectangle – MER). Jeho strany sú zarovnané s hlavnou a vedľajšou osou objektu.

AABB

Obdĺžnik na obrázku 6.18(b) má hrany rovnobežné s osami x a y . Nazýva sa osovo zarovnaný opísaný obdĺžnik (axis aligned bounding box – AABB). Ako vidíme tento príznak nie je invariantný voči rotácii objektu.

topológia

Obrázok 6.19: Objekt zložený z 11 súvislých oblastí (jednotlivé písmená a diakritika) so 6 dierami (písmená o, p, g, a). Eulerovo číslo tohto objektu je 5.

Hranatosť

Pomocou MER vieme určiť hranatosť objektu, čiže ako veľmi sa podobá na svoj minimálny ohraničujúci obdĺžnik. Hodnota hranatosti je daná pomerom obsahov objektu A a jeho MER A_{MER}

$$Hr = \frac{A}{A_{MER}}. \quad (6.18)$$

6.3.5 Topologické príznaky

Eulerovo číslo

Eulerovo číslo charakterizuje topológiu objektu, nie jeho tvar. Je dané rozdielom počtu súvislých oblastí objektu a počtu dier v objekte. Na obrázku 6.19 vidíme objekt zložený z 11 súvislých oblastí (jednotlivé písmená a diakritika) so 6 dierami (písmená o, p, g, a). Eulerovo číslo tohto objektu je 5.

Deravosť

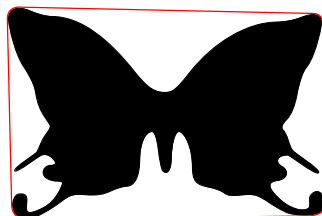
Diery v objektoch opisuje aj pomer ich obsahu A_D k obsahu vyplneného objektu A_{VO}

$$Der = \frac{A_D}{A_{VO}}. \quad (6.19)$$

Tento príznak sa dá využiť pri optickom rozpoznávaní textu.

Konvexnosť

Región je konvexný, ak pre každé dva body regiónu patrí regiónu aj ich spojnica. Konvexný obal objektu je najmenší konvexný útvar obsahujúci daný objekt (obr. 6.20).



Obrázok 6.20: Konvexný obal objektu.

Tvar objektu môžeme charakterizovať vzťahom

$$Conv_1 = \frac{O_{CH}}{O}, \quad (6.20)$$

kde $Conv_1$ charakterizuje „mieru konvexity“ objektu. Čím je objekt „konvexnejší“, tým viac sa jeho obvod O približuje obvodu konvexného obalu O_{CH} .

Podobným príznakom opisujúcim „mieru konvexity“ je pomer obsah objektu A a konvexného obalu A_{CH}

$$Conv_2 = \frac{A}{A_{CH}}. \quad (6.21)$$

Oba príznaky nadobúdajú hodnoty z intervalu $(0, 1)$. Konvexný útvar má oba príznaky rovné 1.

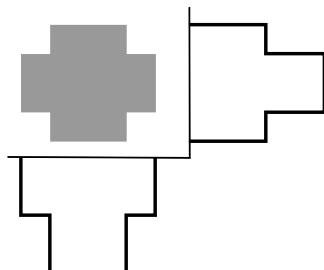
6.3.6 Morfologické príznaky

Kostra

Medzi hlavné morfologické príznaky opisujúce tvar objektu patrí kostra (pozri kap. 4.12). Jej problémom však je, že šum na hranici objektu veľmi ovplyvňuje výsledný tvar kostry.

Hrúbka

Kostru môžeme vytvárať pomocou vzdialenostnej transformácie alebo postupnou eróziou objektu. Hrúbka objektu je potom daná počtom erózií potrebných na zmiznutie objektu alebo maximálnou hodnotou vzdialenostnej transformácie.



Obrázok 6.21: Horizontálna a vertikálna projekcia jednoduchého objektu.

6.4 Šedotónový obraz

So šedotónovým obrazom sme sa zoznámili v kapitole 4.2. Zo šedotónového obrazu vieme získať bohatšiu informáciu o tvare objektu ako z binárneho.

6.4.1 Projekcie

Projekcia (alebo profil) je definovaná ako súčet intenzít v danom smere. Väčšinou sa používajú projekcie v x -ovom a y -ovom smere (obr. 6.21), z ktorých vieme určiť šírku a výšku AABB.

$$Proj_x(i) = \sum_{j=1}^N I(i, j), \quad (6.22)$$

$$Proj_y(j) = \sum_{i=1}^M I(i, j), \quad (6.23)$$

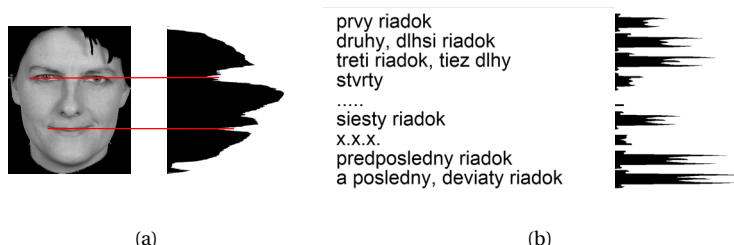
kde $M \times N$ sú rozmery obrazu a $I(i, j)$ sú intenzity (v binárnom obraze $\{0, 1\}$).

Pomocou projekcií sa dá napríklad určiť umiestnenie očí vo vysegmentovanej tvári (obr. 6.22(a)), alebo vymedziť riadky pri rozpoznávaní textu (obr. 6.22(b)).

6.4.2 Opisné štatistiky

Histogram

Histogram intenzít bodov obrazu je funkcia, ktorá pre každú úroveň intenzity udáva počet pixelov v obraze s touto intenzitou. Aby sme mohli



Obrázok 6.22: Použitie projekcií na (a) určenie umiestnenia očí a úst vo vysegmentovanej tvári a (b) vymedzenia riadkov pri rozpoznávaní textu.

pomocou histogramu porovnávať rôzne veľké obrazy, používame normalizované hodnoty relatívnych početností hodnôt intenzít v danom obraze

$$P(i) = N(i) / N, \quad (6.24)$$

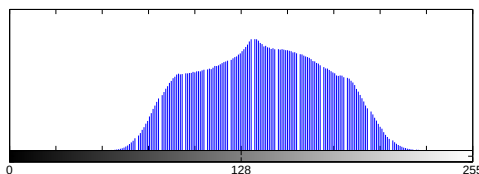
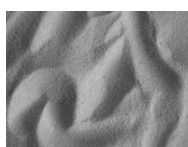
kde $N(i)$ je počet pixelov s intenzitou i a N je počet všetkých bodov obrazu.

Z tvaru histogramu vieme určiť napríklad, či je obraz nízko kontrastný (histogram je „úžky“, t. j. zaberá len malý interval intenzít – obr. 6.23(a)) alebo určiť objekt v popredí (ak je histogram bimodálny – objekt je tvorený pixelmi s tou istou hodnotou intenzity, pozadie zasa pixelmi s inou konštantnou intenzitou – obr. 6.23(b)). Ďalšie vlastnosti môžeme opísať štatistickými momentami.

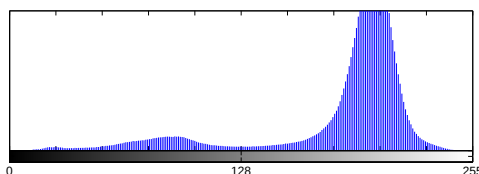
Štatistické momenty

Štatistické momenty opisujú tvar objektu. Detailnejšie sú opísané v dodatkoch A.3.1 a A.5.1. Momenty nižších rádoov opisujú geometrické vlastnosti objektu – ťažisko, obsah, zmenu tvaru. Tieto momenty sa počítajú priamo z objektu, prípadne z jeho hranice. Momenty vyšších rádoov sa používajú na opis projekcií objektu (šikmosť a špicatosť).

Existuje niekoľko metód ako kombinovať momenty, aby sme dostali príznaky invariantné voči rotácii, posunutiu a škálovaniu. Najznámejšími sú Huove momenty, ktoré sú kombináciou normalizovaných centrálnych momentov (pozri kap. A.5.1). Poznáme aj rotačné momenty, ktoré sú definované v polárnych súradniciach a ortogonálne momenty, ktoré miesto neortogonálnej bázy $x^p y^q$ používajú ortogonálne Legendreove, Čebyševove alebo Zernikeove polynómy.



(a) Histogram málo kontrastného obrazu zaberá len malý interval intenzít.



(b) Pomocou bimodálneho histogramu vieme oddeliť objekt od pozadia.

Obrázok 6.23: Použitie histogramu na určenie vlastností obrazu, alebo objektov v obraze.

6.4.3 Energia a entropia

Histogram intenzít alebo projekcie môžeme charakterizovať aj energiou E a entropiou H (kap. 11.5.4)

$$E = \sum_{i=1}^B (P(i))^2, \quad (6.25)$$

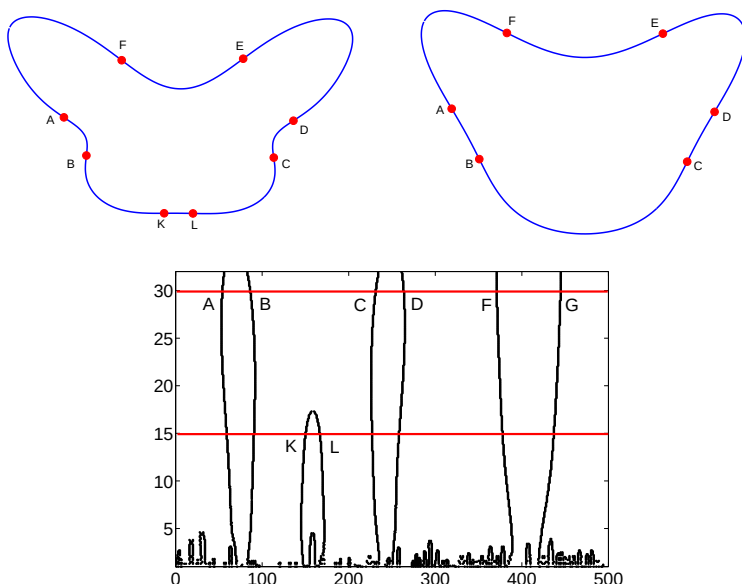
$$H = - \sum_{i=1}^B P(i) \log_2 P(i). \quad (6.26)$$

6.5 MPEG-7

Norma MPEG-7 definuje niekoľko príznakov opisujúcich tvar objektu. Dva opisujú 2D objekty a dva 3D objekty. V tejto kapitole si priblížime deskriptor kontúry (ContourShape) a deskriptor regiónu (RegionShape) pre 2D objekty.

6.5.1 Tvar kontúry

Deskriptor tvaru kontúry je založený na reprezentácii kontúry v škálovom priestore krivosti (Curvature Scale Space – CSS). Táto reprezentácia je veľmi kompaktná – priemerná veľkosť je 14 bytov.



Obrázok 6.24: Deskriptor tvaru kontúry. V hornom riadku je kontúra v dvoch štádiách vyhladenia s vyznačenými inflexnými bodmi. Zodpovedajúce hodnoty σ sú v CSS diagrame vyznačené červenou. Vidíme, že body K a L postupným vyhladením zanikli.

Kontúru objektu postupne vyhladzujeme Gaussovým filtrom, skúma-me krivosť výslednej krivky a zaznamenávame inflexné body kontúry do CSS diagramu. V CSS diagrame sú na horizontálnej osi indexy bodov na kontúre ($1, \dots, N$) a na vertikálnej osi jednotlivé hodnoty σ . Výsledný deskriptor tvoria potom lokálne maximá grafu.

So zvyšujúcim sa parametrom σ Gaussovho jadra sa krivka postupne vyhladzuje a počet inflexných bodov klesá. Na obrázku 6.24 vidíme kontúru v dvoch štádiách vyhladenia s vyznačenými inflexnými bodmi. Vidíme, že body K a L postupným vyhladením zanikli. Zodpovedajúce hodnoty σ sú v CSS diagrame vyznačené červenou.

Nevýhodou tohto deskriptora je, že aj malé „vlnky“ na kontúre, ktoré ľudský vizuálny systém nevyhodnotí ako podstatné, dokážu vytvoriť veľké maximá v CSS diagrame. Tento problém sa rieši orezaním CSS diagramu.

6.5.2 Tvar regiónu

Deskriptor RegionShape používa množinu ART (Angular Radial Transform) koeficientov F_{nm} rádu n a m

$$F_{nm} = \int_0^1 \int_0^{2\pi} V_{m,n}(\rho, \theta) f(\rho, \theta) d\rho d\theta, \quad (6.27)$$

kde $f(\rho, \theta)$ je obrazová funkcia v polárnych súradniciach a $V_{m,n}(\rho, \theta)$ je separovateľná ART bázo­vá funkcia

$$V_{m,n}(\rho, \theta) = \frac{1}{\pi} e^{jm\theta} \cos(\pi n\rho). \quad (6.28)$$

Výsledný deskriptor používa 35 koeficientov F_{nm} ($n < 3, m < 12$) normalizovaných koeficientom $|F_{00}|$ a kvantovaných na 4 bity, spolu teda má 140 bitov.

Príznyaky farby

Farebný obraz je väčšinou reprezentovaný v trojrozmernom priestore farieb. Existujú rôzne priestory farieb (pozri kap. 20), každý z nich sa používa pri iných úlohách počítačového videnia.

7.1 Dominantná farba

Najjednoduchším opisom farebného obrazu je prevládajúca, dominantná farba. Dominantných farieb môže byť viac, ak každá pokrýva dostatočne veľkú časť obrazu. Deskriptorom dominantných farieb je väčšinou množina dvojíc (farba, percentuálne zastúpenie v obraze).

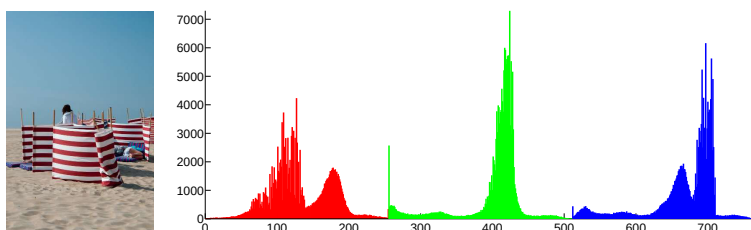
Problémom je, že nevieme, ako sú rovnako zafarbené body umiestnené v obraze.

7.2 Histogram farieb

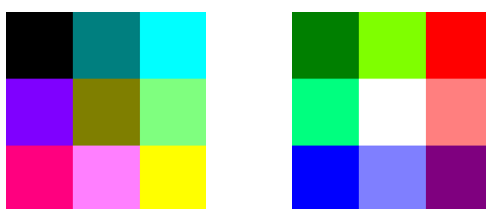
Na farebný obraz v priestore RGB sa môžeme pozerat' ako na 3 šedotónové obrazy. Z týchto potom extrahujeme príznaky opísané v kapitole 6.4 a zoradíme ich za seba.

V prípade histogramu farieb určíme histogramy jednotlivých farebných kanálov $P_c = [N_{c=0}, \dots, N_{c=255}]$, kde $c \in \{R, G, B\}$ (obr. 7.1). Výsledný histogram farieb potom bude

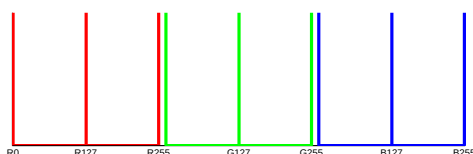
$$P_{RGB} = [P_R | P_G | P_B]. \quad (7.1)$$



Obrázok 7.1: Obráz a jeho histogram jednotlivých farebných kanálov.



(a) Dva obrázky obsahujúce rôzne farby.

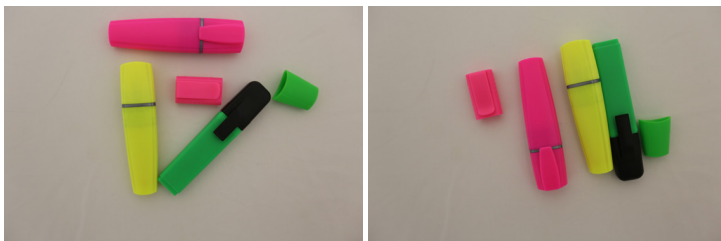


(b) Rovnaký histogram oboch obrazov.

Obrázok 7.2: Ilustrácia problému s histogramom farieb po zložkách. Obrazy obsahujúce rôzne farby môžu mať rovnaký histogram.

Nevýhodou tohoto prístupu je, že farebne rôzne obrázky môžu mať rovnaký deskriptor. Táto situácia je znázornená na obrázku 7.2, keď histogramy červenej, zelenej a modrej zložky obrazu sú rovnaké, ale farby na obrazoch sú rôzne. Riešením tohto problému je kvantizácia farieb (výber reprezentatívnych farieb napríklad pomocou metódy K priemerov – pozri kap. 15.1) a následné určenia histogramu jednotlivých farieb.

Ďalšou nevýhodou histogramu zložiek aj histogramu kvantovaných farieb je, že nevystihujú priestorové rozloženie farieb v obraze (obr. 7.3).



Obrázok 7.3: Dva obrazy obsahujúce rovnaké farby, ale na rôznych miestach obrazu.

7.2.1 Histogram priestorového rozloženia farieb

Aby sme opísali aj rozloženie farieb v obraze, použijeme histogram priestorového rozloženia farieb, kde pre každú farbu okrem počtu pixelov majúcej túto farbu, zaznamenáme aj priemerné umiestnenie $\mathbf{b}_i = (\bar{x}_i, \bar{y}_i)$ (priemer x -ovej a y -ovej súradnice bodov s danou farbou) a štandardnú odchýlku umiestnenia

$$\bar{x}_i = \frac{1}{N \cdot A_i} \sum_{c(\mathbf{p})=C_i} x, \quad (7.2)$$

$$\bar{y}_i = \frac{1}{M \cdot A_i} \sum_{c(\mathbf{p})=C_i} y, \quad (7.3)$$

$$\sigma_i = \sqrt{\frac{1}{A_i} \sum_{c(\mathbf{p})=C_i} d(\mathbf{p}, \mathbf{b}_i)}, \quad (7.4)$$

kde A_i je obsah oblasti majúcej farbu C_i , $\mathbf{p} = (x, y)$ je bod obrazu a $M \times N$ je rozmer obrazu.

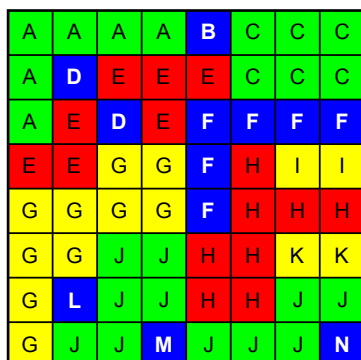
Ak je štandardná odchýlka malá, vieme, že farba je sústredená na malú oblasť obrazu. Ak je veľká, farba je rozmiestnená po celom obraze.

7.2.2 Ďalšie opisy priestorového rozloženia farieb

Ďašími alternatívami sú rozdelenie obrazu na bloky a extrakcia histogramu v týchto blokoch, prípadne quad-tree rozdelenie obrazu, kde každý list stromu má vlastný histogram.

Farebný korelogram

Farebný korelogram je podobný matici opakovaných výskytov (kap. 8.2). Do matice sa zapisujú počty výskytov dvojíc (c_i, c_j) farieb bodov v určenej



Obrázok 7.4: Štyri kvantované farby a určenie spojitých oblastí A–N.

Tabuľka 7.1: Počty bodov v oblastiach. Počet väčší ako prah je zvýraznený.

Región	A	B	C	D	E	F	G	H	I	J	K	L	M	N
Farba														
Veľkosť	6	1	6	2	7	6	10	8	2	11	2	1	1	1

vzdialenosti k

$$M_k(i, j) = |\{(\mathbf{p}) \mid f(\mathbf{p}) = c_i \ \& \ f(\mathbf{q}) = c_j \ \& \ |d(\mathbf{p}, \mathbf{q}) = k|\}|, \quad (7.5)$$

kde d je vzdialenosť bodov.

Vektor farebnej koherencie

Farebnú koherenciu definujeme ako stupeň príslušnosti farebných bodov k spojitkej oblasti. Bod, ktorý je svojou farbou osamotený, prípadne patrí do malého súvislého regiónu, je nekoherentný. Body, ktoré svojou farbou tvoria dostatočne veľké spojité oblasti, sú koherentné.

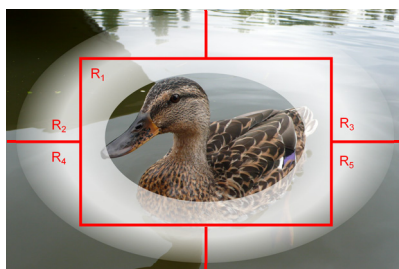
Výpočet vektora farebnej koherencie začína kvantovaním farieb do palety niekoľkých reprezentatívnych farieb. Potom sa určia spojité regióny jednotlivých farieb palety (obr. 7.4).

Následne spočítame body v jednotlivých regiónoch (tab. 7.1) a určíme, či je daný región koherentný alebo nie, podľa zadanej prahovej veľkosti regiónu. V našom príklade berieme prah 4.

Nakoniec pre každú farbu palety spočítame body v koherentných a nekoherentných regiónoch. Označme α_i počet pixelov v koherentných a β_i

Tabuľka 7.2: Počty koherentných a nekoherentných pixelov jednotlivých farieb.

Farba				
α	15	23	6	10
β	0	0	6	4

**Obrázok 7.5:** Príklad neurčitých regiónov. Oblasť neurčitosti príslušnosti k jednotlivým blokom je modelovaná elipsou.

v nekoherentných regiónoch farby i (tab. 7.2). Potom vektor farebnej koherencie je

$$\{(\alpha_1, \beta_1), \dots, (\alpha_K, \beta_K)\}, \quad (7.6)$$

kde K je počet farieb palety. V našom prípade $\{(15, 0), (23, 0), (6, 6), (10, 4)\}$.

Farebné momenty neurčitých regiónov

Obraz môžeme rozdeliť aj na neurčité (fuzzy) bloky, kde príslušnosť bodu k bloku nie je daná striktnie, ale pomocou funkcie príslušnosti $R_r \in \langle 0, 1 \rangle$. Príklad takéhoto rozdelenia je na obrázku 7.5. Oblasť neurčitosti príslušnosti k jednotlivým blokom je modelovaná elipsou.

Funkcia členstva sa potom použije aj pri extrakcii prvých troch centrálnych momentov pre jednotlivé farebné kanály. Regióny teda opisujú tri trojice (priemerná farba, variácia a šikmosť) váhované príslušnosťou k regiónu.

7.3 MPEG-7

Norma MPEG-7 definuje štyri deskriptory opisujúce farebnú informáciu v statických obrazoch. Ide o deskriptor dominantnej farby, škálovateľný

Tabuľka 7.3: Počty tried HSV histogramu

Počet tried	16	32	64	128	256
Odtieň	4	8	8	8	16
Saturácia	2	2	2	4	4
Hodnota	2	2	4	4	4

deskriptor farby, deskriptor rozloženia farieb a deskriptor štruktúry farieb. Farebné priestory, ktoré norma opisuje, sú *RGB*, *HSV*, *YCrCb* a nový *HMMD* (*hue-min-max-diff*) priestor, pričom *hue* má rovnaký význam ako v priestore *HSV*, *max* a *min* sú maximum a minimum *R*, *G*, *B* hodnôt a *diff* je rozdiel medzi týmito hodnotami.

7.3.1 Deskriptor dominantnej farby

Tento deskriptor opisuje *N* najdominantnejších reprezentatívnych farieb v obraze. Pre každú farbu obsahuje jej súradnice vo farebnom priestore a percentuálne zastúpenie v obraze. Reprezentatívne farby sú určené zhľukovacou metódou (napríklad metódou *K* priemerov – kap. 15.1).

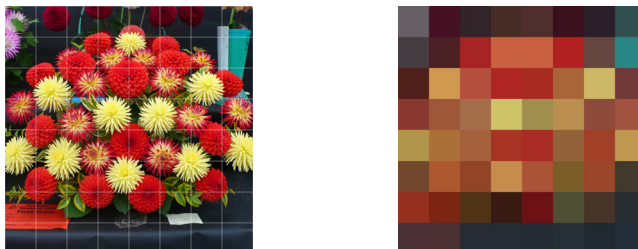
Okrem týchto základných informácií obsahuje deskriptor aj opis rozloženia farby v obraze, a to konkrétne informácie o variancii každej farby a priestorovej koherentnosti (ako „veľmi“ sú spojené body majúce dominantnú farbu). Norma nestanovuje, ako presne túto koherentnosť počítame. Môžeme použiť už spomenutý vektor farebnej koherencie.

Deskriptor sa používa najmä pri výbere a prezeraní obrazov z databázy podľa jednej, či viacerých významných farieb.

7.3.2 Škálovateľný deskriptor farby

Tento deskriptor kóduje histogram hodnôt *HSV* pomocou Haarovej transformácie.

Histogram má 256 tried. Odtieň je kvantovaný do 16, saturácia a hodnota do 4 tried. Hodnoty v histograme sú orezané na 11-bitové čísla. Tieto sú potom nelineárne namapované na 4-bitové čísla. Toto mapovanie preferuje nízke hodnoty, ktoré sa vyskytujú s vyššou frekvenciou. Táto reprezentácia je stále dosť priestorovo náročná (1024 bitov/histogram). Aby sme potrebné miesto zmenšili, je na tento histogram potom aplikovaná 1D Haarova transformácia (HT). Výsledkom je škálovateľný deskriptor, ktorý podľa počtu použitých koeficientov HT reprezentuje histogramy s rôznym počtom tried (pozri tabuľku 7.3).



Obrázok 7.6: Obráz rozdelený do 8×8 blokov a priemerná farba v blokoch.

1	2	3	4	5	6	7	8
9	10	11	12	13	14	15	16
17	18	19	20	21	22	23	24
25	26	27	28	29	30	31	32
33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48
49	50	51	52	53	54	55	56
57	58	59	60	61	62	63	64

Obrázok 7.7: Cik-cak prechod bloku koeficientov v deskriptore rozloženia farieb.

7.3.3 Deskriptor rozloženia farieb

Tento deskriptor sa počíta v priestore $YCbCr$ a používa sa hlavne pri vyhl'adávaní obrazov podľa náčrtku. Základom je vektor reprezentatívnych farieb. Obráz sa rozdelí na 8×8 blokov a vypočíta sa priemerná farba týchto blokov (obr. 7.6). Táto informácia je potom transformovaná 8×8 diskretnou kosínusovou transformáciou (pozri kap. 1.5) do 3×64 koeficientov DCT_Y , DCT_{Cb} , DCT_{Cr} . Výsledný deskriptor dostaneme cik-cak prechodom bloku koeficientov (obr. 7.7), aby sme koeficienty pre nízke frekvencie umiestnili vedľa seba.

7.3.4 Deskriptor štruktúry farieb

Tento deskriptor je generalizáciou histogramu farieb, ktorý obsahuje aj informáciu o rozložení farieb v obraze. Na opis štruktúry farieb využíva štrukturálny element rozmeru 8×8 a do histogramu zapisuje počet štrukturálnych elementov, ktoré danú farbu obsahujú.

Tabuľka 7.4: Počty tried HSV histogramu v deskriptore štruktúry farieb.

Komponent	Podpriestor <i>diff</i>	Počet kvantovacích úrovní pre K CSD tried			
		$K=184$	120	64	32
Odtieň	0	1	1	1	1
	1	8	4	4	4
	2	12	12	6	3
	3	12	12	4	2
	4	24	12	4	2
Intenzita	0	8	8	8	8
	1	4	4	4	2
	2	4	4	4	4
	3	4	4	4	2
	4	2	4	4	2

Obraz je prevedený do *HMMD* priestoru, a farby sú kvantované do palety veľkosti 32, 64, 120 alebo 184. Paletu vytvoríme nasledovne. Celý priestor *HMMD* je rozdelený na 5 podpriestorov podľa hodnôt súradnice *diff*. V každom z týchto podpriestorov sú súradnice *odtieň* a *intenzita* (definovaná ako $(\max + \min)/2$) kvantované podľa požadovaného počtu farieb v palete tak, ako je uvedené v tabuľke 7.4.

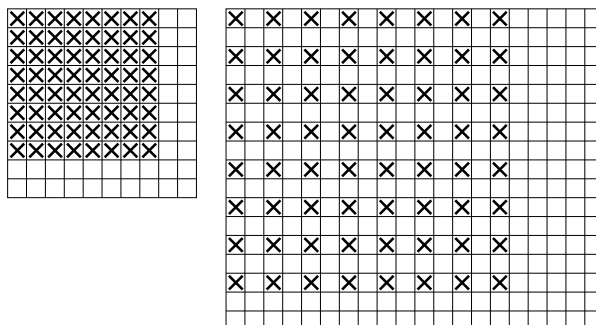
Štrukturálny element má konštantnú veľkosť pre rôzne veľké obrazy. Škálovaný je priestor, ktorý tento element pokrýva (obr. 7.8). Nech W a H sú šírka a výška obrazu, potom $E \times E$ je priestor, ktorý štrukturálny element pokrýva, kde

$$p = \max\{0; [0, 5\log_2 WH - 8]\}, \quad (7.7)$$

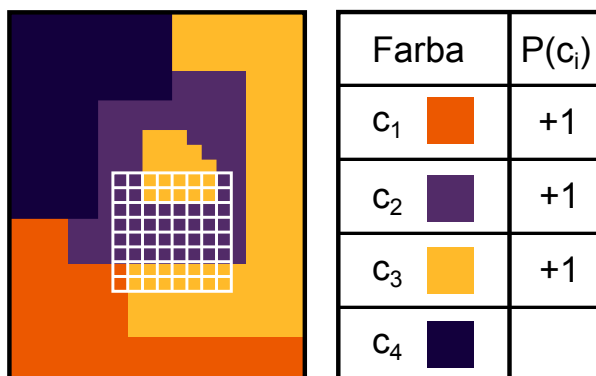
$$K = 2^p, \quad (7.8)$$

$$E = 8K. \quad (7.9)$$

Keď vieme, aký priestor štrukturálny element pokrýva, môžeme vypočítať histogram. Pre každú pozíciu štrukturálneho elementu zistíme, ktoré farby obsahuje. Týmto farbám potom zvýšime hodnotu v histograme (obr. 7.9). Výsledný histogram je následne normalizovaný počtom pozícií štrukturálneho elementu a nelineárne kvantovaný do 8 bitov/trieda.



Obrázok 7.8: Vľavo je štruktúrálly element veľkosti 8×8 pre malé obrazy. Pre väčšie obrazy (vpravo) sa priestor, pokrytý týmto elementom, škáluje.



Obrázok 7.9: Vytvorenie histogramu. Pre každú pozíciu štruktúrného elementu zistíme, ktoré farby obsahuje. Týmto farbám potom zvýšime hodnotu v histograme. Aj keď sa farba c_3 nachádza v štruktúrnóm elemente $2 \times$, histogram sa zvýši o 1.



Príznyaky textúry

Textúra je veľmi ťažko definovateľná vlastnosť obrazu. Pre nás textúra bude rozpoznateľná charakteristická variácia intenzít v regióne obrazu, ktorá sa dá opísať a dať sa podľa nej určiť hranice tohto regiónu.

Zmena intenzity môže byť rýchla, pomalá, môže mať prevládajúci smer, alebo môže byť porovnateľná vo všetkých smeroch, môže sa pravidelne opakovať alebo byť nepravidelná atď. Pravidelnosť je najpoužívanejšou vlastnosťou textúr.

Na opis textúr existuje mnoho metód využívajúcich napríklad rôzne filtre v priestorovej a frekvenčnej oblasti, autoregresívne modely, Mar-



Obrázok 8.1: Príklady rôznych textúr.

kovove náhodné polia, Gaborove filtre, koeficienty vlnkovej transformácie, metódy fraktálnej geometrie, priestorové vzťahy intenzít, lokálne binárne vzory (local binary patterns – LBP), Haarove koeficienty a iné.

8.1 Autokorelácia

Jedným z prvých prístupov k charakterizácii textúr bola autokorelácia, ktorá opisuje lokálne zmeny intenzity, ako aj opakovateľnosť vzorov. Nie je však schopná opísať mieru zrnitosti. Korelácia súvisí s konvolúciou (pozri kap. 2.1).

Autokorelácia je korelácia signálu s rovnakým rovnakým signálom

$$\rho(u, v) = \frac{MN}{(M-u)(N-v)} \frac{\sum_{x=1}^{M-u} \sum_{y=1}^{N-v} I(x, y) I(x+u, y+v)}{\sum_{x=1}^M \sum_{y=1}^N I^2(x, y)},$$

kde $I(x, y)$ je textúra rozmerov $M \times N$ a (u, v) posun v smere x a y .

8.2 Matica opakovaných výskytov

Nedostatky autokorelácie odstraňuje matica opakovaných výskytov (co-occurrence matrix, MOV). Ako už jej názov naznačuje, matica obsahuje počty výskytov dvojíc intenzít v určenom smere a vzdialenosti.

Matica je štvorcová, počet riadkov a stĺpcov sa rovná počtu úrovní jasu v obraze

$$M_{\mathbf{d}}(i, j) = |\{(x, y) \mid f(x, y) = i \ \& \ f(x + d_x, y + d_y) = j\}|, \quad (8.1)$$

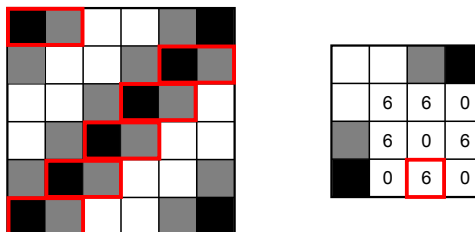
kde $\mathbf{d} = (d_x, d_y)$ je vektor posunutia. Na obrázku 8.2 je ukážka matice opakovaných výskytov pre jednoduchý príklad textúry. Vektor posunutia je $\mathbf{d} = (1, 0)$.

Pre praktické použitie sa MOV normalizuje počtom výskytov jednotlivých dvojíc a používa sa symetrická MOV daná ako

$$M_{\mathbf{d}}^S(i, j) = M_{\mathbf{d}}(i, j) + M_{-\mathbf{d}}(i, j). \quad (8.2)$$

Najčastejšie sa používajú vektory posunutia $(1, 0)$, $(1, 1)$, $(0, 1)$ a $(-1, 1)$.

Z takto vytvorených MOV sa následne extrahujú štatistické príznaky ako entropia, energia, momenty, inverzné momenty, maximálna hodnota, korelácia, kontrast, homogénnosť a ďalšie. Niektoré často používané príznaky sú uvedené v tabuľke 8.1, kde L je počet úrovní jasu v obraze.



Obrázok 8.2: Príklad pravidelnej textúry a jej matica opakovaných výskytov pre vektor posunutia $\mathbf{d} = (1, 0)$. Červenou je vyznačených 6 výskytov usporiadanej dvojice (čierna, šedá).

Tabuľka 8.1: Príznačky počítané z MOV rozmeru $L \times L$.

Kontrast	$C = \frac{1}{(L-1)^2} \sum_{k=0}^{L-1} k^2 \sum_{i,j; i-j =k} M_{\mathbf{d}}(i, j)$
Energia	$H = \sum_{i=1}^L \sum_{j=1}^L M_{\mathbf{d}}^2(i, j)$
Homogénosť	$S = \sum_{i=1}^L \sum_{j=1}^L \frac{M_{\mathbf{d}}(i, j)}{1 + i - j }$

8.3 Diskrétna Fourierova transformácia

Fourierove (kap. 1.2) fázové a amplitúdové spektrá sa dajú využiť na opis textúry. Nech $F(u, v)$ je výsledok Fourierovej transformácie skúmanej textúry. Potom

$$F(u, v) = R(u, v) + iI(u, v), \quad (8.3)$$

kde $R(u, v)$ je reálna a $I(u, v)$ imaginárna zložka spektra.

Spektrá využívané pri opise textúr sú uvedené v tabuľke 8.2.

Tabuľka 8.2: Fourierove spektrá.

Amplitúdové spektrum	$ F(u, v) = \sqrt{R^2(u, v) + I^2(u, v)}$
Fázové spektrum	$\phi(u, v) = \arctan \frac{I(u, v)}{R(u, v)}$
Výkonové spektrum	$P(u, v) = F(u, v) ^2 = R^2(u, v) + I^2(u, v)$

8.4 Gaborove filtre

Ľudským vizuálnym systémom sú inšpirované Gaborove filtre. 2D Gaborov filter je Gaussove jadro modulované plošnou sínusovou vlnou

$$h(x, y, \lambda, \theta, \psi, \sigma_x, \sigma_y) = g(x', y', \theta, \sigma_x, \sigma_y) e(x', \theta, \lambda, \psi), \quad (8.4)$$

kde $g(x', y', \theta, \sigma_x, \sigma_y)$ je Gaussove jadro

$$g(x', y', \theta, \sigma_x, \sigma_y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp \left[- \left(\frac{x'^2}{2\sigma_x^2} + \frac{y'^2}{2\sigma_y^2} \right) \right], \quad (8.5)$$

$e(x', \theta, \lambda, \psi)$ je komplexná sínusoida

$$\begin{aligned} e(x', \theta, \lambda, \psi) &= \exp(2\pi i f x' + \psi) = \\ &= \exp \left[i \left(2\pi \frac{x'}{\lambda} + \psi \right) \right] = \\ &= \cos \left(2\pi \frac{x'}{\lambda} + \psi \right) + i \sin \left(2\pi \frac{x'}{\lambda} + \psi \right) \end{aligned} \quad (8.6)$$

a

$$(x', y') = (x \cos \theta + y \sin \theta, -x \sin \theta + y \cos \theta). \quad (8.7)$$

Gaborov filter sa dá rozložiť na reálnu a imaginárnu zložku nasledovne

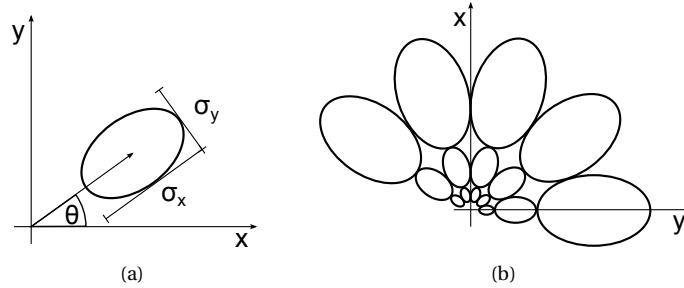
$$h_{Re}(x, y, \Omega) = \frac{1}{2\pi\sigma_x\sigma_y} \exp \left[- \left(\frac{x'^2}{2\sigma_x^2} + \frac{y'^2}{2\sigma_y^2} \right) \right] \cos \left(2\pi \frac{x'}{\lambda} + \psi \right), \quad (8.8)$$

$$h_{Im}(x, y, \Omega) = \frac{1}{2\pi\sigma_x\sigma_y} \exp \left[- \left(\frac{x'^2}{2\sigma_x^2} + \frac{y'^2}{2\sigma_y^2} \right) \right] \sin \left(2\pi \frac{x'}{\lambda} + \psi \right). \quad (8.9)$$

Parametre $\Omega = \{\lambda, \theta, \psi, \sigma_x, \sigma_y\}$ sú opísané v tabuľke 8.3 a graficky znázornené na obrázku 8.3(a).

Tabuľka 8.3: Parametre Gaborových filtrov

λ	vlnová dĺžka
θ	orientácia, smer šírenia vlny
ψ	fázový posun harmonickej zložky
σ_x, σ_y	smerodajná odchýlka, určuje elipticitu Gaussovho jadra



Obrázok 8.3: Hlavné parametre Gaborovho filtra a banka filtrov tvorená zmenou parametrov.

Pre klasifikáciu textúr sa používa množina Gaborových filtrov s rôznymi parametrami, aby sa zachytili rôzne vlastnosti textúr (škála, orientácia). Príklad takejto množiny, ktorá sa nazýva banka filtrov je na obrázku 8.3(b).

Obrázok 8.4 znázorňuje niekoľko realizácií Gaborovho filtra s rôznymi parametrami Ω a obrázok 8.5 aplikáciu týchto filtrov na originálny obraz. Príznačky textúry potom získame tak, že na obraz $I(x, y)$ aplikujeme reálnu a imaginárnu zložku každého z banky filtrov

$$J_{Re}(x, y, \Omega) = I(x, y) * h_{Re}(x, y, \Omega), \quad (8.10)$$

$$J_{Im}(x, y, \Omega) = I(x, y) * h_{Im}(x, y, \Omega). \quad (8.11)$$

Energia E a magnitúda M odozvy filtra je daná nasledovne

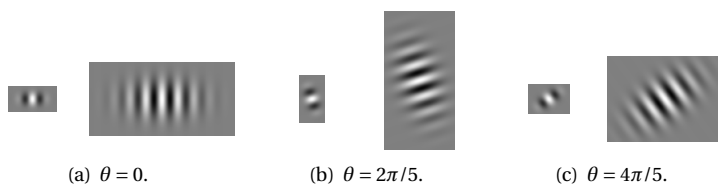
$$E(x, y, \Omega) = J_{Re}^2(x, y, \Omega) + J_{Im}^2(x, y, \Omega), \quad (8.12)$$

$$M(x, y, \Omega) = \sqrt{E}(x, y, \Omega). \quad (8.13)$$

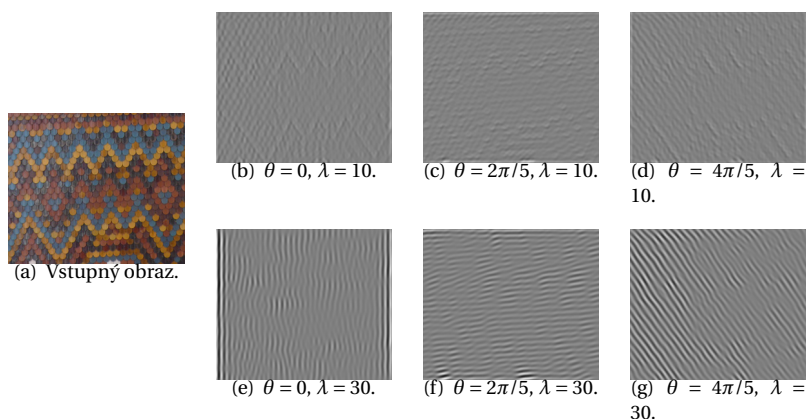
Priemerná energia celého obrazu

$$\bar{E}(x, y, \Omega) = \frac{1}{N} \sum_x \sum_y E(x, y, \Omega), \quad (8.14)$$

kde N je počet pixelov obrazu. Príznakový vektor môžu tvoriť hodnoty priemerov a štandardných odchýlok magnitúd odoziev jednotlivých filtrov.



Obrázok 8.4: Niekoľko realizácií Gaborovho filtra pre rôzne hodnoty θ a vlnové dĺžky $\lambda = \{10, 30\}$.



Obrázok 8.5: Vstupný obrázok filtrovaný filterami z obrázka 8.4.

8.5 MPEG-7

Norma MPEG-7 uvádza tri deskriptory textúry na rôzne použitie. Deskriptor na prehľadávanie textúr charakterizuje hlavné vlastnosti textúr – smerovosť, regulárnosť a zrnitosť. Deskriptor homogénnych textúr počíta lokálne štatistiky frekvenčných vlastností textúry a používa sa na vyhľadávanie podobných textúr. V prípade nehomogénnych textúr môžeme využiť histogram lokálnych hrán. V nasledujúcich častiach opíšeme tieto deskriptory detailnejšie.

8.5.1 Prehľadávanie textúr

Deskriptor na prehľadávanie textúr je 12-bitový deskriptor, ktorý opisuje dvomi bitmi regularitu textúry, šiestimi (2×3) bitmi smery, ktoré najviac



Obrázok 8.6: Textúry s rôznou mierou pravidelnosti zoradené od najmenej po najviac pravidelnú.



Obrázok 8.7: Textúry s rôznou mierou zrnitosti.

vystihujú orientáciu textúry a štyrmi (2×2) bitmi škály, ktoré najviac vystihujú hrubosť (zrnitosť) textúry. Textúry môžu mať viacero charakteristických smerov a škál, preto deskriptor opisuje 2 maximálne hodnoty.

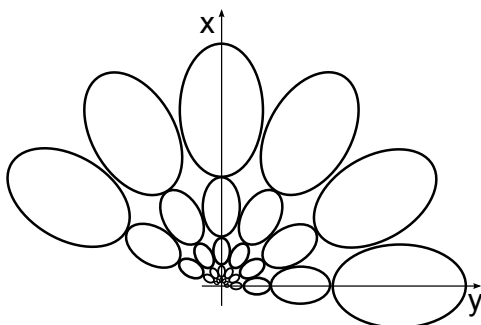
Pravidelnosť textúry je hodnotená číslami $\{0, 1, 2, 3\}$, kde 0 značí ne-regulárnu (náhodnú) textúru. Periodická textúra s pevne definovaným smerom a zrnitosťou je hodnotená číslom 3 (obr. 8.6).

Smer (orientácia) a škála textúry sa dajú zistiť napr. pomocou banky Gaborových filtrov (kap. 8.4). Smer opisujú hodnoty $\{0, \dots, 6\}$, kde 0 znamená, že textúra nemá dominantný smer. Čísla $\{1, \dots, 6\}$ určujú smery od 0° po 150° v 30° krokoch. Ak má textúra len jeden dominantný smer, zapíše sa jeho hodnota do oboch políčok smeru, rovnako ako keď má jednu dominantnú škálu. Škála (zrnitosť) je hodnotená na stupnici $\{0, 1, 2, 3\}$, kde 0 hovorí o jemnej textúre a 3 o veľmi zrnitej textúre. Príklad rôzne zrnitých textúr je na obrázku 8.7.

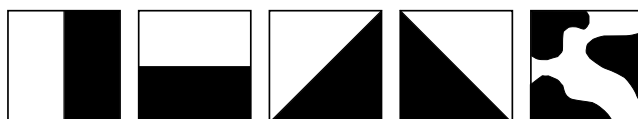
8.5.2 Homogénne textúry

Tento deskriptor používa banku 30 Gaborových filtrov, ktoré pokrývajú 6 orientácií a 5 škál (obr. 8.8). Z odoziev týchto 30 filtrov sa vypočíta energia a jej štandardná odchýlka a ďalej priemer a štandardná odchýlka intenzity obrazu. Deskriptor teda pozostáva zo 62 údajov

$$\{E_1, \text{Std}(E_1), \dots, E_{30}, \text{Std}(E_{30}), \bar{I}, \text{Std}(I)\}. \quad (8.15)$$



Obrázok 8.8: Banka 30 Gaborových filtrov pre deskriptor homogénnych textúr.



(a) Štyri typy smerových hrán, jedna izotropná oblasť.

1	-1	1	1	$\sqrt{2}$	0	0	$\sqrt{2}$	2	-2
1	-1	-1	-1	0	$-\sqrt{2}$	$-\sqrt{2}$	0	-2	2

(b) Filtre zodpovedajúce typom hrán.

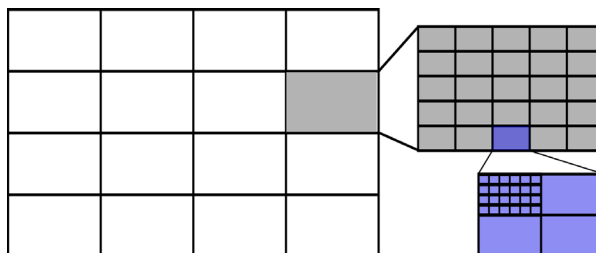
Obrázok 8.9: Rôzne typy detegovaných textúr.

8.5.3 Histogram lokálnych hrán

Tento deskriptor zachytáva distribúciu piatich typov hrán v textúre. Štyri typy sú smerové hrany (horizontálne, vertikálne a 2 diagonálne) a piaty typ zachytáva izotropné oblasti (obr. 8.9(a)).

Postup vytvárania deskriptora je na obrázku 8.10. Pôvodný obraz sa rozdelí na 4×4 bloky a z každého bloku sa vypočíta histogram lokálnych hrán. Veľkosť výsledného deskriptora je teda $16 \text{ blokov} \times 5\text{-hodnotový histogram} = 80$ hodnôt.

Histogram každého bloku sa vypočíta ďalším rozdelením tohoto bloku na daný počet podblokov. Veľkosť týchto podblokov sa mení s rozlíšením obrazu, pretože ich počet je konštantný. Každý podblok je následne aproximovaný 2×2 super-pixelovou oblasťou priemerných intenzít, na



Obrázok 8.10: Pôvodný obraz sa rozdelí na 4×4 bloky a z každého bloku sa vypočíta histogram lokálnych hrán. Histogram každého bloku sa vypočíta ďalším rozdelením tohoto bloku na daný počet podblokov. Každý podblok je následne aproximovaný 2×2 super-pixelovou oblasťou priemerných intenzít.

ktorú sú aplikované hranové filtre uvedené na obrázku 8.9(b). Ak je maximálna odozva týchto filtrov väčšia ako stanovený prah, zvýši sa zodpovedajúca trieda histogramu bloku.



Lokálne príznaky

Dôležitým krokom pri rozpoznávaní objektov je extrakcia príznakov. V tejto kapitole predstavíme nový typ príznakov – lokálne príznaky. Porovnáme lokálne a globálne príznaky na konkrétnych príkladoch, budeme sa venovať všeobecnému opisu tohto druhu príznakov a zdefinujeme základné časti algoritmu extrakcie lokálnych príznakov. Ďalej opíšeme niektoré konkrétne algoritmy ako napríklad Harrisov detektor rohov, SIFT, SURF a ORB a porovnávanie príznakových vektorov. Pozrieme sa na tieto príznaky v kontexte registrácie obrazov a predstavíme si niektoré typické aplikácie využívajúce lokálne príznaky ako napríklad spájanie panorám, stereo-rekonštrukcia, alebo vyhľadávanie obrazov. Na konci kapitoly podrobnejšie rozoberieme konkrétne použitie opísaných metód v prípadovej štúdii a načrtneť použitie lokálnych príznakov pre 3D dáta.

9.1 Lokálne vs. globálne príznaky

Ako sme už spomínali v predchádzajúcich kapitolách, v počítačovom videní extrahujeme informáciu o obraze pomocou príznakov. Existuje veľa rôznych druhov príznakov založených napríklad na farbe, textúre alebo tvare. Pri rozpoznávaní objektov rozlišujeme vo všeobecnosti dve úlohy:

1. Rozpoznávanie kategórií objektov (napríklad chrobáky, bicykle)
2. Rozpoznávanie konkrétnych objektov (napríklad Jankova plyšová panda, Manetova maľba: Raňajky v tráve)



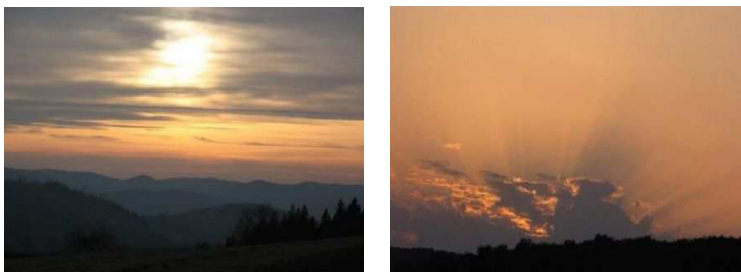
Obrázok 9.1: Príklady dvoch úloh rozpoznávania objektov. Vrchný riadok: Rozpoznávanie kategórie objektov, v tomto prípade chrobákov. Spodný riadok: Rozpoznávanie konkrétneho objektu, v tomto prípade Zamkovského chata v Tatranskom národnom parku.

Pre ilustráciu uvádzame príklady obrazov pre obe úlohy na obrázku 9.1.

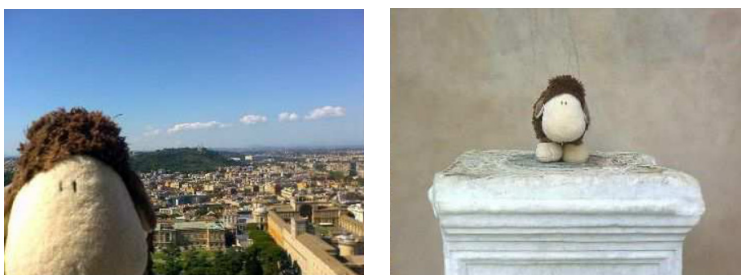
Pri týchto úlohách môžeme využiť globálne príznaky, ktoré analyzujú obraz ako celok a extrahujú jednu hodnotu (alebo vektor hodnôt) pre celú maticu pixelov obrazu a lokálne príznaky, ktoré najskôr hľadajú v obraze zaujímavé body („interesting points“) a extrahujú potom príznaky z okolia týchto bodov. Jeden obraz je teda opísaný takým množstvom hodnôt (vektorov hodnôt), aký je počet zaujímavých oblastí v obraze. Pre lepšiu ilustráciu rozdielu týchto dvoch typov príznakov uvádzame nasledujúci príklad. Máme databázu obrazov (napr. všetky obrázky v databáze Flickr ¹) a chceme v nej vyhľadať obraz. Pri vyhľadávaní obrazu pre dopyt Západ slnka (pozri obrázok 9.2), využijeme globálne príznaky farby, lebo takýto obraz vieme asi najlepšie opísať pomocou dominantných farieb, typických pre západ slnka, a ich výskytu v celom obraze.

Keď sa však pozrieme na vyhľadávanie dopytu ovečka Tilia (pozri obrázok 9.3), tak tušíme, že prístup z prvého príkladu by v tomto prípade

¹<http://www.flickr.com/>



Obrázok 9.2: Príklad vyhl'adaných obrazov pre dopyt „Západ slnka“. Na vyhl'adanie takéhoto typu obrazu použijeme globálne príznaky farby.



Obrázok 9.3: Príklad vyhl'adaných obrazov pre dopyt ovečka Tilia. Na vyhl'adanie takéhoto typu obrazu použijeme lokálne príznaky, ktoré nám v tomto prípade lepšie opíšu hl'adaný objekt.

nefungoval, ak by na obraze boli aj iné objekty, resp., keby bol objekt viditeľný len čiastočne. Keď vyhl'adáme konkrétny objekt, chceme, aby bol identifikovaný aj v prítomnosti iného objektu (objektov), zaradený do zložitejšej scény, aj natočený, škálovaný, čiastočne prekrytý alebo zobrazený pri inom osvetlení. V tejto chvíli prichádzajú na rad lokálne príznaky, ktoré nám pomôžu vyhl'adávať a opisovať konkrétne oblasti patriace objektu v obraze.

9.2 Lokálne príznaky

Počet extrahovaných lokálnych príznakov z obrazu závisí od počtu detegovaných zaujímavých bodov. Ak máme napríklad obraz bielej steny, žiadne lokálne príznaky z neho extrahovať nevieme, lebo na obraze nie sú žiadne lokálne zaujímavé body.

Ako sme už naznačili v predchádzajúcom texte, proces získavania lokálnych príznakov (LP) sa skladá z dvoch častí:

1. Vyhľadávanie zaujímavých bodov (IP),
2. Opis zaujímavých bodov a ich okolia vo forme príznakov (deskriptor).

Ked' hovoríme o LP, je dôležité spomenúť, že väčšina týchto metód sa snaží hľadať a vytvárať také opisné vektory zaujímavých bodov, ktoré zostanú rovnaké, aj keď bude objekt, ktorý hľadáme natočený, škálovaný a podobne. Preto zvyčajne hovoríme o invariantnosti týchto príznakov vzhľadom na nejakú transformáciu.

Príznačky môžu byť invariantné geometricky a rádiometricky. Rádiometrickou invariantnosťou myslíme nezávislosť od osvetlenia scény. Geometrickou invariantnosťou myslíme nezávislosť od rôznych geometrických transformácií. Pri rozpoznávaní objektov môžeme mať na mysli dva druhy transformácií: afinné transformácie a projektívne transformácie. Afinné transformácie zobrazujú priamku na priamku alebo bod a zachovávajú pomery medzi bodmi ležiacimi na priamke. Patria sem všetky podobnostné transformácie (posunutie, otočenie, škálovanie a zrkadlenie) plus skosenie a projekcia a ich skladanie.

Ak v počítačovom videní uvažujeme planárny objekt zosnímaný dvoma kamerami z dostatočnej vzdialenosti (objekt pokrýva iba malú časť obrazu), považujeme transformáciu objektu medzi obrazmi za afinnú. Ak podmienka dostatočnej vzdialenosti neplatí, považujeme transformáciu za projektívnu. V oblasti lokálnych príznakov sa budeme najčastejšie zaoberať metódami invariantnými voči afinným transformáciám.

V nasledujúcej sekcii sa budeme venovať detektorom a deskriptorom lokálnych príznakov, pričom uvedieme aj známe metódy ktoré implementujú oba procesy. Medzi najznámejšie a najpoužívanejšie patria metódy SIFT a SURF. Nové metódy však neustále pribúdajú, preto sa pokúsime naznačiť výskum v tejto oblasti.

9.3 Metódy detekcie a deskripcie lokálnych príznakov

Táto sekcia je rozdelená na dve časti. V prvej si predstavíme niektoré známe detektory lokálnych príznakov. V druhej časti si najskôr predstavíme jeden z prvých (a doteraz používaných) kompletných algoritmov na získanie lokálnych príznakov SIFT a potom metódu SURF, ktorá predstavuje výrazné urýchlenie metódy SIFT bez straty presnosti. Ďalej predsta-

víme metódy na počítanie binárnych deskriptorov ako je BRIEF a napokon metódu ORB. Okrem algoritmov, ktoré budú v tejto časti opísané, existuje veľké množstvo ďalších metód ako sú SUSAN detektor [SB95], Wangov a Bradyho detektor [WB95], alebo deskriptory DAISY [TLF08], FREAK [AOV12] a BRISK [LCS11].

9.3.1 Harrisov detektor rohov

Harrisov detektor rohov bol publikovaný v [HS88]. Tento detektor využíva deriváciu šedotónového obrazu. Pre obrázok $I(x, y)$ sa Harrisov detektor $M(x, y)$ definuje pomocou matice derivácií I_x, I_y (resp. diferencií) intenzity obrazu I :

$$M(\mathbf{x}) = \sum_{x,y} w(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}, \quad (9.1)$$

kde $w(x, y)$ je funkcia okna, ktorá je daná ako vážená suma (Gaussovým jadrom), I_x, I_y sú derivácie v smere osi x a y . Pre detekciu rohov sú dôležité vlastné hodnoty matice M , teda (λ_1, λ_2) , ktoré vyjadrujú varianciu v dvoch hlavných smeroch.

Ak je variancia v oboch smeroch veľká (teda vlastné hodnoty λ_1, λ_2 sú dostatočne veľké), leží bod na rohu, a ak je veľká len jedna z hodnôt, leží na hrane.

Zjednodušené si môžeme predstaviť, že našou snahou je nájsť také (dostatočne malé) oblasti v obraze, ktoré, keby sme vystrihli a priložili k rovnako veľkým oblastiam v ich okolí, rozdiel by bol výrazný.

V [HS88] využívajú namiesto vlastných hodnôt mieru:

$$R(x) = \det(M(\mathbf{x})) - k \text{trace}(M(\mathbf{x}))^2, \quad (9.2)$$

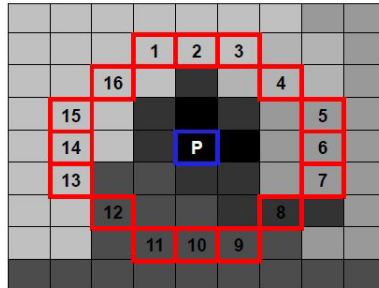
pričom, ak hodnota R pre bod x presiahne zvolený prah, bod je klasifikovaný ako roh, teda detegovaný ako zaujímavý bod. Táto miera bola zjednodušená a zefektívnená v článku [ST94], kde autori použili:

$$R(x) = \min(\lambda_1, \lambda_2), \quad (9.3)$$

Detektor využívajúci túto mieru je známy pod názvom „Good features to track“, alebo Shi-Tomasi detektor.

9.3.2 FAST

Detektor FAST bol publikovaný v [RD06]. Je prvým detekčným algoritmom, ktorý vychádza z AST („Accelerated Segment Test“ – akcelerovaného segmentového testu). Skratka FAST vlastne znamená príznaky z AST



Obrázok 9.4: Detekcia zaujímavých bodov pomocou metódy FAST. P je skúmaný bod, v ktorého okolí vo vzdialenosti $r = 3$ je 11 susediacich bodov (1 – 7, 13 – 16), ktorých intenzita je vyššia ako P . Bod P je detegovaný ako zaujímavý bod.

(„features from AST“). Algoritmus pre každý bod P obrazu preskúma jeho okolie veľkosti Bresenhamovho (rasterizovaného) kruhu s polomerom r . Ak je v okolí n susediacich pixelov, pre ktoré platí, že ich intenzita je vyššia v porovnaní so skúmaným pixelom o viac ako t , alebo menšia o viac ako t , potom je skúmaný pixel považovaný za zaujímavý bod. Algoritmus používa hodnotu polomeru $r = 3$ a najlepšie výsledky dosahuje pri $n = 9$. Na vytvorenie poradia, v ktorom sa budú jednotlivé pixely testovať, sa využívajú rozhodovacie stromy skonštruované z trénovacej množiny dát, v prípade FASTu autori využili „ID3 – Iterative dichotomizer“ (kap. 14.3). Na obrázku 9.4 môžeme vidieť detegciu bodu P s kruhovým okolím polomeru $r = 3$ a počtom potrebných bodov $n = 9$. Body okolia 1 – 6 a 14 – 16 majú väčšiu intenzitu ako P , preto môžeme P považovať za zaujímavý bod.

9.3.3 SIFT

SIFT („Scaleable Invariant Feature Transform“) je technika vytvorená Lowem [Low02], ktorá predstavuje škálovo a rotačne invariantný detektor a deskriptor zaujímavých bodov. V prvom kroku algoritmu SIFT sú detektorom nájdené zaujímavé body v obraze a v druhom kroku je vytvorený deskriptor pre každý z týchto bodov. Na konkrétne implementácie metód detektora a deskriptora sa zameriame v nasledujúcich častiach

Detektor: SIFT využíva na vyhľadávanie zaujímavých bodov tzv. škálový priestor (alebo škálovú pyramídu), ktorý vytvára s využitím rozdielu

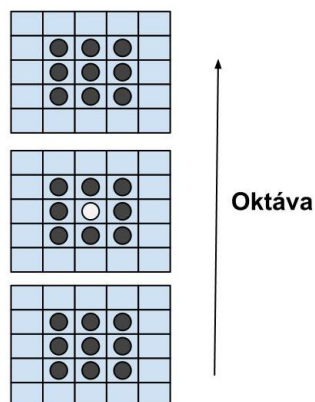


Obrázok 9.5: Škálovacia pyramída so 4 oktávami a 5 obrazmi v každej oktáve. Obrazy sú v rámci každej oktávy opätovne filtrované Gaussiánom. Nová oktáva obsahuje obrazy s polovičnou šírkou a výškou oproti predchádzajúcej oktáve.

Gaussiánov (Difference of Gaussians – DoG). Obraz je podvzorkovaný do k veľkostí – „oktáv“. V každej oktáve je vytvorených n ďalších obrazov opätovným filtrovaním pôvodného obrazu Gaussiánom. Príklad 1D Gaussiánu:

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2}, \quad (9.4)$$

kde hodnota σ označuje štandardnú odchýlku a jej hodnota pre všetky vyhladzovania je $\sqrt{2}$. Teda pri prvom vyhladzovaní v prvej oktáve dostaneme z pôvodného obrazu A, obraz B pomocou Gaussiánu so štandardnou odchýlkou $\sigma = \sqrt{2}$. Nasledujúci obraz C dostaneme z B pomocou vyhladenia Gaussiánom so $\sigma = \sqrt{2}$, obraz C by sme z A dostali pomocou vyhladenia Gaussiánom, opäť so

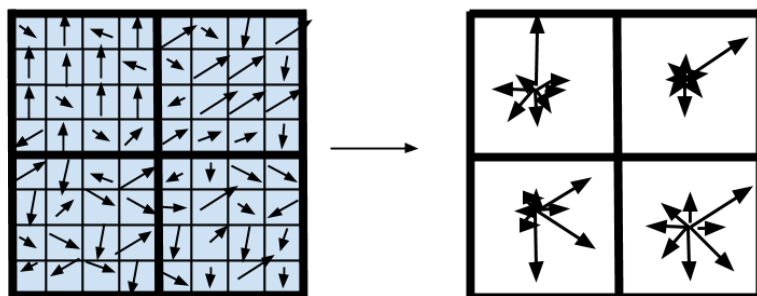


Obrázok 9.6: Lokálne extrémny boli nájdené porovnávaním pixelu (biely bod) s okolím $3 \times 3 \times 3$ (čierne body), teda 3×3 okolím v aktuálnom, predchádzajúcom a nasledujúcom obraze.

$\sigma = 2$. Škálová pyramída so 4 oktávami a 5 obrazmi v každej oktáve je zobrazená na obraze 9.5.

V rámci každej oktávy sa potom všetky susedné obrazy odčítajú, čím vznikne $n - 1$ obrazov v každej oktáve. Rozdiel Gaussiánov, ako bol opísaný v tejto časti, sa využíva preto, lebo predstavuje aproximáciu škálovo normalizovaného Laplaciánu z Gaussiánu (pozri kapitolu o hranových detektoroch). Výhodou použitia Gaussiánu je jeho separovateľnosť, teda možnosť použitia dvoch 1D filtrov namiesto jedného 2D filtra, pre zníženie komplexnosti výpočtu. Vo vytvorenom škálovom priestore sa potom hľadajú lokálne extrémny v okolí $3 \times 3 \times 3$. Okolie $3 \times 3 \times 3$ znamená, že sa v rámci obrazu, v ktorom sa skúmaný bod nachádza, hľadá v okolí 3×3 , ale taktiež sa hľadá aj v okolí 3×3 na nasledovnom a predchádzajúcom obraze v rámci škálového priestoru (pozri obrázok 9.6). Tieto lokálne extrémny (teda minimá alebo maximá) sú identifikované ako zaujímavé body. Oktáva, v ktorej bol zaujímavý bod nájdený, bude predstavovať „škálu“ daného zaujímavého bodu a určí veľkosť okolia, ktoré sa použije pri extrakcii deskriptora.

Deskriptor: Pre každý zaujímavý bod je vypočítaná hodnota deskriptora z histogramov orientácií pre 16 štvorcových oblastí v jeho okolí (veľ-



Obrázok 9.7: Príklad vytvorenia deskriptora z okolia kľúčového bodu pri metóde SIFT. Na obraze sú použité štyri (2×2) regióny namiesto šesťnást' (4×4) . Každú oblasť vzorkujeme 16 bodmi, v ktorých sa vypočíta veľkosť a orientácia gradientu (malé šípky v ľavom obraze). Z týchto sa potom vytvoria výsledné histogramy orientácií (šípky v pravom obraze).

kosť oblastí je závislá od škály daného zaujímavého bodu). V prvom kroku je dôležité zistiť dominantnú orientáciu každého kľúčového bodu a potom vytvoriť deskriptor bodu relatívne k tejto orientácii. Každú oblasť potom navzorkujeme 16 bodmi a v každom z nich (A_{ij}) vypočítame veľkosť (M_{ij}) a orientáciu R_{ij} gradientu nasledovne:

$$M_{ij} = \sqrt{(A_{ij} - A_{i+1,j})^2 + (A_{ij} - A_{i,j+1})^2}, \quad (9.5)$$

$$R_{ij} = \text{atan2}(A_{ij} - A_{i+1,j}, A_{i,j+1} - A_{ij}). \quad (9.6)$$

kde atan2 je hodnota arcus tangens s oborom hodnôt $(-\pi, \pi)$

$$\text{atan2}(y, x) = \begin{cases} \arctan\left(\frac{y}{x}\right) & x > 0, \\ \arctan\left(\frac{y}{x}\right) + \pi & y \geq 0, x < 0, \\ \arctan\left(\frac{y}{x}\right) - \pi & y < 0, x < 0, \\ +\frac{\pi}{2} & y > 0, x = 0, \\ -\frac{\pi}{2} & y < 0, x = 0, \\ \text{nedef.} & y = 0, x = 0. \end{cases} \quad (9.7)$$

Týmto hodnotám sú priradené váhy pomocou Gaussovho kruhového okna (pozri obrázok 9.7). Z váhovaných hodnôt v 16 bodoch sú potom vytvorené histogramy orientácií, pričom každý histogram

pokrýva jednu zo šesnástich (4×4) oblastí z okolia zaujímavého bodu. Dĺžka šípky v histograme predstavuje sumu veľkostí gradientov v blízkosti daného smeru v oblasti. V každom histograme je 8 smerov, čo pri počte 4×4 oblastí znamená 128 hodnôt deskriptora pre jeden zaujímavý bod.

9.3.4 SURF

SURF („Speeded Up Robust Features“) je technika, ktorú autori Bay et al. [Bay+08] považujú za vylepšenie SIFTu a iných podobných techník. SURF obsahuje podobne ako SIFT detektor na nájdenie a deskriptor na opísanie zaujímavých bodov.

Detektor: SURF, podobne ako SIFT deteguje zaujímavé body v škálovej pyramíde, namiesto Gausiánov však využíva aproximáciu Gausiánov druhej parciálnej derivácie v smere y a xy pomocou „box filtrov“. Na vytvorenie škálového priestoru sa pri tejto metóde nevyužíva zmenšovanie obrazu a následná konvolúcia, ale naopak, konvolúcia pôvodného obrazu s postupne sa zväčšujúcimi filtermi. Na nájdenie zaujímavých bodov je na obraz aplikované potlačenie nemaximálnych hodnôt v okolí $3 \times 3 \times 3$ (okolie $3 \times 3 \times 3$ je okolie spomínané pri technike SIFT).

Deskriptor: Pre vytvorenie deskriptora zaujímavých bodov je potrebné v prvom kroku zistiť jeho orientáciu a v druhom kroku vytvoriť štvorcovú oblasť zarovnanú podľa tejto orientácie (z nej sa bude deskriptor extrahovať). Pre určenie orientácie zaujímavého bodu sa najskôr vypočítajú odozvy Haarovho waveletu v x -ovom a y -ovom smere v kruhovom okolí bodu s priemerom $6s$, pričom s je veľkosť aktuálnej škály. Vypočítaným odozvám sú potom priradené váhy pomocou Gaussovho kruhového okna a sú reprezentované ako vektor v priestore. Dominantná orientácia zaujímavého bodu je najdlhší z vektorov vypočítaných ako suma vertikálnych a horizontálnych odoziev, získaných pomocou otočného okna ako suma všetkých odoziev v danej orientácii. V ďalšom kroku je vytvorená v okolí zaujímavého bodu štvorcová oblasť zarovnaná podľa orientácie. Oblasť je rozdelená na 4×4 subregióny, pričom v každom sú vypočítané 4 príznaky na 5×5 pravidelne rozmiestnených bodoch. Príznakmi sú: suma odoziev Haarovho waveletu v horizontálnom a vertikálnom smere a suma absolútnych hodnôt odoziev Haarovho waveletu v horizontálnom a vertikálnom smere. Des-

kriptor zaujímavého bodu potom obsahuje 4 hodnoty pre každú zo 16 oblastí (spolu 64 hodnôt).

9.3.5 BRIEF

Deskriptor: BRIEF alebo Binárne robustné invariantné elementárne príznaky (Binary Robust Invariant Elementary Features) je deskripčná metóda, ktorá bola navrhnutá a popísaná v [Cal+10]. BRIEF je binárny deskriptor, ktorý opisuje záplatu priloženú k obrazu, kde každý bit zodpovedá jednému výsledku binárneho testu na rozdiel intenzity. Môžeme si to predstaviť tak, že v okolí nejakého bodu obrázku rozsypeme n zápaliiek. Potom sa pozrieme pod koncové a počiatočné body zápaliiek a pre každú zápalku porovnáme ich hodnotu. Výsledný deskriptor dĺžky n bude obsahovať 1, ak pre danú zápalku bola hodnota intenzity pod počiatočným bodom vyššia a 0 naopak. Pre záplatu p (pre príklad záplaty pozri obrázok 9.8), je vektor príznakov dĺžky n bitov definovaný

$$f_n(p) = \sum_{1 \leq i \leq n} 2^{i-1} \tau(p; \mathbf{x}_i, \mathbf{y}_i), \quad (9.8)$$

kde

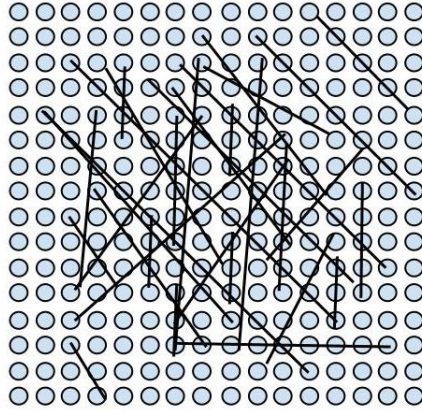
$$\tau(p; \mathbf{x}, \mathbf{y}) = \begin{cases} 0 & \text{ak } p(\mathbf{x}) \leq p(\mathbf{y}), \\ 1 & \text{ak } p(\mathbf{x}) > p(\mathbf{y}), \end{cases} \quad (9.9)$$

kde $p(\mathbf{x})$ je intenzita v bode $\mathbf{x}$.

9.3.6 ORB

Detektor: Detektor, ktorý využíva metóda ORB, sa nazýva oFAST a je modifikáciou metódy FAST. Rozdiel spočíva v tom, že detektor oFAST prahuje rozdiel intenzít medzi centrálnym pixelom a pixelami v jeho kruhovom okolí s priemerom 9 pixelov. Na nájdenie n zaujímavých bodov algoritmus nastaví najskôr nižší prah tak, aby vyhovelo viac ako n bodov. Následne tieto body utriedi na základe Harrisovej miery 9.3, predstavenej v [HS88] a vyberie n najlepších bodov.

Deskriptor: Ako deskriptor využíva ORB modifikáciu BRIEFu, nazvanú rBRIEF, teda BRIEF s rotáciou. Rotácia detegovaného zaujímavého bodu je extrahovaná s použitím ťažiska intenzity okolia. Metóda vychádza z toho, že vektor zo zaujímavého bodu do ťažiska (ťažiska intenzity okolia) indukuje rotáciu. Keď zvolíme okolie zaují-



Obrázok 9.8: Príklad záplaty s 32 čiarami, vytvárajúcej deskriptor dĺžky 32 bitov.

navého bodu ako záplatu s veľkosťou r , môžeme vypočítať centroid C záplaty ako

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right), \quad (9.10)$$

kde m_{pq} sú momenty záplaty, pre ktoré platí, že

$$m_{pq} = \sum_{x,y} x^p y^q I(x, y). \quad (9.11)$$

Orientácia takejto záplaty je potom

$$\theta = \text{atan2}(m_{01}, m_{10}). \quad (9.12)$$

V časti venovanej BRIEFu sme spomínali binárne testy na rozdiel intenzít a spôsob, ako sa počítajú. Pre každú množinu príznakov obsahujúcu N binárnych testov na pozícii $\mathbf{x}_i, \mathbf{y}_i$ môžeme definovať maticu $2 \times N$

$$S = \begin{pmatrix} \mathbf{x}_1 & \dots & \mathbf{x}_n \\ \mathbf{y}_1 & \dots & \mathbf{y}_n \end{pmatrix}. \quad (9.13)$$

Pre rBRIEF autori použili orientáciu θ a jej zodpovedajúcu maticu rotácie R_θ na zostrojenie matice

$$S_\theta = R_\theta S. \quad (9.14)$$

Vektor príznakov pre metódu rBRIEF sa potom vypočíta takto

$$g_n(p, \theta) := \sum_{(\mathbf{x}_i, \mathbf{y}_i) \in S_\theta} 2^{i-1} \tau(p; \mathbf{x}_i, \mathbf{y}_i). \quad (9.15)$$

Pridanie rotačnej invariance však spôsobuje problém so stratou variancie v rBRIEFe. Na zníženie korelácie medzi jednotlivými binárnymi testami autori navrhli metódu učenia sa na získanie „dobrej“ podmnožiny binárnych testov. Algoritmom je greedy („pažravé“) prehľadávanie množiny nekorelovaných testov s priemerom blízko 0,5 vo všetkých možných binárnych testoch na záplate s veľkosťou 31×31 pixelov.

9.3.7 Porovnávanie deskriptorov

Dôležitým krokom pri využívaní lokálnych príznakov je aj vzájomné porovnávanie extrahovaných deskriptorov zaujímavých bodov medzi obrazmi a objektami. Klasickým prístupom je v tomto smere metóda brute force, teda porovnávanie každý s každým. Pri tejto metóde deskriptyory porovnáваме postupne so všetkými deskriptormi z nejakej databázy, alebo extrahovanými z druhého obrazu a počítame vzdialenosti medzi nimi, pričom deskriptor s najmenšou vzdialenosťou je najlepším kandidátom na zhodu. Keďže takýto spôsob porovnávania je vďaka svojej vysokej zložitosti časovo náročný, v praxi sa používajú rôzne vylepšenia s využitím usporiadania deskriptorov do stromov, alebo s použitím binárneho hashovania.

9.4 Registrácia objektov

Doteraz sme v tejto kapitole hovorili najmä o využití lokálnych príznakov pri detekcii a rozpoznávaní objektov. V počítačovom videní je často okrem rozpoznania objektov dôležitá aj registrácia. Len pre príklad uvedieme, že registrácia je dôležitým krokom napríklad v rozšírenej realite (augmented reality), kde slúži na určenie správneho umiestnenia virtuálnych objektov v reálnom prostredí. Podľa Zitovej [ZF03] poznáme štyri typy registrácie

1. Multitemporálna registrácia,
2. Multimodálna registrácia,
3. Registrácia scény a modelu,
4. Multipohľadová registrácia.

Lokálne príznaky sa dajú priamočiaro použiť pri poslednom type registrácie, pričom pri ostatných typoch je ich použitie závislé od konkrétnych podmienok. Ak by sme napríklad chceli zarovnávať jeden rez z počítačovej tomografie (CT) s rezom z rovnakého miesta z magnetickej rezonancie (MRI) pri spracovaní medicínskych dát (pri multimodálnej analýze), nebolo by možné použiť metódy ako SIFT bez väčších modifikácií. Dôvodom je fakt, že dáta zosnímané týmito dvoma metódami zvyčajne majú úplne iné štruktúry (CT – kosti, MRI – mäkké tkanivá), a preto lokálne zaujímavé body v týchto obrazoch nekorešpondujú.

Proces registrácie obsahuje okrem detekcie príznakov aj ďalšie kroky. Celkovo sú fázy registrácie štyri

1. Detekcia príznakov,
2. Párovanie príznakov,
3. Zistenie modelu transformácie,
4. Prezorkovanie a transformácia modelu.

Prvé dve fázy sme predstavili v predchádzajúcich častiach tejto kapitoly. Pre viac informácií o ďalších dvoch fázach pozri [ZF03].

9.5 Využitie lokálnych príznakov

Lokálne príznaky sa, okrem spomínaného rozpoznávania objektov, využívajú aj pre mnohé iné aplikácie z oblastí počítačového videnia, počítačovej grafiky alebo napríklad aj robotiky.

9.5.1 Spájanie panorám

Jednou z prvých aplikácií bolo „zošívanie“ panorám, kde lokálne príznaky slúžia na nájdenie korešpondenčných bodov medzi jednotlivými snímkami. Nájdenie dvojíc korešpondujúcich bodov je pri zošívaní panorám dôležité kvôli vzájomnej registrácii dvoch snímok. Niektoré aplikácie na spájanie panorám stále využívajú manuálne alebo poloautomatické zadávanie korešpondujúcich bodov kvôli väčšej robustnosti. Problém hľadania korešpondujúcich bodov na dvoch fotografiách obsahuje z transformácií: transláciu, rotáciu a v určitých prípadoch aj škálovanie. Voľnými nástrojmi na vytváranie panorám s využitím lokálnych príznakov, sú napríklad Autopano² alebo Autostich³.

²<http://autopano.kolor.com/>

³<http://cs.bath.ac.uk/brown/autostitch/autostitch.html>

9.5.2 Stereo rekonštrukcia

Ďalším populárnym typom aplikácie lokálnych príznakov je 3D rekonštrukcia. Podobne ako v predchádzajúcom prípade lokálne príznaky slúžia na nájdenie korešpondencií medzi obrazmi, z ktorých chceme rekonštruovať scénu (objekt). Tieto obrazy sú typicky nasnímané pomocou dvoch kamier (stereorekonštrukcia), ktoré sú umiestnené v známej vzájomnej pozícii určenej v procese kalibrácie kamery.

9.5.3 Vlastná kalibrácia kamery

V počítačovom videní zvyčajne aproximujeme reálnu kameru modelom štrbinovej kamery. Kalibrácia kamery je proces zisťovania vnútorných a vonkajších parametrov kamery, ktoré reprezentujú geometriu štrbinovej kamery. Vonkajších parametrov kamery je šesť a hovoria o polohe a orientácii kamery v 3D priestore. Sú to 3 parametre rotácie a 3 parametre translácie. Vnútorných parametrov kamery je päť a sú nezávislé od polohy a orientácie kamery v priestore. Metódy kalibrácie rozdelujeme na

1. Kalibráciu zo známej scény (použitie kalibračného objektu so známou geometriou, napríklad šachovnice).
2. Kalibráciu z neznámej scény (použitie viacerých snímok).

Pri kalibrácii z neznámej scény, teda autokalibrácii, sa používa niekoľko rôznych snímok rovnakej scény namiesto známeho kalibračného objektu. Z týchto snímok sa extrahujú dvojice korešpondujúcich bodov a z nich sa vypočíta fundamentálna matica. Fundamentálna matica je matica 3×3 , ktorá dáva do súvisu korešpondujúce body v stereo zobrazení. Z fundamentálnej matice sa potom určia vnútorné parametre kamery a relatívna pozícia kamery (až na projektívnu transformáciu v 3D priestore). Tak ako v predchádzajúcich aplikáciách, lokálne príznaky nám poslúžia na vyhľadanie korešpondencií v snímkach. Použitie metódy SIFT na vlastnú kalibráciu bolo opísané v [LH06]. Pre podrobnejšie informácie o kalibrácii kamery odporúčame [SHB98].

9.5.4 Sledovanie objektov

Podobne sa lokálne príznaky využívajú aj pri sledovaní (tracking) objektov vo videu. Tu sa hľadajú korešpondencie medzi lokálne zaujímavými bodmi detegovanými na sledovanom objekte a nasledujúcimi snímkami

sekvencie. Tieto korešpondencie potom umožňujú sledovať pohyb (bodov) objektu medzi snímkami. Špecifické pre tieto tri typy aplikácií je, že snímky, na ktorých hľadáme korešpondenciu, sú zväčša nasnímané v krátkom časovom odstupe (alebo v rovnakom čase pri stereorekonštrukcii), pod rovnakým osvetlením, v rovnakom prostredí a v podobnej škále.

9.5.5 Vyhľadávanie obrazov

Lokálne príznaky našli využitie aj pri CBIR („Content based image retrieval“), teda vyhľadávaní obrazov na základe ich obsahu. Vyhľadávanie obrazov je neoddeliteľnou súčasťou vyhľadávačov od roku 2001, keď Google zaviedol Image search. Pôvodne sa obrazy vyhľadávali iba na základe ich opisu, resp. na základe textu, ktorý sa nachádzal na stránke v blízkosti obrazu. S príchodom CBIR sa začali obrazy vyhľadávať aj na základe vizuálneho obsahu, teda na základe výskytu a rozloženia farieb alebo textúry. V súčasnosti CBIR systémy obsahujú oveľa sofistikovanejšie možnosti na vyhľadávanie, ako je napríklad vyhľadávanie na základe používateľom zadanej kontúry, alebo vyhľadávanie obrazov obsahujúcich používateľom zadané objekty. V tejto fáze prichádzajú na rad lokálne príznaky, ktoré sa používajú na detekciu hľadaných objektov.

Predstavili sme niektoré klasické aplikácie pre využitie lokálnych príznakov, existuje však mnoho ďalších možností na využitie tohto typu príznakov, ktoré tu neboli predstavené a nechávame ich na čitateľovej fantázii. V nasledujúcej časti pre úplnosť rozoberieme krok po kroku konkrétny prípad použitia.

9.6 Prípád použitia: Klasifikácia malieb

Opíšeme prípad použitia lokálnych príznakov pre rozpoznávanie malieb na fotografiách (budeme vychádzať z [Hal10]). Predstavme si, že sme sa práve vrátili z výstavy v galérii a prezeráme si fotografie, ktoré sme tam vytvorili. Pri prezeraní zistíme, že fotografia mal'by, ktorá sa nám najviac páčila, a ktorú sme chceli použiť ako tapetu na počítači sa nám veľmi nevydarila. Rozhodneme sa teda vyhľadať kvalitnejšiu fotografiu vo virtuálnej galérii na internete. Vo väčšine týchto galérií sa vyhľadávajú mal'by len na základe názvu, autora alebo roku vyhotovenia. Čo však robiť, ak nevieme ani autora ani názov mal'by? Predstavme si na vyhľadanie názvu obrazu nasledujúcu funkciu. Táto funkcia má na vstupe fotografiu a na

výstupe názov maľby z databázy, ktorá je čo najpodobnejšia zadanej fotografii. Databáza obrazov, nazvime ju databáza originálov, s ktorými je náš obraz porovnávaný, obsahuje pre každý obraz jeho názov a súbor hodnôt extrahovaných deskriptorov detegovaných lokálne zaujímavých bodov (pre príklad dopytových obrazov a obrazov z databázy fotografií pozri obr. 9.9).

Ak by sme ako metódu na extrakciu lokálnych príznakov použili metódu SIFT, potom by pre každý obraz v databáze existoval súbor veľkosti $n \times 128$ celých čísel, pričom n je počet zaujímavých bodov v obraze.

Vo fáze vyhľadávania by sme najskôr vypočítali lokálne príznaky pre vstupný (dopytový) obraz a tieto postupne porovnali s lokálnymi príznakmi (uloženými vo forme súborov) všetkých obrazov databázy. Výsledkom každého porovnania súborov je počet zhôd medzi deskriptormi zaujímavých bodov dopytového obrazu a obrazu z databázy. Najlepší kandidát na správnu klasifikáciu (teda názov obrazu) je potom obraz s najväčším počtom zhôd so vstupným obrazom.

Vo všeobecnosti môžeme za základnú metriku pri porovnávaní deskriptorov pri metóde SIFT, resp. pri metódach používajúcich celočíselné deskripty, považovať Euklidovskú vzdialenosť. Využiť sa však teoreticky dajú aj iné metriky, ako napríklad Manhattanská vzdialenosť. Pri metódach, ktoré využívajú binárne deskripty (BRIEF, ORB), teda opisujú zaujímavé body vektormi obsahujúcimi len 0 a 1, sa za štandardnú metriku považuje Hammingova vzdialenosť. Pre porovnanie Euklidovská vzdialenosť je vyjadrená nasledovne. Majme dva body v Euklidovskom priestore $\mathbf{p} = (p_1, p_2, \dots, p_n)$ a $\mathbf{q} = (q_1, q_2, \dots, q_n)$. Potom vzdialenosť medzi $\mathbf{p}$ a $\mathbf{q}$ je

$$d_e(\mathbf{p}, \mathbf{q}) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2}. \quad (9.16)$$

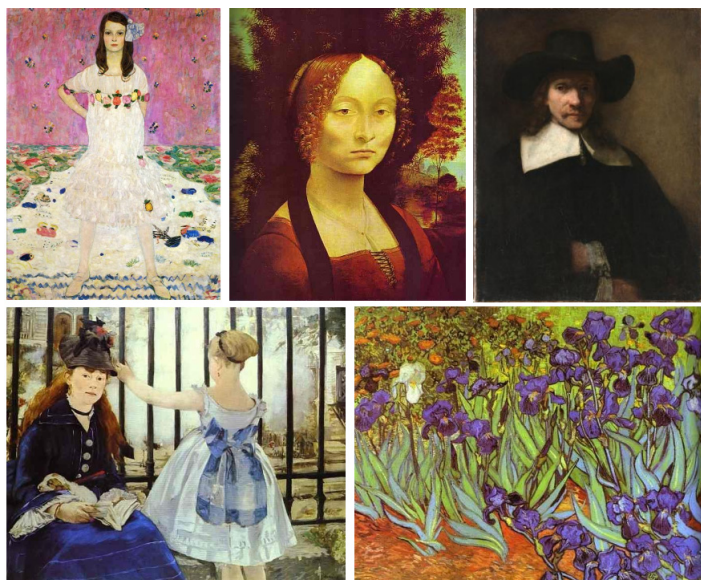
Keď však $\mathbf{p} = (p_1, p_2, \dots, p_n)$ a $\mathbf{q} = (q_1, q_2, \dots, q_n)$ sú dva binárne reťazce, môžeme ich podobnosť vyjadriť ako Hammingovu vzdialenosť nasledovne

$$d_h(\mathbf{p}, \mathbf{q}) = \sum_{i=1}^n \delta(p_i, q_i) / n, \quad (9.17)$$

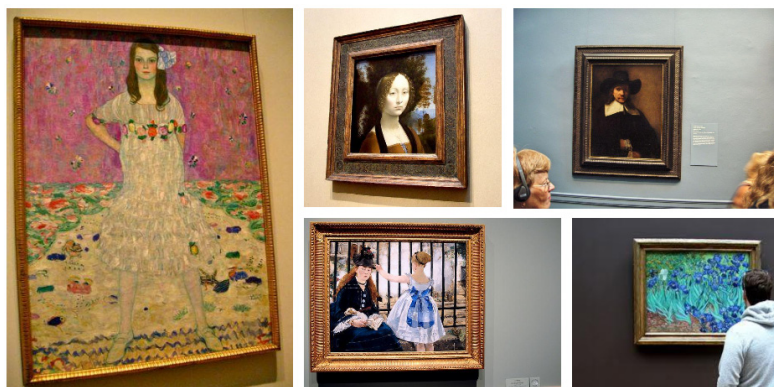
kde

$$\delta(\mathbf{p}, \mathbf{q}) = \begin{cases} 0 & \text{ak } (p_i > 0 \wedge q_i > 0) \vee (p_i = 0 \wedge q_i = 0), \\ 1 & \text{ak } = \text{inak.} \end{cases} \quad (9.18)$$

Výhodou Hammingovej vzdialenosti je jej veľmi rýchly a efektívny výpočet. Pri porovnávaní veľkej databázy príznakov predstavuje výrazné



(a)



(b)

Obrázok 9.9: Príklad obrazov z databázy originálov (a) a dopytových obrazov (b). Databáza originálov obsahuje obrazy zosnímané vo vysokom rozlíšení pod konštantným osvetlením a s minimálnym šumom. Dopytové obrazy boli zosnímané turistami pri návšteve galérie a slúžia ako dopyt pre vyhľadanie názvu, resp. korešpondujúceho originálu z databázy originálov.

zrýchlenie oproti použitiu Euklidovskej vzdialenosti. Táto skutočnosť je jedným z hlavných dôvodov úspechu nových metód využívajúcich binárne deskriptory.

Na rozhodnutie, či je určená vzdialenosť deskriptorov dostatočne malá, aby sa dali deskriptori považovať za zhodné, sa používa buď jednoduché prahovanie, alebo napríklad metóda druhého najbližšieho suseda („second nearest neighbour“). Táto metóda najskôr utriedi jednotlivé porovnávané deskriptory podľa vzdialenosti a za zhodu považuje najbližší deskriptor vtedy, ak jeho vzdialenosť je 0,6 krát menšia ako vzdialenosť druhého najbližšieho deskriptora.

Ak ako názov hľadaného obrazu zvolíme vždy názov kandidáta z najväčším počtom zhôd bezohľadu na počet zistených zhôd sa môže stať, že aj obraz, ktorý v databáze nemáme, označíme názvom najlepšieho kandidáta. Pri finálnom vyhodnotení sa v týchto prípadoch zvýši počet obrazov zaradených do nesprávnej triedy. Pre zníženie počtu zlých klasifikácií treba používať prah na minimálny počet zhôd pre správnu klasifikáciu. Ak použijeme prah príliš nízky, môže sa stať, že množstvo obrazov nepatriacich do žiadnej kategórie bude nesprávne klasifikované ako do kategórie patriace, ako v spomínanom prípade („false positive“). Keď však nastavíme prah príliš vysoko, bude veľa obrazov z databázy klasifikovaných, ako keby do nej nepatril („false negative“). Aby sme sa vyhli tejto situácii, správnu hodnotu prahu môžeme natrénovať na malej časti testovacej databázy (kap. 16).

9.7 3D lokálne príznaky

Doteraz sme sa v kapitole venovali len lokálnym príznakom v obrazoch, teda 2D príznakom. S príchodom cenovo dostupných zariadení, ako napríklad Kinect od Microsoftu, poskytujúcich RGBD dáta (3 farebné kanály R, G, B a hĺbková informácia), sa objavili aj metódy extrahujúce lokálne príznaky v RGBD alebo 3D dátach. Pri RGBD dátach sa ako možnosť priamočiarej extrakcie lokálnych príznakov, namiesto informácie o intenzite, ponúka využitie hĺbkovej informácie („depth image“). Na prvý pohľad je zrejmé, že tak, ako sme predpokladali zmenu intenzity na hranách v 2D obraze, môžeme predpokladať, že na hranách 3D objektov nachádzajúcich sa v scéne bude nastane zmena hĺbkovej informácie.

K informácii z hĺbkovej mapy môžeme pristupovať aj inak, prostredníctvom jej projekcie do mračna bodov. Lokálne príznaky sa potom namiesto z 2D dát extrahujú z 3D mračna bodov. V takejto reprezentácii môžeme jednotlivé mračná bodov z jednej scény navzájom registrovať

a spájať. Registrácia mračien bodov sa v štandardne robí v dvoch krokoch. V prvom sa snažíme mračná približne zarovnať použitím jednej transformácie a v druhom kroku použijeme iteratívnu metódu (napríklad „ICP – Iterative Closest point“ [BM92]) na finálne zarovnanie. V prvom kroku sa na približné zarovnanie dajú použiť modifikované metódy na extrakciu 2D lokálnych príznakov. Po zovšeobecnení môžeme povedať, že spájanie mračien bodov je 3D analógia zošívania 2D panorám.

Algoritmy ako SIFT či SURF, pôvodne vyvinuté na 2D dáta, sa uplatnili aj v mračnách bodov alebo iných reprezentáciách 3D objektov, ako sú napr. volumetrické dáta. Napr. v [Kno+10] autori navrhli a implementovali modifikáciu algoritmu SURF – 3D (SURF pre voxelizované (voxel – 3D pixel) dáta), ktorý namiesto kruhového okolia bodu, pracuje so sférickým okolím. Pre prácu s mračnami bodov (spájanie mračien bodov) odporúčame dynamicky sa rozvíjajúcu Open Source knižnicu PCL (Point Cloud Library)⁴.

⁴<http://pcl.org>

Predspracovanie príznakov

V predchádzajúcich kapitolách sme videli, že príznakov opisujúcich obraz je mnoho. Čím viac príznakov extrahujeme z obrazu, tým viac informácie o obraze máme, čím zvyšujeme presnosť následnej klasifikácie. Zároveň však stúpa zložitosť extrakcie, zvyšuje sa pravdepodobnosť, že viaceré príznaky zachytávajú rovnakú informáciu a rastie zložitosť tréningovania klasifikátora. Tomuto javu sa hovorí „preklatie“ dimenzionality.

Riešením je zníženie počtu príznakov. K zníženiu môžeme pristúpiť dvomi základnými spôsobmi. Prvým z nich je *výber príznakov* (kap. 11), kde z pôvodných príznakov vyberieme podmnožinu, ktorá rovnako dobre klasifikuje dáta ako pôvodná množina príznakov. Druhým spôsobom je *redukcia príznakov* (kap. 12), pri ktorej transformujeme pôvodnú množinu do menejrozmernej množiny.

Predtým, než pristúpime k zníženiu dimenzie, musíme rozhodnúť, či príznaky treba predspracovať, alebo nie. Cieľom predspracovania príznakov je úprava dát do formátu vhodného pre používaný klasifikátor, prípadne transformácia dát, ktorá zlepší výsledok klasifikácie.

Podľa typu dát delíme príznaky na nemetrické a metrické.

Nemetrické sa ďalej delia na kvalitatívne a poradové.

- Kvalitatívne (nominálne) príznaky umožňujú rozlíšiť klasifikačné triedy, ale neumožňujú ich porovnať. Takéto dáta nájdeme napríklad pri klasifikácii geometrických útvarov podľa tvaru (kruh, štvorec, ...) alebo fotografií ľudí podľa pohlavia (žena, muž).

- Poradové (ordinálne) príznaky umožňujú usporiadanie príznakov podľa ich hodnoty, ale nehovoria o vzťahu týchto hodnôt. Ak napríklad zaznamenáme pre objekty len ich poradie podľa obsahu a nie konkrétnu hodnotu obsahu objektov, nevieme povedať, o koľko alebo koľko krát je objekt na prvom mieste väčší ako objekt na druhom mieste.

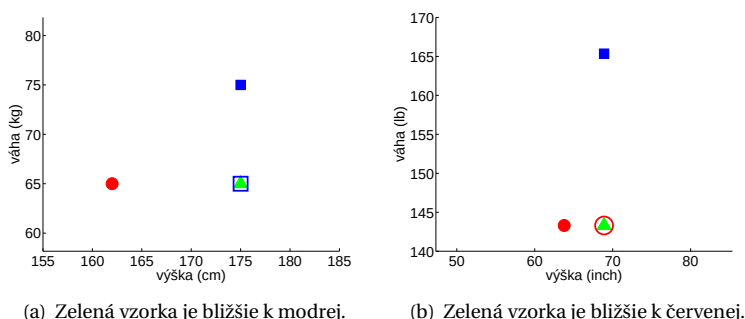
Metrické príznaky môžu byť diskkrétne alebo spojité a delia sa na intervalové a pomerové.

- Intervalové (kardinálne) príznaky sú výsledkom merania v určených jednotkách. Dokážeme z nich určiť, o koľko je jedna hodnota vyššia ako druhá. Všetky inštancie jedného príznaku musia byť merané v rovnakých jednotkách. Nemôžeme merať obvod jedného objektu v pixeloch a druhého v centimetroch. Ak intervalovú premennú reprezentujeme histogramom, ide už o poradovú premennú.
- Pomerové príznaky sú vyjadrené v škále s absolútnym nulovým bodom. Vieme z nich určiť koľkokrát je jedna hodnota väčšia ako druhá. Pomer je nezávislý na jednotkách merania, avšak meranie aj vzťahná hodnota musia byť vyjadrené v rovnakých jednotkách.

Pri spracovaní dát sú častým problémom chýbajúce údaje, ktoré sa nepodarilo pre niektoré objekty získať, alebo sa stratili v priebehu spracovania. Tento problém môžeme riešiť viacerými prístupmi. Príznak, ktorému chýbajú niektoré hodnoty, môžeme pre všetky objekty ignorovať. Ak však rôznym objektom chýbajú hodnoty rôznych príznakov, nemusí byť tento prístup vhodný. Chýbajúci údaj potom môžeme nahradiť priemernou, najčastejšou alebo náhodnou hodnotou. Prípadne môžeme príznaky, ktoré sú riedko obsadené, spojiť do jedného. Potom k nemu pristupujeme ako k nominálnemu príznaku.

V nameraných dátach sa môžu vyskytovať aj hodnoty, ktoré sa objavujú s veľmi nízkou pravdepodobnosťou. Tieto prvky nazývame odľahlé údaje alebo odľahlé pozorovania (outliery). Pri normálnom rozdelení je obvyklé považovať za odľahlú tú hodnotu, ktorá je vzdialená od strednej hodnoty viac ako o 2,4 smerodajných odchýlok. Pravdepodobnosť, že hodnota bude považovaná za odľahlú, je 0,025.

K odľahlým údajom musíme pristupovať premyslene. Ak ich vylúčime z dát, riskujeme stratu užitočnej informácie. Na druhej strane, ak ich zahrnieme do tréningu klasifikátora, riskujeme kontamináciu údajov a zníženie presnosti klasifikácie. Existujú robustné štatistické metódy,



Obrázok 10.1: Vplyv výberu jednotiek merania na vzdialenosť vzoriek.

ktoré berú do úvahy odľahlé údaje, ale minimalizujú ich účinok. Príkladom robustnej opisnej štatistiky je modus alebo medián na rozdiel od priemeru, ktorý robustný nie je.

10.1 Príznačky bez rozmeru

Pri intervalových príznakoch je vzdialenosť v priestore príznakov ovplyvnená výberom jednotiek. Je rozdiel, či meriame dĺžku v centimetroch alebo palcoch, ako vidno na obrázku 10.1. V niektorých riešených problémoch môže byť výhodné príznaky najprv normalizovať, čím získame pomerovú škálu (bezrozmerné príznaky). Normalizácia však môže zmenšiť mieru separácie tried. Treba zistiť, či sú normalizované príznaky vhodnejšie a rozhodnúť sa kvalifikovane.

Bezrozmerný príznak je určený pomer príznaku a určitej referenčnej hodnoty. Referenčná hodnota môže byť najväčšia/najmenšia hodnota v množine, prípadne iná významná hodnota vzťahujúca sa k množine dát (napr. fyzikálne konštanty, jednotky SI, ...).

10.2 Škálovanie

Bezrozmerné príznaky dostaneme aj škálovaním. Poznáme niekoľko základných škálovacích metód.

Lineárne škálovanie na jednotkový interval

Pri škálovaní na jednotkový interval sa maximálna hodnota dát U zobrazí na jednotku a minimálna hodnota L na nulu

$$\tilde{x} = \frac{x - L}{U - L} \quad (10.1)$$

Lineárne škálovanie na jednotkovú dĺžku vektora

Viacrozmerné príznaky môžeme normalizovať na jednotkové vektory

$$\tilde{x}_i = \frac{x_i}{\|x\|}. \quad (10.2)$$

Lineárne škálovanie na nulový priemer a jednotkovú varianciu

Pri normalizácii môžeme využiť aj štatistické vlastnosti príznakov tak, že novovzniknuté dáta budú mať nulový priemer a jednotkovú varianciu

$$\tilde{x}_i = \frac{x_i - \hat{\mu}}{\hat{\sigma}}, \quad (10.3)$$

kde $\hat{\mu}$ je výberový priemer a $\hat{\sigma}$ výberová štandardná odchýlka.

Pravidlo 3σ

Využitím pravidla 3σ môžeme dáta z normálneho rozdelenia škálovať tak, aby 99 % dát ležalo v intervale $\langle 0, 1 \rangle$

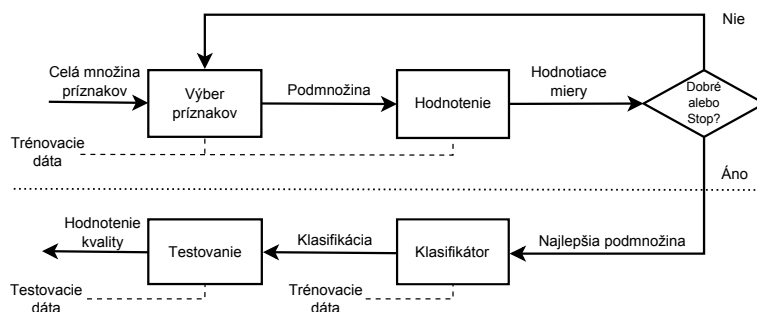
$$\tilde{x}_i = \frac{\frac{x_i - \mu}{3\sigma} + 1}{2}. \quad (10.4)$$

Výber príznačkov

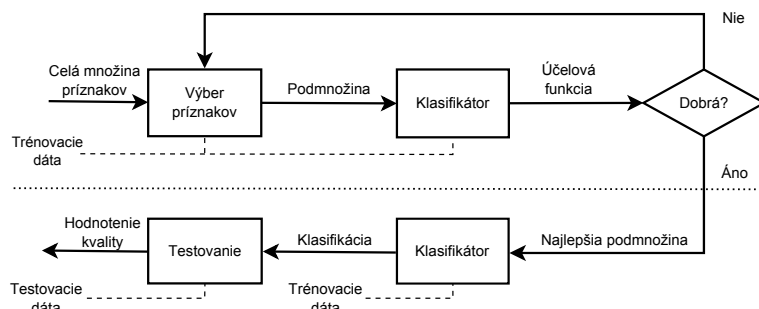
Pri výbere príznačkov ponecháme v príznačkovom vektore, ktorý bude vstupovať do klasifikácie, len relevantné príznačky, ktoré zachytávajú najviac informácie o objektoch. K identifikácii relevantných príznačkov pristupujeme dvomi spôsobmi, ktoré sa líšia tým, či na výber používame konkrétny klasifikátor, alebo nie.

11.1 Filter a obálka

Filter (obr. 11.1) pri výbere príznačkov nepoužíva klasifikátor, ale určuje všeobecnú vhodnosť príznačkov na základe štatistických, geometrických alebo informačno-teoretických mier.



Obrázok 11.1: Postup výberu príznačkov pri použití filtra.



Obrázok 11.2: Postup výberu príznačov pri použití obálky.

Na druhej strane obálka (wrapper, obr. 11.2) zisťuje vhodnosť príznačov na základe výsledkov daného klasifikátora. Príznačky s lepšími výsledkami potom tvoria vybranú podmnožinu.

V oboch prípadoch môžeme hodnotiť jednotlivé príznačky samostatne, alebo hodnotiť celé podmnožiny príznačov.

Hodnotenie jednotlivých príznačov je efektívne, ľahko implementovateľné, ale nemusí odhaliť nadbytočnosť niektorých príznačov, prípadne koreláciu medzi príznačkami. Taktiež je problémom určiť správne počet príznačov, ktoré budú vybraté do výslednej podmnožiny.

Hodnotenie podmnožín je náročnejšie. Pri D príznačkoch máme 2^D podmnožín, ktoré treba ohodnotiť.

Najjednoduchší výber príznačov je dopredná alebo spätná metóda. Rôzne kombinácie dopredného a spätného výberu, metódy exponenciálneho prehľadávania (vetvenie a orezávanie, lúčové hľadanie atď.) alebo randomizované algoritmy (simulované žiňanie, genetické algoritmy atď.) patria už medzi sofistikovanejšie prístupy.

11.2 Dopredný výber

Ak chceme z D príznačov vybrať K ($K < D$), pri jednokrakovom prístupe ohodnotíme všetky príznačky zvolenou mierou a priamo vyberieme K najlepších. Tento prístup málokedy funguje, pretože neberie do úvahy vzťahy medzi príznačkami. Môže nastať situácia, keď príznačky, ktoré boli samostatne ohodnotenú ako najlepší a najhorší, spolu najlepšie klasifikujú dáta. Túto možnosť jednokrakový prístup neodhalí.

Pri iteratívnom prístupe začíname s prázdnu množinou vybraných príznačov $\tilde{X}$. V každom kroku vyberieme príznaček, ktorý najviac obohatí

množinu $\tilde{X}$ a tento do množiny pridáme. Postup opakujeme, kým nedosiahneme K príznakov (pozri algoritmus 11.1). Nevýhodou tohto prístupu je, že niektoré príznaky sa môžu stať nadbytočnými postupom času (po pridaní iných), ale už ich nevieme odstrániť.

```

začni prázdnu množinou  $\tilde{X} = \emptyset$ 
opakuj
  pre každý príznak  $x_i \in X \setminus \tilde{X}$ 
    prerátaj skóre  $\tilde{X} \cup \{x_i\}$ 
  vyber príznak s najvyšším skóre
kým nevyberieš  $K$  príznakov

```

Algoritmus 11.1: Iteratívny dopredný výber príznakov.

11.3 Spätný výber

Pri spätnom výbere naraz alebo postupne odstraňujeme príznaky. Pri jednokrokovom prístupe ohodnotíme všetky príznaky zvolenou mierou a odstránime $D - K$ najhorších.

Pri iteratívnom prístupe začíname s množinou $\tilde{X}$ obsahujúcou všetky príznaky. Postupne z nej odoberáme príznaky, ktorých odstránenie maximalizuje ohodnotenie zvyšných príznakov. Postup opakujeme, kým nedosiahneme K príznakov (pozri algoritmus 11.2). Nevýhodou je, že už odstránené príznaky nemôžeme opätovne pridať, aj keby ich pridanie klasifikáciu zlepšilo.

```

začni celou množinou príznakov  $\tilde{X} = X$ 
opakuj
  pre každý príznak  $x_i \in \tilde{X}$ 
    prerátaj skóre  $\tilde{X} \setminus \{x_i\}$ 
  odstráň príznak, ktorý maximalizuje skóre
kým neodstrániš  $D-K$  príznakov

```

Algoritmus 11.2: Iteratívny spätný výber príznakov.

11.4 Kombinovaný výber

Plus L mínus R výber je generalizáciou dopredného a spätného výberu. Odstraňuje ich nevýhody tým, že zavádza čiastočný backtracking. Problémom zostáva určenie vhodných hodnôt L a R . Pri tejto metóde opakovane pridávame L a odoberáme R príznakov v stanovenom poradí (pozri algoritmus 11.3).

```
- L>R: začni prázdnu množinou
    opakuj
        iteratívne pridaj L príznakov
        iteratívne odober R príznakov
    kým nemáš K príznakov
- L<R: začni celou množinou príznakov
    opakuj
        iteratívne odober R príznakov
        iteratívne pridaj L príznakov
    kým nemáš K príznakov
```

Algoritmus 11.3: Kombinovaný výber príznakov.

11.5 Ohodnotenie príznakov

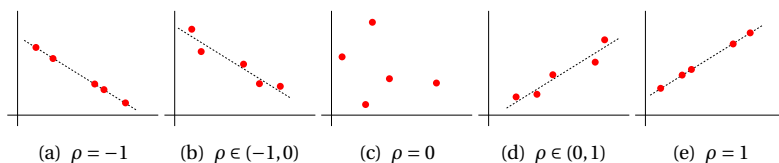
Vhodnosť príznakov závisí na zvolenej miere podobnosti príznakov.

Pri filtri sa používa konzistencia príznakov, medzitriedna vzdialenosť, štatistická závislosť alebo informačno-teoretické miery. Pri jedнокrokovom výbere sa používajú univariátne metódy, ohodnocujúce vždy jeden príznak. Pri iteratívnom postupe používame multivariátne metódy, kde ohodnocujeme množinu príznakov.

Pri obálke sa vyhodnocuje dosiahnutá chyba klasifikátora.

11.5.1 Štatistická závislosť

Mierou štatistickej závislosti je korelačný koeficient. Základná myšlienka takéhoto hodnotenia príznakov je, že príznak (alebo množina príznakov) vhodný na klasifikáciu je silno korelovaný s predikovanou triedou a zároveň nekorelovaný s inými príznakmi.



Obrázok 11.3: Hodnoty korelačného koeficientu ρ blízke ± 1 značia lineárnu závislosť.

Pearsonov (lineárny) korelačný koeficient dvoch premenných X a Y je

$$\rho_{X,Y} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y}, \quad (11.1)$$

kde $\text{Cov}(X, Y)$ je kovariancia premenných (pozri dodatok A.5.3). Tento koeficient nadobúda hodnoty $\langle -1, 1 \rangle$, pričom $\rho_{X,Y} = \pm 1$, ak premenné sú lineárne závislé a $\rho_{X,Y} = 0$, ak sú nekorelované (obr. 11.3).

Nulový korelačný koeficient neznamená nezávislosť dát. Platí to len v prípade dát s Gaussovým rozdelením.

Silná závislosť poukazuje na zbytočnosť niektorých dát.

Nízka korelácia neznamená absenciu vzťahu medzi dátami. Medzi dátami môže existovať nelineárny vzťah.

Korelácia však neznamená kauzalitu. Kladná korelácia znamená, že dané dva javy sa vyskytujú často spolu, ale neznamená, že by jeden jav bol príčinou druhého. Príčinou oboch môže byť iný (aj nepozorovaný) jav.

Multivariálny prístup je nasledovný. Majme množinu $\tilde{X}$ obsahujúcu vzorky reprezentované k vybranými príznakmi, ktoré chceme ohodnotiť. Nech $\bar{r}_{xy}$ je priemer korelačných koeficientov medzi týmito k príznakmi a klasifikačnou triedou a $\bar{r}_{xx}$ je priemer korelačných koeficientov všetkých dvojíc príznakov v $\tilde{X}$. Potom ohodnotenie množiny $\tilde{X}$ pomocou korelačných koeficientov je dané nasledovne

$$J(\tilde{X}) = \frac{k \bar{r}_{xy}}{\sqrt{k + k(k-1) \bar{r}_{xx}}}. \quad (11.2)$$

11.5.2 Konzistencia

Nekonzistencia nastáva, ak vzorky s rovnakými príznakmi v podmnožine patria rôznym triedam. Mieru konzistencie môžeme vyjadriť pomocou

maximálneho počtu nekonzistentných vzoriek patriacich jednej klasifikačnej triede.

Nech v množine $\tilde{X}$ vzoriek reprezentovaných K vybranými príznakmi existuje M rovnakých vzoriek $\mathbf{x}$. Označme m_i počet týchto vzoriek patriacich do triedy ω_i . Potom platí, že $\sum_{i=1}^C m_i = M$, kde C je počet klasifikačných tried. Označme

$$NJ(\mathbf{x}) = M - \max_i m_i, \quad (11.3)$$

potom miera konzistencie množiny $\tilde{X}$ je

$$J(\tilde{X}) = 1 - \frac{\sum_{\mathbf{x} \in \text{Unique}(\tilde{X})} NJ(\mathbf{x})}{N}, \quad (11.4)$$

kde $\text{Unique}(\tilde{X})$ je množina všetkých vzájomne rôznych vzoriek z $\tilde{X}$.

11.5.3 Medzitriedna vzdialenosť

Príznak i budeme preferovať pred príznakom j , ak lepšie separuje klasifikačné triedy, t. j. ak ich vzdialenosť je väčšia pri použití príznaku i .

Množinu príznakov ohodnotíme nasledovne. Označme $d_{\tilde{X}}(\mathbf{x}, \mathbf{y})$ zvolenú metriku (Euklidovská, City Block – Manhattan, Chessboard ... pozri kapitolu 4.2.1). Potom v danej metrike je vzdialenosť dvoch tried pri použití K príznakov

$$D_{\tilde{X}}(\omega_i, \omega_j) = \frac{1}{|\omega_i||\omega_j|} \sum_{\mathbf{x} \in \omega_i} \sum_{\mathbf{y} \in \omega_j} d_{\tilde{X}}(\mathbf{x}, \mathbf{y}). \quad (11.5)$$

Potom množinu $\tilde{X}$ ohodnotíme ako

$$J(\tilde{X}) = \sum_{i=1}^C P(\omega_i) \sum_{j=i+1}^C P(\omega_j) D_{\tilde{X}}(\omega_i, \omega_j). \quad (11.6)$$

11.5.4 Informačno-teoretické miery

Príznak i je vhodnejší na klasifikáciu než príznak j , ak je väčší jeho informačný zisk. Na určenie informačného zisku musíme zdefinovať potrebné pojmy z teórie informácií.

Entropia

V teórii informácií entropia určuje mieru neurčitosti daného javu. Koľko informácie získame, keď sa dozvieme, akú hodnotu y nadobudla náhod-

ná premenná Y ? Miera takejto informácie zodpovedá „stupňu prekvapenia“. Keď sa dozvieme, že premenná Y nadobudla vysoko nepravdepodobnú hodnotu, dostaneme viac informácie, ako keď sa dozvieme, že nadobudla veľmi pravdepodobnú hodnotu. Matematicky to môžeme vyjadriť ako

$$h(y) = -\log_2 p(y), \quad (11.7)$$

kde $p(y) = P(Y = y)$. Záporné znamienko zabezpečuje nezápornú hodnotu informácie.

Ak teraz chceme zistiť, koľko informácie je celkovo v náhodnej premennej, vypočítame to spriemerovaním informácie všetkých hodnôt, ktoré môže náhodná premenná nadobudnúť. Informačná (Shannonova) entropia diskretnej náhodnej premennej Y , ktorá nadobúda hodnoty $\mathcal{Y} = \{y\}$ s pravdepodobnosťou $P(Y = y) = p(y)$ je daná vzťahom

$$H(Y) = - \sum_{y \in \mathcal{Y}} p(y) \log_2 p(y). \quad (11.8)$$

Príklad 11.1: Určíme entropiu binárnej náhodnej premennej Y , pričom $p(1) = P(y = 1) = p$ a $p(0) = P(y = 0) = 1 - p$. Entropia tejto premennej je

$$\begin{aligned} H(Y) &= -(p(1) \log_2 p(1) + p(0) \log_2 p(0)) = \\ &= -p \log_2 p - (1 - p) \log_2 (1 - p). \end{aligned} \quad (11.9)$$

Hodnota entropie je závislá od hodnoty p . Pre $p = 1$ je entropia nulová, čo zodpovedá faktu, že premenná Y nadobúda hodnotu 1 s pravdepodobnosťou 100 % (žiadna neurčitost'). Rovnako pre $p = 0$, kde Y nadobúda hodnotu 0. Pre ostatné hodnoty p je entropia nenulová a maximálna je pre $p = 0,5$, ako vidno na obrázku 11.4. Tento výsledok opäť korešponduje s faktom, že neurčitost' hodnoty, ktorú premenná Y nadobudne, je najvyššia, ak hodnoty 1 a 0 majú rovnakú pravdepodobnosť výskytu. ■

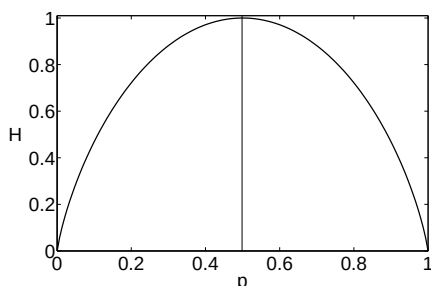
Podmienená entropia

Špecifická podmienená entropia $H(Y|X = x)$ je entropia príznaku (javu) Y počítaná z tých vzoriek, v ktorých príznak X dosahuje hodnotu x

$$H(Y|X = x) = - \sum_{y \in \mathcal{Y}} P(Y = y|X = x) \log_2 P(Y = y|X = x). \quad (11.10)$$

Potom podmienená entropia $H(Y|X)$ je priemernou špecifickou podmienenou entropiou Y

$$H(Y|X) = \sum_{x \in \mathcal{X}} P(X = x) H(Y|X = x). \quad (11.11)$$



Obrázok 11.4: Entropia ako funkcia pravdepodobnosti pre náhodnú premennú s dvomi možnými hodnotami 1 a 0. Entropia je maximálna, keď premenná dosiahne hodnoty 1 a 0 s rovnakou pravdepodobnosťou $p = 0,5$.

Vzájomná informácia

Vzájomná informácia (v literatúre nazývaná aj informačný zisk) udáva, ako sa zníži neurčitosť premennej Y , ak poznáme premennú X

$$I(Y; X) = H(Y) - H(Y|X). \quad (11.12)$$

Pre vzájomnú informáciu platí

$$\begin{aligned} I(Y; X) &= H(Y) - H(Y|X) = \\ &= H(X) - H(X|Y) = I(X; Y). \end{aligned} \quad (11.13)$$

Vzájomná informácia pomáha odhaliť aj nelineárne vzťahy medzi premennými. Ak Y a X sú nezávislé, tak $I(Y; X) = 0$, lebo $H(Y|X) = H(Y)$.

Pri hodnotení množiny príznakov $\tilde{X}$, použijeme multivariátnu vzájomnú informáciu $I(Y; \tilde{X})$

$$J(\tilde{X}) = I(Y; \tilde{X}) = H(Y) - H(Y|\tilde{X}).$$

11.5.5 Hodnotiace miery

Hodnotiace miery používané pre metódu obálky merajú tzv. prediktívnu schopnosť množiny príznakov. Účelovou funkciou, ktorú vyhodnocujeme pre rôzne klasifikátory, je väčšinou odhad chyby klasifikácie (detailne opísaný v kapitole 16), niekedy môže zahŕňať aj cenu natrénovania klasifikátora.

Redukcia príznakov

Pri redukcii transformujeme pôvodnú množinu vektorov príznakov do menej-dimenzionálnej. Pri transformácii sa stráca pôvodný význam príznakov (dĺžka hrany, obsah objektu atď.), na druhej strane premietnutie do menej dimenzionálneho priestoru môže zvýšiť separovateľnosť tried, a tak zlepšiť klasifikáciu.

Vektory príznakov môžeme transformovať bez znalosti klasifikačnej triedy tak, že minimalizujeme stratu informácie pri transformácii. Jednou z takýchto metód je metóda hlavných komponentov.

V prípade, že poznáme klasifikačnú triedu, hľadáme také premietanie, ktoré maximalizuje medzitriednu vzdialenosť. Medzi tieto metódy patrí Fisherova lineárna diskriminačná analýza.

12.1 Metóda hlavných komponentov

Metóda hlavných komponentov (principal component analysis – PCA) otočí súradnicovú sústavu tak, aby prvá os bola v smere najväčšej variability a ďalšie boli na ňu kolmé v smeroch najväčšej zvyšnej variability.

V nasledujúcom texte budeme pod vektorom rozumieť stĺpcový vektor.

Majme N D -rozmerných vektorov príznakov $\mathbf{x}_1, \dots, \mathbf{x}_N$. Postupne nájdeme ortonormálne vektory $\{\mathbf{b}_1, \dots, \mathbf{b}_D\}$, t. j. také, že $\mathbf{b}_i^T \mathbf{b}_j = \delta_{ij}$, kde δ_{ij} je Kroneckerova δ .

Priemetom vektora $\mathbf{x}_i$ do smeru $\mathbf{b}_1$ je $x'_{i1} = \mathbf{b}_1^T \mathbf{x}_i$.

Priemerom všetkých vektorov premietnutých do smeru $\mathbf{b}_1$ je $\bar{x}'_1 = \mathbf{b}_1^T \bar{\mathbf{x}}$, kde $\bar{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i$ a varianciou premietnutých vektorov je

$$\begin{aligned} Var_1 &= \frac{1}{N} \sum_{i=1}^N (x'_{i1} - \bar{x}'_1)^2 = \frac{1}{N} \sum_{i=1}^N (\mathbf{b}_1^T \mathbf{x}_i - \mathbf{b}_1^T \bar{\mathbf{x}})^2 = \\ &= \frac{1}{N} \sum_{i=1}^N (\mathbf{b}_1^T (\mathbf{x}_i - \bar{\mathbf{x}}))^2 = \mathbf{b}_1^T \mathbf{\Sigma} \mathbf{b}_1, \end{aligned} \quad (12.1)$$

kde $\mathbf{\Sigma}$ je kovariančná matica pôvodných príznakov.

Chceme teda nájsť taký vektor $\mathbf{b}_1$, aby sme maximalizovali varianciu Var_1 . Jedná sa o maximalizačnú úlohu s obmedzením – vektor $\mathbf{b}_1$ je jednotkový, preto použijeme Lagrangeovu metódu neurčitých koeficientov.

Majme maximalizačnú úlohu

$$\begin{aligned} \max_{\mathbf{b}_1} \quad & \mathbf{b}_1^T \mathbf{\Sigma} \mathbf{b}_1, \\ \text{aby} \quad & \mathbf{b}_1^T \mathbf{b}_1 = 1. \end{aligned}$$

Nech λ je Lagrangeov neurčitý koeficient, potom maximalizujeme funkciu

$$L = \mathbf{b}_1^T \mathbf{\Sigma} \mathbf{b}_1 - \lambda (\mathbf{b}_1^T \mathbf{b}_1 - 1). \quad (12.2)$$

Položením

$$\frac{\partial L}{\partial \mathbf{b}_1} = 2\mathbf{\Sigma} \mathbf{b}_1 - 2\lambda \mathbf{b}_1 \equiv 0 \quad (12.3)$$

dostaneme $\mathbf{\Sigma} \mathbf{b}_1 = \lambda \mathbf{b}_1$, a to znamená, že $\mathbf{b}_1$ je vlastným vektorom a λ vlastným číslom matice $\mathbf{\Sigma}$. Pretože platí $\mathbf{b}_1^T \mathbf{\Sigma} \mathbf{b}_1 = \lambda \mathbf{b}_1^T \mathbf{b}_1 = \lambda$, potom λ musí byť najväčšie vlastné číslo, aby výraz $\mathbf{b}_1^T \mathbf{\Sigma} \mathbf{b}_1$ bol maximálny. Vlastný vektor $\mathbf{b}_1$ prislúchajúci najväčšiemu vlastnému číslu sa nazýva prvý hlavný komponent.

Vektor $\mathbf{b}_2$ nájdeme obdobne. Tentoraz budeme maximalizovať varianciu $Var_2 = \mathbf{b}_2^T \mathbf{\Sigma} \mathbf{b}_2$. Jedná sa o maximalizačnú úlohu s obmedzením – vektor $\mathbf{b}_2$ je jednotkový a kolmý na vektor $\mathbf{b}_1$, preto opäť použijeme Lagrangeovu metódu neurčitých koeficientov.

Majme maximalizačnú úlohu

$$\max_{\mathbf{b}_2} \quad \mathbf{b}_2^T \mathbf{\Sigma} \mathbf{b}_2, \quad (12.4)$$

$$\text{aby} \quad \mathbf{b}_2^T \mathbf{b}_2 = 1, \quad (12.5)$$

$$\mathbf{b}_1^T \mathbf{b}_2 = 0. \quad (12.6)$$

Nech λ a μ sú Lagrangeove neúčitý koeficienty, potom maximalizujeme funkciu

$$L = \mathbf{b}_2^T \Sigma \mathbf{b}_2 - \lambda(\mathbf{b}_2^T \mathbf{b}_2 - 1) - \mu \mathbf{b}_1^T \mathbf{b}_2. \quad (12.7)$$

Položením

$$\frac{\partial L}{\partial \mathbf{b}_2} = 2\Sigma \mathbf{b}_2 - 2\lambda \mathbf{b}_2 - \mu \mathbf{b}_1 \equiv 0 \quad (12.8)$$

a po úprave dostaneme $\Sigma \mathbf{b}_2 = \lambda \mathbf{b}_2$, teda $\mathbf{b}_2$ je tiež vlastným vektorom matice Σ . Keďže $\mathbf{b}_1$ prislúcha najväčšiemu vlastnému číslu, aby sme maximalizovali Var_2 , $\mathbf{b}_2$ musí prislúchať druhému najväčšiemu vlastnému číslu. Vektor $\mathbf{b}_2$ sa nazýva druhý hlavný komponent.

Týmto postupom dostaneme, že j -ty hlavný komponent $\mathbf{b}_j$ je vlastný vektor prislúchajúci v poradí j -temu najväčšiemu nenulovému vlastnému číslu. Vektory prislúchajúce nulovým vlastným číslam nie sú jednoznačne určené.

Poznáme teda smery otočených bázových vektorov $\{\mathbf{b}_1, \dots, \mathbf{b}_D\}$ a chceme zistiť, kde má ležať nový počiatok súradnicovej sústavy, aby suma štvorcov vzdialeností medzi pôvodnými a premietnutými vektormi bola minimálna.

Nech je počiatočným bodom nejaký bod $\mathbf{p}$, potom súradnice (v pôvodnom priestore) premietnutého vektora $\mathbf{x}_i$ sú $\mathbf{y}_i = \mathbf{p} + \sum_{j=1}^N y_{ij} \mathbf{b}_j$. Suma štvorcov vzdialeností (pri predpoklade ortonormality) je

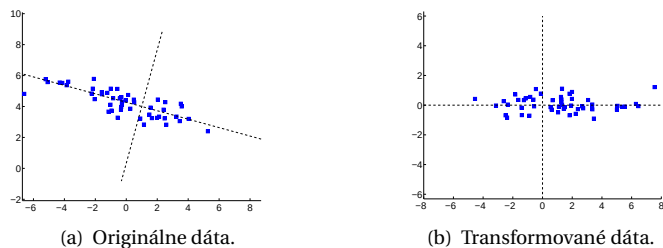
$$\begin{aligned} E &= \sum_{i=1}^N \|\mathbf{x}_i - \mathbf{y}_i\|^2 = \sum_{i=1}^N \|\mathbf{x}_i - (\mathbf{p} + \sum_{j=1}^N y_{ij} \mathbf{b}_j)\|^2 = \\ &= \sum_{i=1}^N \|\mathbf{x}_i - \mathbf{p}\|^2 - 2 \sum_{j=1}^N \sum_{i=1}^N y_{ij} \mathbf{b}_j^T (\mathbf{x}_i - \mathbf{p}) + \sum_{j=1}^N \sum_{i=1}^N y_{ij}^2. \end{aligned} \quad (12.9)$$

Lahko zistíme, že podmienkami minimalizácie sú $y_{ij} = \mathbf{b}_j^T (\mathbf{x}_i - \bar{\mathbf{x}})$ a $\mathbf{p} = \bar{\mathbf{x}}$.

12.1.1 Použitie PCA

Pri použití PCA postupujeme nasledovne:

1. Najprv originálne vektory príznačkov $\{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ vycentrujeme, t. j. odčítame od nich priemer $\bar{\mathbf{x}}$ a posunieme počiatok súradnicovej sústavy do bodu $\bar{\mathbf{x}}$ (dvojrozmerný prípad je ukázaný na obrázku 12.1). Kovariančná matica Σ je rovnaká pre centrované aj pôvodné príznaky, preto môžeme používať PCA na centrované príznaky.

**Obrázok 12.1:** Transformácia dát pomocou PCA.

2. Následne vypočítame maticu Σ

$$\Sigma = \frac{1}{N} \sum_{i=1}^N (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T = \frac{1}{N} \mathbf{X} \mathbf{X}^T, \quad (12.10)$$

kde $[\mathbf{X}]_{D \times N} = [\mathbf{x}_1 - \bar{\mathbf{x}}, \dots, \mathbf{x}_N - \bar{\mathbf{x}}]$.

3. Vypočítame vlastné vektory matice Σ

$$\Sigma \mathbf{b}_j = \lambda_j \mathbf{b}_j. \quad (12.11)$$

4. a nakoniec vypočítame súradnice premietnutých vektorov

$$\mathbf{x}'_i = \sum_{j=1}^D \mathbf{b}_j^T (\mathbf{x}_i - \bar{\mathbf{x}}) = \mathbf{B}^T (\mathbf{x}_i - \bar{\mathbf{x}}) = \mathbf{B}^T \mathbf{X}, \quad (12.12)$$

kde $\mathbf{B} = [\mathbf{b}_1, \dots, \mathbf{b}_D]$.

V praxi sa často stretávame s prípadmi, že počet vektorov príznakov je oveľa menší než dimenzia týchto vektorov $N < D$. Hodnosť matice $\mathbf{X}$ je $r = \text{rank}(\mathbf{X}) = \text{rank}(\Sigma) \leq N$ a najviac r vlastných hodnôt je nenulových. Vtedy je neefektívne počítat' vlastné čísla $D \times D$ matice $\Sigma = \frac{1}{N} \mathbf{X} \mathbf{X}^T$, keď typické algoritmy majú výpočtovú zložitosť $O(D^3)$.

Nech $\mathbf{X}$ je $D \times N$ matica centrovaných príznakov, potom platí

$$\begin{aligned} \Sigma \mathbf{b}_j &= \lambda_j \mathbf{b}_j, \\ \frac{1}{N} \mathbf{X} \mathbf{X}^T \mathbf{b}_j &= \lambda_j \mathbf{b}_j, \\ \frac{1}{N} \mathbf{X}^T \mathbf{X} \mathbf{X}^T \mathbf{b}_j &= \lambda_j \mathbf{X}^T \mathbf{b}_j, \\ \frac{1}{N} \mathbf{X}^T \mathbf{X} \mathbf{p}_j &= \lambda_j \mathbf{p}_j, \end{aligned} \quad (12.13)$$

kde $\mathbf{p}_j = \mathbf{X}^T \mathbf{b}_j$ je vlastný vektor $N \times N$ matice $\frac{1}{N} \mathbf{X}^T \mathbf{X}$. Vypočítame teda vlastné vektory $\mathbf{p}_j$ so zložitou $O(N^3)$ a z nich určíme vektory $\mathbf{b}_j$:

$$\begin{aligned}\frac{1}{N} \mathbf{X}^T \mathbf{X} \mathbf{p}_j &= \lambda_j \mathbf{p}_j, \\ \frac{1}{N} \mathbf{X} \mathbf{X}^T \mathbf{X} \mathbf{p}_j &= \lambda_j \mathbf{X} \mathbf{p}_j, \\ \boldsymbol{\Sigma} \mathbf{X} \mathbf{p}_j &= \lambda_j \mathbf{X} \mathbf{p}_j,\end{aligned}\tag{12.14}$$

kde $\mathbf{X} \mathbf{p}_j$ je vlastný vektor matice $\boldsymbol{\Sigma}$ prislúchajúci vlastnému číslu λ_j , a teda $\mathbf{b}_j \propto \mathbf{X} \mathbf{p}_j$ a zostáva nám učiť normalizačnú konštantu, aby $\|\mathbf{b}_j\| = 1$ za predpokladu, že $\|\mathbf{p}_j\| = 1$

$$\mathbf{b}_j = \frac{1}{\sqrt{N \lambda_j}} \mathbf{X} \mathbf{p}_j.\tag{12.15}$$

12.1.2 Ako môže PCA znížiť dimenziu?

Predpokladajme, že $\boldsymbol{\Sigma}$ má plnú hodnotu a všetky vlastné vektory sú jednoznačne určené. Potom vezmeme len prvých K vlastných vektorov. Zostáva otázkou, aké K zvoliť. Existuje niekoľko metód, ktoré určujú vhodné K . Predpokladajme, že λ_j sú usporiadané zostupne.

Scree graf

Podiel variability zachytenej v j -tom hlavnom komponente je

$$\frac{\lambda_j}{\sum_{j=1}^D \lambda_j}.\tag{12.16}$$

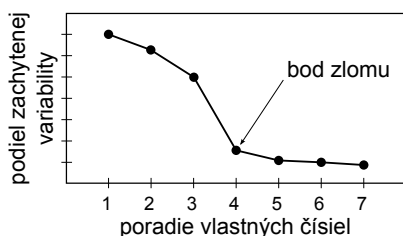
Keď tieto podiely nanesieme na y -ovú os grafu, pričom x -ová os zodpovedá zostupne usporiadaným vlastným číslam, ide o tzv. „scree“ graf (obr. 12.2), kde optimálne K je určené bodom zlomu.

Kumulatívny pomer zachytenej variability

Môžeme použiť aj kumulatívny pomer zachytenej variability

$$\frac{\sum_{j=1}^K \lambda_j}{\sum_{j=1}^D \lambda_j}.\tag{12.17}$$

Pováčšine sa K volí také, aby tento pomer bol väčší ako 0,9 alebo 0,95.



Obrázok 12.2: Scree graf. Vidíme, že v bode zlomu, od štvrtého najväčšieho vlastného čísla začínajú vlastné čísla zachycujúce málo variability. Optimálny počet vlastných čísel je 3.

Kaiserovo kritérium

Ďalšou možnosťou je Kaiserovo kritérium, ktoré volí také K , že hodnota λ_{K+1} je menšia ako priemer hodnôt λ

$$\lambda_{K+1} < \frac{1}{D} \sum_{j=1}^D \lambda_j. \quad (12.18)$$

Andersonov test

Anderson zaviedol testovanie hypotézy, že posledných $D - K$ vlastných čísel je rovnakých

$$H_0 : \lambda_{K+1} = \lambda_{K+2} = \dots = \lambda_D, \quad (12.19)$$

$$H_1 : \text{neplatí } H_0. \quad (12.20)$$

Testovacia štatistika je

$$V = n(D - K) \ln \frac{a}{c}, \quad (12.21)$$

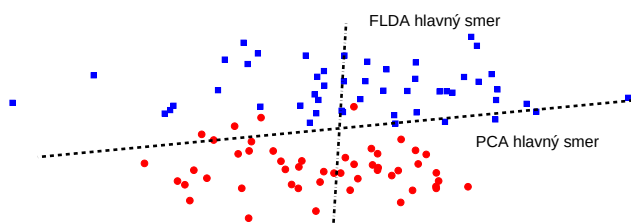
kde

$$a = \frac{\sum_{j=K+1}^D \lambda_j}{D - K} \quad (12.22)$$

a

$$c = \left(\prod_{j=K+1}^D \lambda_j \right)^{\frac{1}{D-K}}. \quad (12.23)$$

Táto veličina má χ^2 rozdelenie s $0,5(D - K + 2)(D - K - 1)$ stupňami voľnosti.



Obrázok 12.3: Dáta dvoch klasifikačných tried s vyznačenými hlavnými smermi určenými metódami PCA a FLDA.

12.1.3 Vzťah medzi PCA a SVD

Na výpočet vlastných vektorov matice Σ môžeme použiť aj singulárny rozklad matice (singular value decomposition – SVD, pozri dodatok B.1).

Ak k matici centrovanej príznačkov $\mathbf{X} = [\mathbf{x}_1 - \bar{\mathbf{x}}, \dots, \mathbf{x}_N - \bar{\mathbf{x}}]$ zostrojíme maticu $\mathbf{Y} = \frac{1}{\sqrt{N}}\mathbf{X}^T$, potom $\mathbf{Y}^T\mathbf{Y} = \frac{1}{N}\mathbf{X}\mathbf{X}^T = \Sigma$.

Teda ak pre SVD matice $\mathbf{Y}$ platí

$$\mathbf{Y} = \mathbf{U}\mathbf{S}\mathbf{V}^T, \quad (12.24)$$

tak $\mathbf{V}$ sú vlastné vektory matice $\mathbf{Y}^T\mathbf{Y} = \Sigma$ a $\mathbf{S}$ má na diagonále odmocninu vlastných čísel matice Σ .

Takto môžeme použiť SVD namiesto PCA a nemusíme vyčíslovať maticu Σ .

12.2 Fisherova lineárna disriminačná analýza

PCA a SVD nemusia byť vždy vhodnými metódami na znižovanie dimenzie, ak potrebujeme príznaky klasifikovať do tried. Patria medzi neriadené metódy, lebo o dátach nemáme informáciu, do ktorej klasifikačnej triedy patria, snažíme sa len odhaliť smery najväčšej variability v dátach.

Metódy, ktoré využívajú informáciu o klasifikačných triedach, nazývame riadené. Medzi riadené metódy patrí aj Fisherova lineárna disriminačná analýza (FLDA). Príklad rozdielu medzi PCA a FLDA transformáciou dát je na obrázku 12.3.

Metóda FLDA maximalizuje medzitriednu variabilitu a zároveň minimalizuje vnútrotriednu variabilitu premietnutých príznakov. Teda príznaky z jednej triedy budú premietnuté blízko pri sebe a príznaky z rôznych tried čo najďalej od seba.

Majme N D -rozmerných vektorov príznakov $\mathbf{x}_1, \dots, \mathbf{x}_N$, ktoré patria do C klasifikačných tried $\{\omega_j\}_{j=1}^C$, pričom triede ω_j patrí N_j príznakov.

Priemer vektorov patriacich triede ω_j je

$$\bar{\mathbf{x}}_j = \frac{1}{|\omega_j|} \sum_{\mathbf{x} \in \omega_j} \mathbf{x} = \frac{1}{N_j} \sum_{\mathbf{x} \in \omega_j} \mathbf{x}. \quad (12.25)$$

Variabilita v jednotlivých triedach je

$$\mathbf{S}_j = \sum_{\mathbf{x} \in \omega_j} (\mathbf{x} - \bar{\mathbf{x}}_j)(\mathbf{x} - \bar{\mathbf{x}}_j)^T. \quad (12.26)$$

Celkový priemer všetkých príznakov je

$$\bar{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i. \quad (12.27)$$

Celková variabilita v dátach je

$$\mathbf{S} = \sum_{i=1}^N (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T. \quad (12.28)$$

Táto variabilita sa dá zapísať ako súčet medzitriednej ($\mathbf{S}_M$) a vnútrotriednej variability ($\mathbf{S}_V$):

$$\mathbf{S} = \mathbf{S}_M + \mathbf{S}_V, \quad (12.29)$$

kde

$$\mathbf{S}_M = \sum_{j=1}^C N_j (\bar{\mathbf{x}}_j - \bar{\mathbf{x}})(\bar{\mathbf{x}}_j - \bar{\mathbf{x}})^T, \quad (12.30)$$

$$\mathbf{S}_V = \sum_{j=1}^C \sum_{\mathbf{x} \in \omega_j} (\mathbf{x} - \bar{\mathbf{x}}_j)(\mathbf{x} - \bar{\mathbf{x}}_j)^T = \sum_{j=1}^C \mathbf{S}_j. \quad (12.31)$$

Povedzme, že chceme nájsť taký smer $\mathbf{w}$, ktorý by nám pomohol klasifikovať príznaky do jednotlivých tried.

Priemetom vektora $\mathbf{x}_i$ do smeru $\mathbf{w}$ je $x'_i = \mathbf{w}^T \mathbf{x}_i$.

Priemerom vektorov patriacich triede ω_j premietnutých do smeru $\mathbf{w}$ je $\bar{x}'_j = \mathbf{w}^T \bar{\mathbf{x}}_j$.

O variabilitách premietnutých príznačov platí:

$$\mathbf{Q} = \mathbf{Q}_M + \mathbf{Q}_V, \quad (12.32)$$

kde

$$\mathbf{Q} = \sum_{i=1}^N (\mathbf{w}^T \mathbf{x}_i - \mathbf{w}^T \bar{\mathbf{x}})(\mathbf{w}^T \mathbf{x}_i - \mathbf{w}^T \bar{\mathbf{x}})^T = \mathbf{w}^T \mathbf{S}_M \mathbf{w}, \quad (12.33)$$

$$\mathbf{Q}_M = \sum_{j=1}^C N_j (\mathbf{w}^T \bar{\mathbf{x}}_j - \mathbf{w}^T \bar{\mathbf{x}})(\mathbf{w}^T \bar{\mathbf{x}}_j - \mathbf{w}^T \bar{\mathbf{x}})^T = \mathbf{w}^T \mathbf{S}_M \mathbf{w}, \quad (12.34)$$

$$\mathbf{Q}_V = \sum_{j=1}^C \sum_{\mathbf{x} \in \omega_j} (\mathbf{w}^T \mathbf{x} - \mathbf{w}^T \bar{\mathbf{x}}_j)(\mathbf{w}^T \mathbf{x} - \mathbf{w}^T \bar{\mathbf{x}}_j)^T = \mathbf{w}^T \mathbf{S}_V \mathbf{w}. \quad (12.35)$$

Fisherova lineárna disriminačná analýza hľadá vektor $\mathbf{w}$, ktorý maximalizuje hodnotu

$$J(\mathbf{w}) = \frac{\mathbf{w}^T \mathbf{S}_M \mathbf{w}}{\mathbf{w}^T \mathbf{S}_V \mathbf{w}}, \quad (12.36)$$

kde Fisherovo kritérium J je skalár pre konkrétny vektor $\mathbf{w}$.

To znamená, že $\mathbf{w}$ maximalizuje medzitriednu a zároveň minimalizuje vnútrotriednu variabilitu premietnutých príznačov.

Položením

$$\frac{dJ}{d\mathbf{w}} = (\mathbf{w}^T \mathbf{S}_M \mathbf{w}) 2 \mathbf{S}_V \mathbf{w} - (\mathbf{w}^T \mathbf{S}_V \mathbf{w}) 2 \mathbf{S}_M \mathbf{w} \equiv 0 \quad (12.37)$$

dostaneme $\mathbf{S}_M \mathbf{w} = J \mathbf{S}_V \mathbf{w}$. Keď predpokladáme, že $\mathbf{S}_V$ je regulárna, potom $\mathbf{S}_V^{-1} \mathbf{S}_M \mathbf{w} = J \mathbf{w}$, a teda J je vlastné číslo a $\mathbf{w}$ je vlastný vektor matice $\mathbf{S}_V^{-1} \mathbf{S}_M$. Vidíme, že J maximalizujeme tak, že za $\mathbf{w}$ vezmeme vlastný vektor zodpovedajúci najväčšiemu vlastnému číslu.

V prípade C klasikačných tried môžeme určiť $C - 1$ projekčných vektorov $\mathbf{w}_j$, ktoré usporiadame do projekčnej matice $\mathbf{W} = [\mathbf{w}_1 | \mathbf{w}_2 | \dots | \mathbf{w}_{C-1}]$. Projekciou príznaču $\mathbf{x}_i$ je potom $\mathbf{y}_i = \mathbf{W}^T \mathbf{x}_i$.

Fisherove kritérium je v tomto prípade

$$J(\mathbf{W}) = \frac{|\mathbf{W}^T \mathbf{S}_M \mathbf{W}|}{|\mathbf{W}^T \mathbf{S}_V \mathbf{W}|}, \quad (12.38)$$

kde J je skalár pre konkrétnu maticu $\mathbf{W}$. $\mathbf{W}^T \mathbf{S}_M \mathbf{W}$ a $\mathbf{W}^T \mathbf{S}_V \mathbf{W}$ sú matice rozmerov $(C - 1) \times (C - 1)$, preto použijeme determinant, aby sme dostali podiel skalárov.

Riešením FLDA sú také vektory $\mathbf{w}_j$, ktoré sú zovšeobecnenými¹ vlastnými vektormi prislúchajúcimi zovšeobecneným vlastným číslam matíc $\mathbf{S}_M$ a $\mathbf{S}_V$, a teda spĺňajú

$$(\mathbf{S}_M - \mathbf{S}_V \lambda) \mathbf{w} = 0. \quad (12.39)$$

Zovšeobecnené vlastné čísla a vektory matíc $\mathbf{S}_M$ a $\mathbf{S}_V$ môžeme vypočítať ako vlastné čísla a vektory matice $\mathbf{S}_V^{-1} \mathbf{S}_M$, ktorá však nie je symetrická, preto je táto úloha náročná. Alebo môžeme použiť nasledovný spôsob.

Maticu $\mathbf{S}_V$, ktorá je štvorcová a symetrická rozložíme na $\mathbf{S}_V = \mathbf{U} \mathbf{\Phi} \mathbf{U}^T$, kde $\mathbf{\Phi} = \text{diag}(\{\varphi_i\}_{i=1}^D)$ má na diagonále vlastné čísla a $\mathbf{U}$ sú ortonormálne vlastné vektory matice $\mathbf{S}_V$. Zadefinujeme matice

$$\mathbf{S}_V^{1/2} = \mathbf{U} \mathbf{\Phi}^{1/2} \mathbf{U}^T, \text{ kde } \mathbf{\Phi}^{1/2} = \text{diag}(\{\varphi_i^{1/2}\}_{i=1}^D), \quad (12.40)$$

$$\mathbf{S}_V^{-1/2} = \mathbf{U} \mathbf{\Phi}^{-1/2} \mathbf{U}^T, \text{ kde } \mathbf{\Phi}^{-1/2} = \text{diag}(\{\varphi_i^{-1/2}\}_{i=1}^D), \quad (12.41)$$

pre ktoré platí $\mathbf{S}_V = \mathbf{S}_V^{1/2} \mathbf{S}_V^{1/2}$, $\mathbf{S}_V^{-1} = \mathbf{S}_V^{-1/2} \mathbf{S}_V^{-1/2}$ a $\mathbf{S}_V^{-1/2} \mathbf{S}_V^{1/2} = \mathbf{I}$. Preto

$$\mathbf{S}_M \mathbf{w} - \mathbf{S}_V \lambda \mathbf{w} = 0, \quad (12.42)$$

$$\mathbf{S}_M (\mathbf{S}_V^{-1/2} \mathbf{S}_V^{1/2}) \mathbf{w} - \lambda \mathbf{S}_V (\mathbf{S}_V^{-1/2} \mathbf{S}_V^{1/2}) \mathbf{w} = 0 \quad | \quad \mathbf{S}_V^{-1/2}, \quad (12.43)$$

$$\mathbf{S}_V^{-1/2} \mathbf{S}_M \mathbf{S}_V^{-1/2} \mathbf{c} - \lambda \mathbf{c} = 0, \quad (12.44)$$

kde $\mathbf{c} = \mathbf{S}_V^{1/2} \mathbf{w}$. Matica $\mathbf{S}_V^{-1/2} \mathbf{S}_M \mathbf{S}_V^{-1/2}$ je symetrická, preto nájdenie vlastných čísel je jednoduchší problém.

¹Pre štvorcové matice $\mathbf{A}$ a $\mathbf{B}$ je zovšeobecnený vlastný vektor $\mathbf{c}$ prislúchajúci zovšeobenému vlastnému číslu λ definovaný ako nenulové riešenie rovnice $(\mathbf{A} - \mathbf{B}\lambda)\mathbf{c} = 0$.



Zoznam použitej literatúry

- [AMK00] S. Abbasi, F. Mokhtarian a J. Kittler. „Enhancing CSS-based shape retrieval for objects with shallow concavities“. *Image and Vision Computing* 18.3 (2000), str. 199–211.
- [AOV12] A. Alahi, R. Ortiz a P. Vandergheynst. „FREAK: Fast retina keypoint“. *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. IEEE. 2012, str. 510–517.
- [Ban+09] H. Bang a kol. „Robust object recognition using a color co-occurrence histogram and the spatial relations of image patches“. *Artificial Life and Robotics* 13.2 (2009), str. 488–492.
- [Bay+08] H. Bay a kol. „Speeded-Up Robust Features (SURF)“. *Computer Vision and Image Understanding* 110.3 (2008), str. 346–359.
- [BH83] J. F. Boyce a W. J. Hossack. „Moment Invariants for Pattern Recognition“. *Pattern Recogn. Lett.* 1.5-6 (júl 1983), str. 451–456.
- [Bis06] C. M. Bishop. *Pattern Recognition and Machine Learning*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006.
- [BM92] P. Besl a N. McKay. „A Method for Registration of 3-D Shapes“. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14 (1992), str. 239–256.

- [Cal+10] M. Calonder a kol. „BRIEF: Binary Robust Independent Elementary Features“. *Computer Vision ? ECCV 2010*. Ed. K. Daniilidis, P. Maragos a N. Paragios. Zv. 6314. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2010, str. 778–792.
- [CBL09] S. A. Chatzichristofis, Y. S. Boutalis a M. Lux. „Img(Rummager): An Interactive Content Based Image Retrieval System“. *Similarity Search and Applications, International Workshop on* (2009), str. 151–153.
- [CC00] L. d. F. Costa a R. Cesar. *Shape Analysis and Classification: Theory and Practice*. Boca Raton, FL, USA: CRC Press, Inc., 2000.
- [CD99] M. Comer a E. Delp. „Segmentation of textured images using a multiresolution Gaussian autoregressive model“. *Image Processing, IEEE Transactions on* 8.3 (1999), str. 408–420.
- [Cin+99] L. Cinque a kol. „Color-based image retrieval using spatial-chromatic histograms“. *Multimedia Computing and Systems, 1999. IEEE International Conference on*. Zv. 2. 1999, 969–973 vol.2.
- [CJ83] G. R. Cross a A. K. Jain. „Markov Random Field Texture Models“. *Pattern Analysis and Machine Intelligence, IEEE Transactions on PAMI*-5.1 (1983), str. 25–39.
- [CM09] J. Chamorro-Martinez a P. Martinez-Jimenez. „A comparative study of texture coarseness measures“. *Image Processing (ICIP), 2009 16th IEEE International Conference on*. 2009, str. 1337–1340.
- [Dav04] E. R. Davies. *Machine Vision: Theory, Algorithms, Practicalities*. 3rd. Elsevier, 2004.
- [DHS01] R. O. Duda, P. E. Hart a D. G. Stork. *Pattern Classification (2nd Edition)*. Wiley-Interscience, 2001.
- [EH11] B. S. Everitt a T. Hothorn. *An Introduction to Applied Multivariate Analysis with R*. New York: Springer-Verlag, 2011.
- [FSZ09] J. Flusser, T. Suk a B. Zitová. *Moments and Moment Invariants in Pattern Recognition*. Wiley & sons, 2009.
- [GSP06] M. Gavrielides, E. Sikudova a I. Pitas. „Color-based descriptors for image fingerprinting“. *Multimedia, IEEE Transactions on* 8.4 (2006), str. 740–748.

- [Hal10] Z. Haladová. „Rozpoznávanie a klasifikácia maliarskych plátien na fotografiách turistov“. Diplomová práca. Bratislava: Univerzita Komenského, 2010.
- [HGS08] T. Hurtut, Y. Gousseau a F. Schmitt. „Adaptive image retrieval based on the spatial organization of colors“. *Computer Vision and Image Understanding* 112.2 (2008), str. 101–113.
- [HH87] P. Hebák a J. Hustopecký. *Vícerozměrné statistické metody s aplikacemi*. SNTL/Alfa, 1987.
- [HS88] C. Harris a M. Stephens. „A combined corner and edge detector“. *In Proc. of Fourth Alvey Vision Conference*. 1988, str. 147–151.
- [HSD73] R. Haralick, K. Shanmugam a I. Dinstein. „Textural Features for Image Classification“. *Systems, Man and Cybernetics, IEEE Transactions on SMC-3.6* (1973), str. 610–621.
- [Hu62] M. K. Hu. „Visual pattern recognition by moment invariants“. *Information Theory, IRE Transactions on* 8.2 (1962), str. 179–187.
- [Hua+97] J. Huang a kol. „Image indexing using color correlograms“. *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*. 1997, str. 762–768.
- [JF90] A. Jain a F. Farrokhnia. „Unsupervised texture segmentation using Gabor filters“. *Systems, Man and Cybernetics, 1990. Conference Proceedings., IEEE International Conference on*. 1990, str. 14–19.
- [KBG10] S. Kiranyaz, M. Birinci a M. Gabbouj. „Perceptual Color Descriptor Based on Spatial Distribution: A Top-down Approach“. *Image Vision Comput.* 28.8 (aug. 2010), str. 1309–1326.
- [KCC89] J. M. Keller, S. Chen a R. M. Crownover. „Texture description and segmentation through fractal geometry“. *Computer Vision, Graphics, and Image Processing* 45.2 (1989), str. 150–166.
- [Kno+10] J. Knopp a kol. „Hough Transform and 3D SURF for Robust Three Dimensional Classification“. *Computer Vision – ECCV 2010*. Ed. K. Daniilidis, P. Maragos a N. Paragios. Zv. 6316. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2010, str. 589–602.
- [LC05] K. L. Lee a L. H. Chen. „An efficient computation method for the texture browsing descriptor of MPEG-7.“ *Image Vision Comput.* 23.5 (2005), str. 479–489.

- [LCS11] S. Leutenegger, M. Chli a R. Y. Siegwart. „BRISK: Binary robust invariant scalable keypoints“. *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE. 2011, str. 2548–2555.
- [LH06] J. Liu a R. Hubbard. „Automatic Camera Calibration and Scene Reconstruction with Scale-Invariant Features“. *Advances in Visual Computing*. Ed. G. Bebis a kol. Zv. 4291. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2006, str. 558–568.
- [Low02] D. G. Lowe. „Distinctive Image Features from Scale-Invariant Keypoints“. *International Journal of Computer Vision* 60.2 (2002), str. 91–110.
- [LP89] F. Lamoš a R. Potocký. *Pravdepodobnosť a matematická štatistika; Štatistické analýzy*. Alfa, 1989.
- [NA08] M. Nixon a A. Aguado. *Feature extraction and image processing*. Academic Press. Academic, 2008.
- [PC96] R. Porter a N. Canagarajah. „A robust automatic clustering scheme for image segmentation using wavelets“. *Image Processing, IEEE Transactions on* 5.4 (1996), str. 662–665.
- [PJW00] D. K. Park, Y. S. Jeon a C. S. Won. „Efficient Use of Local Edge Histogram Descriptor“. *Proceedings of the 2000 ACM Workshops on Multimedia*. MULTIMEDIA '00. Los Angeles, California, USA: ACM, 2000, str. 51–54.
- [PR79] D. Proffitt a D. Rosen. „Metrication errors and coding efficiency of chain-encoding schemes for the representation of lines and edges“. *Computer Graphics and Image Processing* 10.4 (1979), str. 318–332.
- [PR92] J. Prokop a A. P. Reeves. „A survey of moment-based techniques for unoccluded object representation and recognition“. *CVGIP: Graphical Models and Image Processing*. 1992, str. 438–460.
- [Pra01] W. K. Pratt. *Digital Image Processing: PIKS Inside*. New York, NY, USA: John Wiley & Sons, Inc., 2001.
- [PZ96] G. Pass a R. Zabih. „Histogram Refinement for Content-Based Image Retrieval“. *IEEE Workshop on Applications of Computer Vision*. 1996, str. 96–102.

- [RD06] E. Rosten a T. Drummond. „Machine learning for high-speed corner detection“. *European Conference on Computer Vision*. Zv. 1. Máj 2006, str. 430–443.
- [RM08] T. Raykov a G. Marcoulides. *An Introduction to Applied Multivariate Analysis*. Taylor & Francis, 2008.
- [RSZ99] A. Rao, R. Srihari a Z. Zhang. „Spatial color histograms for content-based image retrieval“. *Tools with Artificial Intelligence, 1999. Proceedings. 11th IEEE International Conference on*. 1999, str. 183–186.
- [SB95] S. M. Smith a J. M. Brady. „SUSAN - A New Approach to Low Level Image Processing“. *International Journal of Computer Vision* 23 (1995), str. 45–78.
- [SD96] M. A. Stricker a A. Dimai. „Color indexing with weak spatial constraints“. *Storage and Retrieval for Image and Video Databases (SPIE)*. Zv. 2670. 1996, str. 29–40.
- [Sha+09] H. Shao a kol. „Application Research of Homogeneous Texture Descriptor in Content-Based Image Retrieval“. *Information Engineering and Computer Science, 2009. ICIECS 2009. International Conference on*. 2009, str. 1–4.
- [SHB98] M. Sonka, V. Hlavac a R. Boyle. *Image Processing: Analysis and Machine Vision*. Druhé vyd. CL-Engineering, 30.1998.
- [ST94] J. Shi a C. Tomasi. „Good features to track“. *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94., 1994 IEEE Computer Society Conference on*. Jún 1994, str. 593–600.
- [Tea80] M. R. Teague. „Optical calculation of irradiance moments“. *Appl. Opt.* 19.8 (apr. 1980), str. 1353–1356.
- [TK10] S. Theodoridis a K. Koutroumbas. *Pattern Recognition*. Štvrté vyd. Academic Press, 2010.
- [TLF08] E. Tola, V. Lepetit a P. Fua. „A Fast Local Descriptor for Dense Matching“. *Conference on Computer Vision and Pattern Recognition*. 2008.
- [VS82] A. Vossepoel a A. Smeulders. „Vector code probability and metrication error in the representation of straight lines of finite length“. *Computer Graphics and Image Processing* 20.4 (1982), str. 347–364.
- [WB95] H. Wang a M. Brady. „Real-time corner detection algorithm for motion estimation“. *Image and Vision Computing* 13.9 (1995), str. 695–703.

- [Wu+00] P. Wu a kol. „A texture descriptor for browsing and similarity retrieval“. *Signal Processing: Image Communication* 16.1–2 (2000), str. 33–43.
- [Wu+99] P. Wu a kol. „A texture descriptor for image retrieval and browsing“. *Content-Based Access of Image and Video Libraries, 1999. (CBAIVL '99) Proceedings. IEEE Workshop on*. 1999, str. 3–7.
- [YWB74] I. T. Young, J. E. Walker a J. E. Bowie. „An Analysis Technique for Biological Shape“. *Information and Control* 25 (1974), str. 357–370.
- [ZF03] B. Zitová a J. Flusser. „Image registration methods: a survey“. *Image and Vision Computing* 21 (2003), str. 977–1000.
- [ZL04] D. Zhang a G. Lu. „Review of shape representation and description techniques“. *Pattern Recognition* 37 (2004), str. 1–19.
- [ZZ05] H. Zhang a D. Zhao. „Spatial Histogram Features for Face Detection in Color Images“. *Advances in Multimedia Information Processing - PCM 2004*. Ed. K. Aizawa, Y. Nakamura a S. Satoh. Zv. 3331. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2005, str. 377–384.



ČASŤ III

KLASIFIKÁCIA

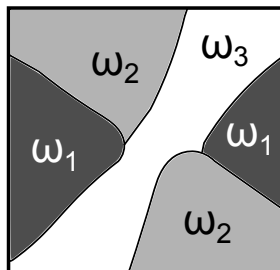
Klasifikácia príznakov

S klasifikáciou sa stretávame každodenne bez toho, aby sme si to uvedomili. Napríklad v obchodnom dome je tovar umiestnený do jednotlivých oddelení na základe určitých klasifikačných kritérií. V oddelení odevov nájdeme napríklad nohavice a košeľe, v oddelení zeleniny rajčiny a šalát, v oddelení ovocia pomaranče a jablká. Táto klasifikácia je robená na základe typu/druhu tovaru. Ak by však tento obchodný dom zaviedol iné klasifikačné kritériá, napríklad za jediný príznak tovaru by zvolil jeho farbu a otvoril by oddelenie zelených tovarov, našli by sme tu spolu jablká, šalát aj košeľe.

Vidíme teda, že alternatívne kritériá výberu príznakov vedú k rôznemu výsledku klasifikácie. Klasifikácia je komplexný proces, ktorého súčasťou je výber príznakov, mier, kritérií, parametrov atď. a nie iba slepé použitie niektorého klasifikátora na všetky dostupné príznaky.

V počítačovom videní klasifikácia znamená porozumenie, z akých typov objektov pochádzajú príznaky, čiže určenie triedy objektu na základe príznakov získaných z objektu. Regresná úloha je podobná klasifikačnej s hlavným rozdielom, že cieľová premenná je číselná. Regresný model môže predpovedať napr. rýchlosť vetra v závislosti od teploty, tlaku a vlhkosti alebo cenu bytu v závislosti od plochy, lokality, vybavenia a pod.

Pri klasifikácii podľa určitého pravidla vieme zaradiť vektor $\mathbf{x}$ opisujúci objekt $\mathcal{O}$ do jednej z možných klasifikačných tried. Toto pravidlo rozdeľuje priestor na oblasti, kde všetky vektory v danej oblasti opisujú objekty z jednej klasifikačnej triedy. Tieto oblasti sa nazývajú rozhodova-



Obrázok 13.1: Rozhodovacie oblasti tried ω_1 a ω_2 nie sú spojité, rozhodovacia oblasť triedy ω_3 je spojitá.

cie oblasti. Rozhodovacia oblasť môže, ale nemusí byť súvislá (obrázok 13.1). Hranice medzi oblasťami sa nazývajú rozhodovacie hranice.

Podľa toho, aké informácie klasifikátory používajú, môžeme ich rozdeliť nasledovne

Príznakové metódy využívajú na klasifikáciu príznaky extrahované z objektov. Medzi tieto metódy patria

- Štatistické metódy, ktorých cieľom je minimalizácia počtu nesprávne klasifikovaných objektov na základe ich štatistických vlastností. Ak pre triedu ω_i platí $P(\mathbf{x}, \omega_i) > P(\mathbf{x}, \omega_{j \neq i})$, zaradíme vektor $\mathbf{x}$ opisujúci objekt $\mathcal{O}$ do triedy ω_i . Bližšie sú opísané v kapitole 14.1.
- Ďalšia množina metód využíva vzdialenosti medzi príznakmi. Sem patria napríklad techniky najbližších susedov (kap. 14.2), diskriminačná analýza (kap. 14.4), mechanizmy podporných vektorov (kap. 14.5), metóda K priemerov (kap. 15.1) alebo hierarchické zhlukovanie (kap. 15.2).
- Medzi biologicky inšpirované metódy patria neurónové siete (kap. 14.6) alebo samoorganizujúce mapy (kap. 15.3).
- Metódou, ktorá využíva pravidlá a dokáže klasifikovať aj nemetrické príznaky, sú rozhodovacie stromy (kap. 14.3).

Syntaktické metódy identifikujú štruktúru objektov. Na opis objektu využívajú elementárne (ďalej nedeliteľné) prvky objektu zvané primitíva a relácie medzi nimi. Takéto relačné štruktúry možno znázorniť pomocou grafov. Základná myšlienka spočíva v tom, že takýto opis objektu tvorí slovo jazyka nad nejakou abecedou. Úlohou

klasifikácie je rozhodnúť pomocou určenia gramatiky, ktorá slovo vygenerovala, do ktorého jazyka skúmané slovo patrí.

Hybridné metódy sú kombináciou predchádzajúcich metód, keď príznakovými metódami identifikujeme primitíva na vytvorenie syntaktického opisu objektu. Mnohé reálne problémy sa riešia hybridným prístupom.

Existuje aj iné roztriedenie klasifikačných metód na riadené a neriadené.

Neriadené metódy. Rozhodovacie oblasti môžeme určovať čisto na základe podobnosti údajov $\mathbf{x}$, bez toho aby sme vedeli, do ktorej triedy opisované objekty skutočne patria. Všetky premenné v modeli sú rovnocenné, žiadna z nich neriadi proces učenia. Takto skonštruované klasifikátory sa nazývajú neriadené a proces ich učenia sa je učením bez učiteľa. Do tejto skupiny patria metóda K priemerov (kap. 15.1), hierarchické zhľukovanie (kap. 15.2), Kohonenove siete (samoorganizujúce mapy) (kap. 15.3)...

Riadené metódy. Ak existuje jedna cieľová premenná, ktorá riadi proces učenia tak, aby ostatné premenné (prediktory) čo najlepšie predpovedali hodnotu cieľovej premennej, hovoríme o riadených klasifikátoroch a o učení sa s učiteľom. Medzi riadené klasifikátory patria: Bayesovské modely (kap. 14.1), techniky najbližších susedov (kap. 14.2), rozhodovacie stromy (kap. 14.3), diskriminačná analýza (kap. 14.4), mechanizmy podporných vektorov (kap. 14.5), neurónové siete (kap. 14.6)...

Pri klasifikácii musíme vyhodnocovať aj dosiahnutú chybu. V prípade dvoch možných klasifikačných tried, chybná klasifikácia nastane, ak vektor $\mathbf{x}$ opisujúci objekt $\mathcal{O}$ patrí regiónu $\mathcal{R}_1$ (klasifikačná trieda ω_1), ale objekt patrí klasifikačnej triede ω_2 a naopak. Matematicky to môžeme zapísať ako

$$\begin{aligned} P(\text{chyba}) &= P(\mathbf{x} \in \mathcal{R}_1, \omega_2) + P(\mathbf{x} \in \mathcal{R}_2, \omega_1) \\ &= \int_{\mathcal{R}_1} P(\mathbf{x}, \omega_2) d\mathbf{x} + \int_{\mathcal{R}_2} P(\mathbf{x}, \omega_1) d\mathbf{x}. \end{aligned} \quad (13.1)$$

Ak sú hodnoty $P(\mathbf{x}, \omega_k)$ porovnateľné (čo je ekvivalentné s podmienkou $P(\omega_k|\mathbf{x}) \ll 1$), zvyšuje sa neistota pri zaradení objektu do tried. V takom prípade je v niektorých klasifikátoroch zaradená možnosť odmietnutia klasifikácie. Napríklad pri rozpoznávaní číslíc bude 11 klasifikačných rozhodnutí $\omega_0, \omega_1, \dots, \omega_9, \omega_{\text{Nezaradené}}$. Nezaradené objekty budú tie,

pre ktoré platí $\forall \omega_k : P(\omega_k | \mathbf{x}) \leq \phi$. Nastavením prahu ϕ vieme ovplyvniť počet nezaradených objektov. Ak nastavíme $\phi = 1$, všetky objektu budú nezaradené a ak pri K klasifikačných triedach nastavíme $\phi < 1/K$, všetky objekty budú zaradené.

V prípadoch, keď je nesprávna klasifikácia rôznych objektov penalizovaná rôzne, nestačí minimalizovať počet nesprávne klasifikovaných objektov. V medicínskej diagnostike je nesprávne zaradenie zdravého pacienta medzi chorých pri nasadení liečby s malými vedľajšími účinkami menej nebezpečné ako neliečenie vážne chorého pacienta. Čiže cena nesprávneho zaradenia do triedy $\omega_{Chorý}$ je nižšia ako cena nesprávneho zaradenia do triedy $\omega_{Zdravý}$. Tomuto problému aj iným metódam ohodnotenia kvality klasifikácie sa budeme venovať v kapitole 16.

Riadené klasifikátory

14.1 Bayesovská teória rozhodovania

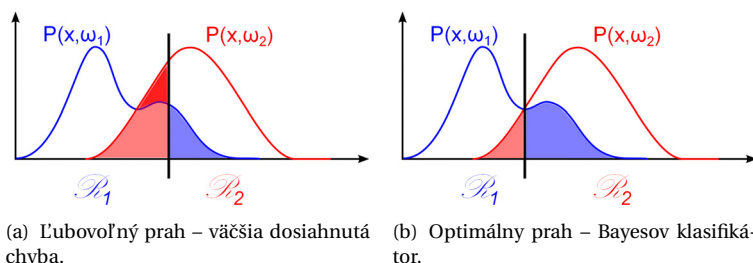
Ak poznáme rozloženie pravdepodobnosti, z ktorého príznaky pochádzajú, alebo ak ho vieme odhadnúť, môžeme na klasifikáciu príznakov použiť Bayesovský prístup. Výhodou tohto prístupu je, že vieme identifikovať prvky, ktoré sa vyskytujú s veľmi nízkou pravdepodobnosťou (odľahlé hodnoty).

Bayesovský klasifikátor klasifikuje objekt do triedy ω_i , do ktorej s najväčšou pravdepodobnosťou patrí vzhľadom na príznaky $\mathbf{x}$. Na to, aby sme túto (aposterioriornu) pravdepodobnosť $P(\omega_i|\mathbf{x})$ vedeli spočítať, potrebujeme poznať

- podmienené rozdelenie pravdepodobnosti $P(\mathbf{x}|\omega_i)$ pre každú klasifikačnú triedu ω_i ,
- apriórnu pravdepodobnosť $P(\omega_i)$ pre každú klasifikačnú triedu ω_i (pred pozorovaním vektora $\mathbf{x}$),
- nepodmienené rozdelenie pravdepodobnosti $P(\mathbf{x})$ – určuje pravdepodobnosť vektora $\mathbf{x}$ bez ohľadu na triedu.

Tieto pravdepodobnosti môžeme určiť zo vzájomnej pravdepodobnosti $P(\mathbf{x}, \omega_i)$.

Naivný Bayesovský klasifikátor vychádza z predpokladu, že efekt, ktorý má hodnota (každého) príznaku na danú triedu, nie je ovplyvnený hodnotami ostatných príznakov.



Obrázok 14.1: Porovnanie dosiahnutej chyby pri ľubovoľnom a optimálnom prahu. Rozdiel oproti optimálnemu prahu je vyznačený tmavšou výplňou.

Na výpočet aposterioriórnych pravdepodobností každej triedy použijeme Bayesovo pravidlo (pozri kapitolu A.1.1)

$$P(\omega_i|\mathbf{x}) = \frac{P(\mathbf{x}|\omega_i)P(\omega_i)}{P(\mathbf{x})}. \quad (14.1)$$

Bayesov klasifikátor potom priradí vzorku $\mathbf{x}$ do triedy ω_i , ak pre každé $j \neq i$ platí

$$\frac{P(\mathbf{x}|\omega_i)P(\omega_i)}{P(\mathbf{x})} \geq \frac{P(\mathbf{x}|\omega_j)P(\omega_j)}{P(\mathbf{x})}. \quad (14.2)$$

Hodnota $P(\mathbf{x}) = \sum_{j=1}^K P(\mathbf{x}|\omega_j)P(\omega_j)$ sa vyskytuje na oboch stranách nerovnice, preto vzorka $\mathbf{x}$ patrí do triedy ω_{i^*} , ktorej rozhodovacia funkcia

$$d_i(\mathbf{x}) = P(\mathbf{x}|\omega_i)P(\omega_i) \quad (14.3)$$

dosiahne najvyššiu hodnotu

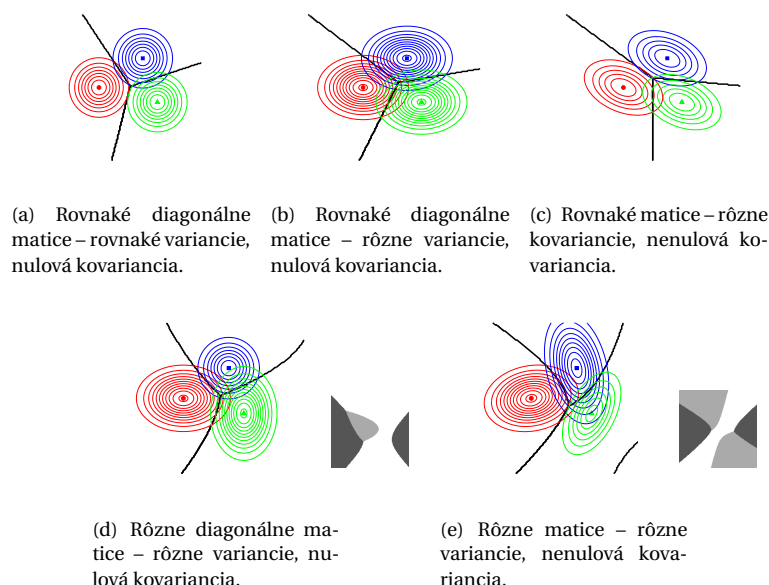
$$i^* = \arg\max_i d_i(\mathbf{x}). \quad (14.4)$$

14.1.1 Optimalita Bayesovho klasifikátora

Pri klasifikácii do dvoch rovnako pravdepodobných tried ($P(\omega_1) = P(\omega_2)$) sa dá dosiahnutá chyba klasifikácie vyjadriť ako

$$P(\text{chyba}) = \int_{\mathcal{R}_1} P(\mathbf{x}, \omega_2) d\mathbf{x} + \int_{\mathcal{R}_2} P(\mathbf{x}, \omega_1) d\mathbf{x}, \quad (14.5)$$

kde $\mathcal{R}_i$ je rozhodovacia oblasť triedy ω_i . Na obrázku 14.1(a) vidíme príklad rozhodovacích oblastí. Chyba, ktorú tento klasifikátor dosiahol, je

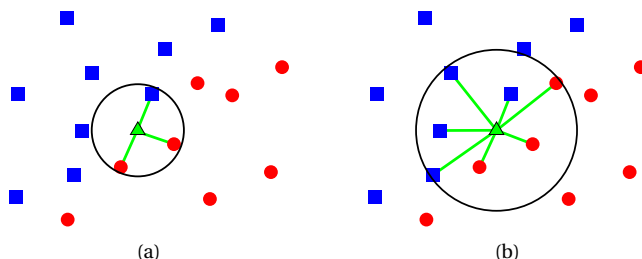


Obrázok 14.2: Rozhodovacie hranice troch klasifikačných tried pre rôzne typy kovariančných matic. V prípade (a), (b) a (c) sú kovariančné matice v triedach rovnaké a rozhodovacie hranice sú lineárne. V prípade (d) a (e) sú kovariančné matice v triedach rôzne a rozhodovacie hranice sú nelineárne. V týchto dvoch prípadoch vidíme aj pohľad na väčšiu oblasť, z ktorej výrez pochádza.

zjednotením červenej a modrej oblasti. Červená vznikne nesprávnou klasifikáciou objektom triedy ω_2 do triedy ω_1 , modrá pri klasifikácii objektov triedy ω_1 do triedy ω_2 . Dosiahnutú chybu minimalizujeme, keď objekty, pre ktoré $P(\mathbf{x}, \omega_1) > P(\mathbf{x}, \omega_2)$ zaradíme do triedy ω_1 a naopak (obr. 14.1(a)). Vieme, že $P(\mathbf{x}, \omega_i) = P(\omega_i|\mathbf{x})P(\mathbf{x}) = P(\mathbf{x}|\omega_i)P(\omega_i)$. Preto do triedy ω_1 zaradíme objekty, pre ktoré $P(\mathbf{x}|\omega_1) > P(\mathbf{x}|\omega_2)$, čo pri predpoklade rovnakej pravdepodobnosti tried zodpovedá Bayesovmu rozhodovaciemu kritériu (14.4).

14.1.2 Rozhodovacie hranice

Tvar rozhodovacej hranice v Bayesovskom rozhodovaní závisí na hodnotách kovariančnej matice (obr. 14.2). V prípade rovnakých kovariančných matic tried sú rozhodovacie hranice lineárne, pri rôznych kovariančných



Obrázok 14.3: Príklad klasifikácie neznámej vzorky metódou najbližších susedov v závislosti na K . (a) Pre $K = 3$ parí neznámy objekt do červenej triedy a (b) pre $K = 7$ do modrej triedy.

maticiach to už neplatí. Obrázok 14.2 obsahuje príklady rôznych typov kovariančných matíc a im zodpovedajúcich rozhodovacích hraníc.

14.2 Techniky najbližších susedov

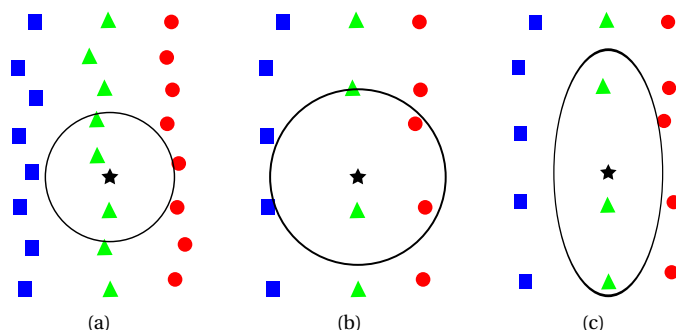
Tieto techniky sa nazývajú podľa toho, že klasifikáciu vzorky ovplyvňujú len jej najbližší susedia. Na obrázku 14.3 vidíme vzorky patriace do dvoch klasifikačných tried a neznámu vzorku, ktorú chceme klasifikovať. V prípade, že o klasifikácii rozhodujú traja najbližší susedia, tak neznáma vzorka patrí do červenej triedy.

Pseudokód metódy najbližších susedov je opísaný algoritmom 14.1. Vidíme, že postup je veľmi jednoduchý, má však veľkú nevýhodu, že pri klasifikácii potrebuje mať k dispozícii všetky trénovacie dáta a pre každú neznámu vzorku musíme spočítať vzdialenosti ku všetkým trénovacím dátam.

```
Nájdí  $K$  najbližších susedov vzorky
Klasifikuj vzorku do triedy,
do ktorej patrí väčšina z  $K$  susedov
```

Algoritmus 14.1: Algoritmus klasifikácie metódou najbližších susedov.

Pri metóde najbližších susedov treba rozhodnúť, ako budeme určovať „blížkosť“ dát, čiže akú metriku použijeme. Toto rozhodnutie urobíme na základe vlastností trénovacích dát. Na obrázku 14.4(a) vidíme, že pri husto vzorkovaných dátach môžeme použiť Euklidovskú metriku



Obrázok 14.4: Príklad vhodnosti metrík pre metódu K najbližších susedov. (a) Pri dostatočne hustom vzorkovaní trénovacích dát je vhodná Euklidovská metrika. (b) Pri redšom vzorkovaní už táto metrika nie je vhodná. (c) Pre takto rozložené dáta sa javí lepšia Mahalanobisova metrika.

a vzorka bude klasifikovaná správne. Toto ale neplatí v prípade riedko vzorkovaných dát na obrázku 14.4(b). V tomto prípade by bolo vhodnejšie zvoliť metriku, ktorá preferuje vzdialenosť v y -ovom smere pred vzdialenosťou v x -ovom smere. Mohli by sme použiť napríklad Mahalanobisovu metriku ako na obrázku 14.4(c).

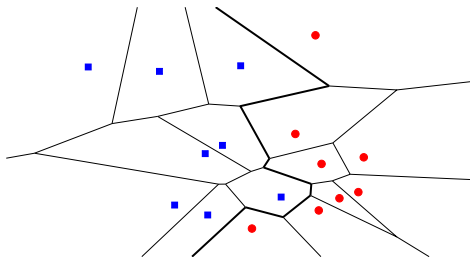
Druhá premenná, ktorú treba určiť, je počet susedov K , ktorí určujú klasifikačnú triedu. Na obrázku 14.3 vidíme, že pri $K = 3$ neznáma vzorka patrí do červenej triedy a pri $K = 7$ do modrej triedy.

Pre $K = 1$ dostávame Voronoiov diagram a hľadáme, do ktorej Voronoiovej bunky vzorka patrí (obrázok 14.5).

Keď sa K zvyšuje, rastie rád Voronoiovhho diagramu, hranica oblastí sa vyhladzuje (obrázok 16.1), môže rásť chyba pri tréningu a klesať chyba pri testovaní. Optimálnu hodnotu K zistíme pomocou vzájomnej validácie (pozri kap. 16).

14.3 Rozhodovacie stromy

Klasifikátory väčšinou určujú príslušnosť na základe vzdialenosti príznakov. Ak sa však medzi príznakmi vyskytujú nominálne dáta, pojem vzdialenosti stráca zmysel. Príkladom môže byť klasifikácia jedla podľa chuti (sladká, slaná, horká ...). V takomto prípade je vhodné použiť na klasifikáciu rozhodovacie stromy. Rozhodnutie o zaradení triedy a postup, ako sme k nemu dospeli, sú usporiadané do stromovej štruktúry, kde uzly



Obrázok 14.5: Metóda K najbližších susedov pre $K = 1$ vytvorí Voronoiov diagram. Pri klasifikácii hľadáme, do ktorej Voronoiovej bunky vzorka patrí.

opisujú testy hodnôt jednotlivých príznakov. Z uzlov vychádza toľko vetiev, koľko rôznych hodnôt test nadobúda, väčšinou sa pracuje s binárnymi stromami. Listy stromu zodpovedajú klasifikačným triedam (obr. 14.6).

Podľa počtu nominálnych hodnôt, ktoré môžu príznaky nadobúdať, vytvárame binárne alebo n-árne stromy. Binárne stromy môžeme vyrobiť aj z príznakov s viacerými hodnotami tak, že sa v uzle budeme postupne pýtať na konkrétne hodnoty (obr. 14.6b)).

Pri vytváraní binárneho rozhodovacieho stromu musíme určiť nasledovné body.

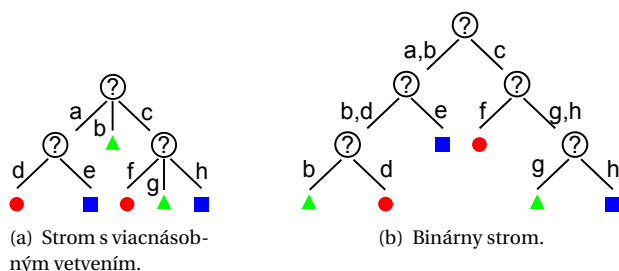
Rozhodovacie porovnanie treba určiť v každom uzle. Porovnanie zisťuje, aký je vzťah hodnoty určeného príznaku k zvolenej prahovej hodnote (napr. $p_j \leq 7$, chuť = sladká). Každý uzol t skúma dáta podmnožiny X_t trénovacej množiny a rozdeľuje ju na dve podmnožiny X_{tA} , X_{tN} . Podmnožina X_{tA} obsahuje dáta, pri ktorých bolo porovnanie splnené (odpoveď na porovnanie je „Áno“), druhá podmnožina obsahuje zvyšné dáta (odpoveď „Nie“). Koreň stromu skúma celú množinu trénovacích dát. V každom uzle platí

$$\begin{aligned} X_{tA} \cap X_{tN} &= \emptyset, \\ X_{tA} \cup X_{tN} &= X_t. \end{aligned}$$

Rozdeľujúce kritérium sa stanovuje v rámci stromu a určuje prahovú hodnotu a príznak, ktoré budú vystupovať v rozhodovacom porovnaní.

Ukončujúce kritérium riadi rast stromu.

Pravidlá určujú klasifikačnú triedu v listoch.



Obrázok 14.6: Príklady rozhodovacích stromov.

Algoritmus 14.2 zobrazuje proces vytvárania jednoduchého binárneho rozhodovacieho stromu, bez orezávania. Orežávaním stromov sa zaoberá kapitola 14.3.3.

```

PROC RozhStrom (X)
  vytvor strom T s jediným vrcholom V
  IF je splnené ukončovacie kritérium THEN
    označ V ako list a zaraď do určenej triedy
  ELSE
    nájdi príznak P a prahovú hodnotu H pomocou
      rozdeľujúceho kritéria
    vyhodnoť rozhodovacie porovnanie
    urči množiny  $X_{tA}$ ,  $X_{tN}$ 
     $TA = \text{RozhStrom}(X_{tA})$ 
     $TN = \text{RozhStrom}(X_{tN})$ 
    spoj uzol V s TA hranou označenou A
    spoj uzol V s TN hranou označenou N
  END IF
  RETURN T

```

Algoritmus 14.2: Základný algoritmus tvorby rozhodovacieho stromu.

Pôvodná myšlienka rozhodovacích stromov pochádza od psychológov Hunta, Marina a Stonea, ktorí v 60. rokoch 20. storočia modelovali koncept ľudského učenia sa. Ich CLS (Concept Learning System) bol binárny rozhodovací strom, ktorý bol navrhnutý tak, aby minimalizoval cenu klasifikácie objektu skladajúcu sa z ceny určenia klasifikačnej triedy a ceny chybnnej klasifikácie.

ID3 (Iterative Dichotomiser 3) je iteratívny algoritmus vyvinutý z CLS. Z náhodne vybranej podmnožiny tréovacích dát (nazývanej okno) sa vytvorí rozhodovací strom a otestuje sa na zvyšných dátach. V prípade, že všetky dáta sú klasifikované správne, proces končí. Ak sú niektoré dáta klasifikované nesprávne, pridajú sa do okna a vytvorí sa nový rozhodovací strom. Takto sa dá nájsť strom klasifikujúci všetky tréovacie dáta správne v priebehu niekoľkých iterácií pre množiny do 30 000 objektov opísaných 50-miestnym vektorom príznakov.

ACLS je generalizáciou ID3. Príznačky môžu nadobúdať neobmedzené celočíselné hodnoty. To umožnilo použitie tohto algoritmu v reálnych problémoch ako je rozpoznávanie obrazov.

C4.5 algoritmus vylepšil ID3 niekoľkými vlastnosťami. Vie narábať so spojitými aj diskretnými dátami, dokáže ošetriť chýbajúce dáta, vie váhovať jednotlivé príznaky a používa spätné orezanie stromu. Jeho ďalším vylepšením je C5.0.

CART (Classification and Regression Trees) konštruuje binárny strom. Uzly sa delia na základe tzv. rozdeľovacieho (twoing) kritéria a strom je následne orezaný. Algoritmus generuje aj regresné stromy, kde v listoch nie je klasifikačná trieda, ale reálne číslo. Pri regresných stromoch sa v uzloch minimalizuje stredná kvadratická chyba.

14.3.1 Rozdeľujúce kritérium

Rozdeľujúce kritérium ohodnocuje príznaky a do uzlov sa priradí test na najdôležitejší príznak. Dôležitosť príznaku môžeme merať „nedokonanosťou rozdelenia“, ktorú vyjadríme hodnotou vzájomnej informácie, entropie atď. Dokonalé rozdelenie v uzle nastane, ak všetky objekty v jednej vetve patria do tej istej klasifikačnej triedy.

Každé z kritérií je vhodné pre iný druh problémov. Viacerí autori však argumentujú, že voľba rozdeľujúceho kritéria v mnohých prípadoch nie je podstatná a má len malý vplyv na kvalitu výslednej klasifikácie.

Entropia

Delenie na základe minimálnej entropie (kap. 11.5.4) sa používa v algoritme ID3.

Giniho index

Giniho index použitý v algoritme CART udáva ako často bude náhodne vybraný objekt nesprávne klasifikovaný, ak bude náhodne klasifikovaný

podľa rozdelenia tried v množine. Majme m klasifikačných tried. Nech je v uzle U relatívna početnosť objektov jednotlivých tried

$$p_i = \frac{|\{\mathbf{x} \mid \mathbf{x} \in \omega_i\}|}{|U|}. \quad (14.6)$$

Potom

$$GI = \sum_{i=1}^m p_i(1 - p_i) = 1 - \sum_{i=1}^m p_i^2, \quad (14.7)$$

kde p_i zodpovedá pravdepodobnosti výberu objektu z triedy i a $(1 - p_i)$ pravdepodobnosti nesprávnej klasifikácie tohto objektu. Ak sú všetky dáta v jednej triede, Giniho index dosahuje minimálnu hodnotu 0.

Kolmogorovove–Smirnovove kritérium

Kolmogorovove–Smirnovove (KS) kritérium hodnotí maximálny rozdiel distribučných funkcií. Pre triedy ω_i a ω_j s kumulatívnymi distribučnými funkciami $F_i(x)$ a $F_j(x)$ je KS kritérium je dané nasledovne

$$KS(\omega_i, \omega_j) = \max_x |F_i(x) - F_j(x)|. \quad (14.8)$$

Viacnásobný test

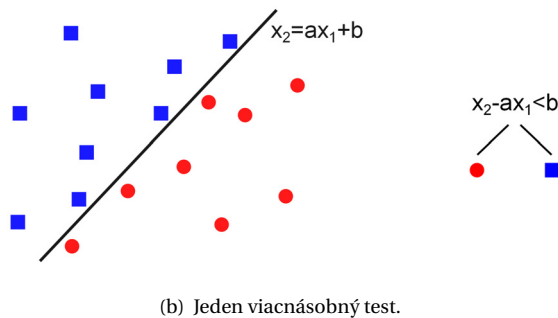
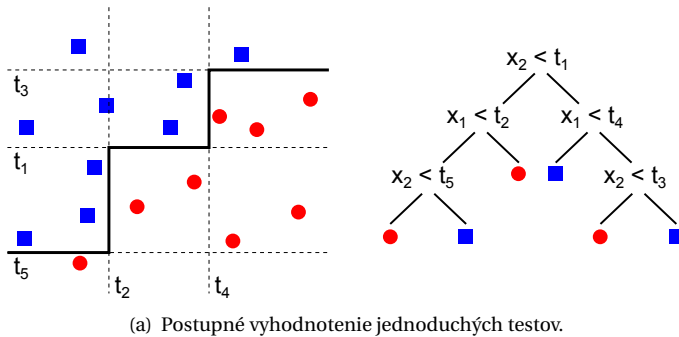
Do testov môže vstupovať viacero príznakov spoločne. Väčšina viacnásobných testov je lineárnou kombináciou príznakov. Na nájdenie najlepšej kombinácie sa používa pažravý algoritmus, lineárne programovanie alebo lineárna diskriminačná analýza.

Použitie viacnásobného testu znižuje hĺbku rozhodovacieho stromu a vyhladzuje rozhodovacie hranice, ako vidno na obrázku 14.7.

14.3.2 Ukončujúce kritérium

Medzi bežné ukončovacie kritériá patria:

- Všetky objekty v skúmanej množine patria do jednej klasifikačnej triedy.
- Strom dosiahol maximálnu stanovenú hĺbku.
- Počet objektov klasifikovaných v danom uzle je menší ako stanovený prah.
- Ohodnotenie najlepšieho príznaku je nižšie ako stanovený prah.



Obrázok 14.7: Viacnásobný test znižuje hĺbku rozhodovacieho stromu.

14.3.3 Orezávanie rozhodovacích stromov

Vylepšením základného algoritmu 14.2 je tzv. orezávanie rozhodovacieho stromu, ktoré zahŕňa odstránenie listov alebo celých podstromov. Ak odstránenie prinesie zvýšenie kvality klasifikácie, zvolené uzly sa odstránia a ich rodič sa zmení na list, do ktorého sa priradí klasifikačná trieda podľa najpravdepodobnejšieho objektu z trénovacej množiny objektov, ktoré po tento list spadajú.

Stromy môžeme orezávať už počas vytvárania, alebo orezávame hotový strom. Orezanie počas generovania rozhodovacieho stromu je zahrnuté v ukončovacom kritériu „Ohodnotenie najlepšieho príznaku je nižšie ako stanovený prah“. Toto kritérium zastaví vetvenie stromu, aj keď dáta v podstromu nepatria len do jednej triedy. Môže však nastať situácia, že delenie o úroveň nižšie by malo pozitívny vplyv na kvalitu klasifikácie.

Pri orezávaní hotových stromov sa dáta rozdelia na tréningovú a testovaciu (validačnú) množinu. Pomocou tréningovej množiny sa vytvorí strom a validačná množina sa použije na hodnotenie kvality klasifikácie.

Orezávanie uzlov

Použitím validačnej množiny sa vyráta chyba klasifikácie stromu. Každý uzol sa berie ako možný kandidát na odstránenie. Pre každý uzol sa zráta, ako sa zmení chyba klasifikácie, keby sa tento uzol odstránil (anglický názov metódy je *reduced-error pruning*). Odstránia sa jednotlivé uzly, pri ktorých sa chyba klasifikácie zníži. Ak takéto uzly neexistujú, odstránia sa uzly, pri ktorých sa chyba nemení. Ak by odstránenie ľubovoľného uzlu prinieslo zvýšenie chyby klasifikácie, všetky uzly zostávajú v strome a proces orezávania sa skončí. Inak sa prepočíta chyba klasifikácie v novovzniknutom orezanom strome a proces sa opakuje.

Orezávanie podstromov

Táto metóda orezávania (po anglicky *cost-complexity pruning*) ohodnocuje zmenu kvality klasifikácie pri odstránení celého podstromu a jeho náhrade listovým uzlom. Ak sa chyba po náhrade listovým uzlom zmenší, daný podstrom je možné odrezať.

V prvej fáze orezávania sa vygenerujú stromy $T_0, \dots, T_N$, kde T_0 je originálny strom a T_N je strom tvorený len koreňom. V druhej fáze sa vypočíta validačná chyba pre všetky vytvorené stromy $T_0, \dots, T_N$. Strom s najnižšou chybou je potom výsledný klasifikačný strom.

Postupnosť stromov $T_0, \dots, T_N$ sa vytvára orezávaním podstromov, ktoré vykazujú najnižšiu chybu (*cost-complexity, CC*) pri odstránení. Majme strom T , ktorý klasifikuje dáta s chybou $Err(T)$. *CC* stromu T je

$$CC(T) = Err(T) + \alpha L(T), \quad (14.9)$$

kde $L(T)$ je počet listov stromu T . Ak odstránime zo stromu T podstrom S , tento novovytvorený strom T_S bude mať

$$CC(T_S) = Err(T_S) + \alpha L(T_S), \quad (14.10)$$

pričom $L(T_S) = L(T) - L(S) + 1$. Tieto dva stromy budú mať rovnakú *CC*, ak

$$\alpha = \frac{Err(T_S) - Err(T)}{L(S) - 1}. \quad (14.11)$$

Strom T_{i+1} je tvorený zo stromu T_i nahradením podstromu, ktorý minimalizuje hodnotu α .

14.3.4 Výhody a nevýhody rozhodovacích stromov

Medzi hlavné výhody rozhodovacích stromov patrí ich jednoduchá interpretácia, aj pre laických používateľov rôznych klasifikátorov. Ďalej je to možnosť klasifikácie numerických a aj kategorických dát a nízka výpočtová náročnosť. Na druhej strane, vytvorenie optimálneho rozhodovacieho stromu je NP úplný problém a používané heuristiky nezaručujú optimalitu vytvoreného stromu. Rozhodovacie stromy sú vhodné v problémoch, kde existuje niekoľko veľmi relevantných príznakov, inak môže vzniknúť priveľmi komplexné stromy obsahujúce viaceré kópie určitých podstromov.

14.4 Diskriminačná analýza

Jednou z úloh diskriminačnej analýzy (DA) je určiť, ktoré príznaky sú dôležité pri klasifikácii objektov do jednotlivých tried (pozri kapitolu 12).

Ďalšou úlohou DA je na základe trénovacej množiny vytvoriť pravidlá, ktoré každý nový objekt zaradia do jednej zo skupín tak, aby chyba klasifikácie bola čo najmenšia. Táto DA môže byť lineárna alebo kvadratická.

V tejto kapitole predpokladáme, že klasifikačné triedy sú lineárne separovateľné, t. j. existuje nadrovina rozdeľujúca priestor tak, že vektory rôznych tried ležia v rôznych polpriestoroch.

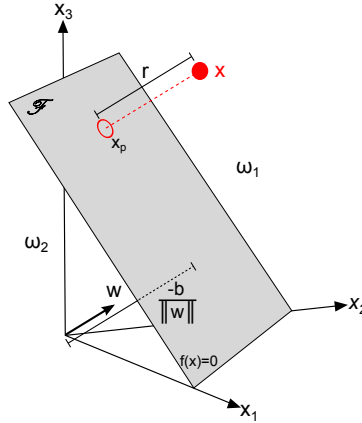
14.4.1 Lineárne diskriminačné funkcie

Výhodou lineárnych diskriminačných funkcií (lineárnych klasifikátorov) je ich jednoduchosť a nízka výpočtová náročnosť. Budeme sa zaoberať funkciami, ktoré sú lineárnou kombináciou zložiek $\mathbf{x}$, čo sa dá napísať ako

$$f(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b, \quad (14.12)$$

kde $\mathbf{w}$ je váhový vektor a b je prah (alebo posunutie).

Pri technikách najbližších susedov sú všetky body trénovacej množiny potrebné na určenie triedy každého nového bodu. Pri lineárnom klasifikátore nie je trénovacia množina po určení (natrénovaní) vektora $\mathbf{w}$ naďalej potrebná.



Obrázok 14.8: Nadrovina $\mathcal{F}$ určená parametrami $\mathbf{w}$ a b . Vzdialenosť bodu $\mathbf{x}$ od nadroviny je r .

14.4.2 Klasifikácia do dvoch tried

Majme vektory príznakov patriace dvom rôznym klasifikačným triedam ω_1 a ω_2 . Rovnica nadroviny je

$$\mathbf{w}^T \mathbf{x} + b = 0. \quad (14.13)$$

Hodnota $f(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b$ je nulová pre body ležiace na rozdeľujúcej nadrovine, záporná resp. kladná pre body ležiace mimo nadroviny.

Úlohou lineárneho klasifikátora je nájsť taký váhový vektor $\mathbf{w}$ (normálový vektor nadroviny) a prah b , aby platilo:

$$f(\mathbf{x}) \geq 0 \text{ pre } \mathbf{x} \in \omega_1, \quad (14.14)$$

$$f(\mathbf{x}) < 0 \text{ pre } \mathbf{x} \in \omega_2. \quad (14.15)$$

Tvrdenie 1. Hodnota $\frac{f(\mathbf{x})}{\|\mathbf{w}\|}$ určuje geometrickú vzdialenosť bodu $\mathbf{x}$ od nadroviny $\mathcal{F}$ určenej parametrami $\mathbf{w}$ a b (pozri obrázok 14.8).

Dôkaz. Nech $\mathbf{x}_p$ je kolmý priemet $\mathbf{x}$ do $\mathcal{F}$. Potom $\mathbf{x}$ sa dá vyjadriť ako

$$\mathbf{x} = \mathbf{x}_p + r \frac{\mathbf{w}}{\|\mathbf{w}\|}, \quad (14.16)$$

kde $|r|$ je vzdialenosť od priemetu. Hodnota r je kladná, ak $\mathbf{x}$ leží v rovnanom polpriestore ako vektor $\mathbf{w}$. Vieme, že $f(\mathbf{x}_p) = 0$ (bod $\mathbf{x}_p$ leží na $\mathcal{F}$).

Potom

$$\begin{aligned}
 f(\mathbf{x}) &= \mathbf{w}^T \mathbf{x} + b \\
 &= \mathbf{w}^T \left(\mathbf{x}_p + r \frac{\mathbf{w}}{\|\mathbf{w}\|} \right) + b \\
 &= \mathbf{w}^T \mathbf{x}_p + b + r \frac{\mathbf{w}^T \mathbf{w}}{\|\mathbf{w}\|} \\
 &= f(\mathbf{x}_p) + r \frac{\|\mathbf{w}\|^2}{\|\mathbf{w}\|} \\
 &= r \|\mathbf{w}\|,
 \end{aligned} \tag{14.17}$$

a teda

$$r = \frac{f(\mathbf{x})}{\|\mathbf{w}\|}. \tag{14.18}$$

□

Podobne vieme ukázať, že b je vzdialenosť $\mathcal{F}$ od počiatku súradnicovej sústavy.

14.4.3 Určenie váh

Aby sme jednoduchšie zistili váhy a prah, zapíšeme lineárny klasifikátor ako

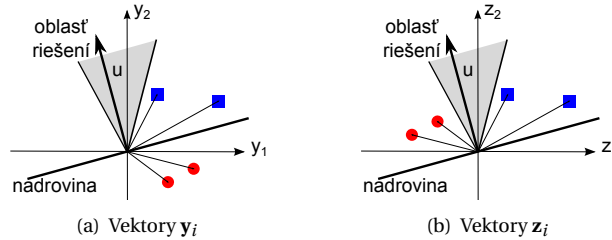
$$f(\mathbf{x}) = \mathbf{u}^T \mathbf{y}, \tag{14.19}$$

kde $\mathbf{u} = [b, \mathbf{w}^T]^T$ a $\mathbf{y} = [1, \mathbf{x}^T]^T$.

Pri určovaní váhového vektora $\mathbf{u}$ zo vzťahu (14.19), nedelíme problém na samostatné hľadanie vektora $\mathbf{w}$ a prahu b .

Takýto prechod do $D + 1$ rozmerného priestoru (nazvime ho $\mathbf{y}$ -priestor) zachováva vzdialenosti pôvodných vektorov $\mathbf{x}$ aj u nových vektorov $\mathbf{y}$. Rozdeľujúca nadrovina $\hat{\mathcal{F}}$ definovaná ako $\mathbf{u}^T \mathbf{y} = 0$ prechádza počiatkom súradnicovej sústavy $\mathbf{y}$ -priestoru, hoci v pôvodnom priestore nadrovina $\mathcal{F}$ mohla byť umiestnená ľubovoľne. Vzdialenosť vektora $\mathbf{y}$ od $\hat{\mathcal{F}}$ je $\frac{f(\mathbf{x})}{\|\mathbf{u}\|}$. Táto vzdialenosť je menšia alebo rovná vzdialenosti vektora $\mathbf{x}$ od $\mathcal{F}$, lebo $\|\mathbf{u}\| \geq \|\mathbf{w}\|$.

Majme N vzoriek $\mathbf{y}_1, \dots, \mathbf{y}_N$ patriacich dvom rôznym klasifikačným triedam ω_1 a ω_2 . Vzorku $\mathbf{y}_i$ zaradíme do triedy ω_1 , ak $\mathbf{u}^T \mathbf{y}_i > 0$ a do triedy



Obrázok 14.9: Vektory $\mathbf{y}_i$ po normalizácii podľa vzťahu (14.20). Všetky dáta ležia v jednom podpriestore určenom nadrovinou.

ω_2 , ak $\mathbf{u}^T \mathbf{y}_i < 0$. Ak by sme zaviedli nasledovnú normalizáciu

$$\mathbf{z}_i = \begin{cases} \mathbf{y}_i & \text{pre } \mathbf{y} \in \omega_1, \\ -\mathbf{y}_i & \text{pre } \mathbf{y} \in \omega_2, \end{cases} \quad (14.20)$$

tak hodnota $\mathbf{u}^T \mathbf{z}_i > 0$ pre všetky $\mathbf{z}_i$ (pozri obrázok 14.9).

Naším problémom je teda nájsť váhový vektor $\mathbf{u}$, aby platilo $\mathbf{u}^T \mathbf{z}_i > 0$ pre všetky $\mathbf{z}_i$. Existuje niekoľko metód ako tento vektor určiť. Zavedieme skalárnu účelovú funkciu $O(\mathbf{u})$, ktorej hodnota je minimálna, keď $\mathbf{u}$ je riešením nášho problému. Nájdenie takého $\mathbf{u}$ sa rieši gradientnou metódou.

Gradientná metóda je iteratívna. Začíname s náhodne zvoleným vektorom $\mathbf{u}_1$, vypočítame gradient ∇O v bode $\mathbf{u}_1$ a aktualizujeme hodnotu $\mathbf{u}$ tak, že vyberieme vektor z okolia $\mathbf{u}_1$ v smere najväčšej negatívnej zmeny gradientu

$$\mathbf{u}_{i+1} = \mathbf{u}_i - \eta(i) \nabla O|_{\mathbf{u}_i}, \quad (14.21)$$

kde $\eta(i)$ je (pozitívna) rýchlosť učenia v kroku i , ktorá určuje veľkosť kroku. Výber vhodných veľkostí je dôležitý, lebo ak zvolíme nízku hodnotu, konvergencia je pomalá a ak zvolíme príliš vysokú hodnotu, proces môže správne riešenie „preskočiť“ a divergovať. Ak predpokladáme, že účelová funkcia je diferencovateľná, môžeme pomocou Taylorovho rozvoja určiť rýchlosť učenia nasledovne. Funkcia $O(\mathbf{a})$ je v okolí bodu $\mathbf{u}_i$ aproximovaná

$$O(\mathbf{a}) \approx O(\mathbf{u}_i) + \nabla O^T (\mathbf{a} - \mathbf{u}_i) + \frac{1}{2} (\mathbf{a} - \mathbf{u}_i)^T \mathbf{H} (\mathbf{a} - \mathbf{u}_i), \quad (14.22)$$

kde $\mathbf{H}$ je matica druhých parciálnych derivácií funkcie O (Hesseho matica). $\mathbf{H}$ a ∇O sú vyhodnotené v bode $\mathbf{u}_i$ (rovnako v ďalšom texte). Pre

$\mathbf{a} = \mathbf{u}_{i+1}$ dostávame

$$O(\mathbf{u}_{i+1}) \approx O(\mathbf{u}_i) - \eta(i) \|\nabla O\|^2 + \frac{1}{2} \eta^2(i) \nabla O^T \mathbf{H} \nabla O. \quad (14.23)$$

Lahko ukážeme, že hodnota (14.23) je minimalizovaná pre

$$\eta(i) = \frac{\|\nabla O\|^2}{\nabla O^T \mathbf{H} \nabla O}. \quad (14.24)$$

Druhou možnosťou na hľadanie riešenia $\mathbf{u}$ je použitie Newtonovho algoritmu, pričom namiesto vzťahu 14.21 použijeme na zmenu hodnoty $\mathbf{u}_i$ vzťah

$$\mathbf{u}_{i+1} = \mathbf{u}_i - \mathbf{H}^{-1} \nabla O. \quad (14.25)$$

14.4.4 Voľba účelovej funkcie

Pozrime sa teraz na výber účelovej funkcie. Naším cieľom je minimalizácia počtu nesprávne klasifikovaných vzoriek. Takáto funkcia je po častiach konštantná, preto nie je vhodná na použitie v gradientnej metóde. Lepšou voľbou je po častiach lineárna perceptrónová funkcia

$$O_P = \sum_{\mathbf{z} \in \mathcal{Z}} (-\mathbf{u}^T \mathbf{z}), \quad (14.26)$$

kde $\mathcal{Z}$ je množina nesprávne klasifikovaných vzoriek. V prípade, keď je táto množina prázdna, definujeme $O_P \equiv 0$. Vieme, že pre nesprávne klasifikované vzorky platí $\mathbf{u}^T \mathbf{z} \leq 0$, a preto O_P nadobúda nezáporné hodnoty. Nulovú hodnotu nadobúda pre váhové vektory, ktoré sú riešením nášho problému, keď nemáme nesprávne klasifikované vzorky.

Pre funkciu definovnú vzťahom 14.26 platí

$$\nabla O_P = \sum_{\mathbf{z} \in \mathcal{Z}} (-\mathbf{z}) \quad (14.27)$$

a vzťah (14.21) nadobudne tvar

$$\mathbf{u}_{i+1} = \mathbf{u}_i + \eta(i) \sum_{\mathbf{z} \in \mathcal{Z}_i} \mathbf{z}, \quad (14.28)$$

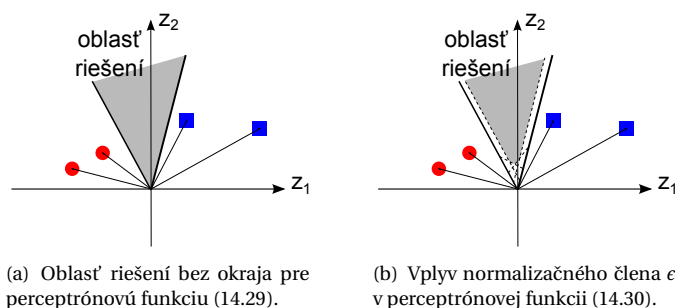
kde $\mathcal{Z}_i$ je množina nesprávne klasifikovaných vzoriek v kroku i . Tomuto prístupu sa hovorí dávkový, pretože vyhodnocuje všetky neklasifikované vzorky naraz.

Jednoduchším riešením je vyhodnocovať vzorky po jednej, ako je to uvedené v algoritme 14.3.

```

inicializuj  $\mathbf{u}_0 = \mathbf{1}$ ,  $i=0$ ,  $k=0$ 
opakuaj
    ak  $\mathbf{z}_k$  je nesprávne klasifikovaný
         $\mathbf{u}_{i+1} = \mathbf{u}_i + \mathbf{z}_k$ 
         $i=i+1$ 
    koniec ak
     $k=(k+1) \bmod N$ 
kým existuje nejaký nesprávne klasifikovaný vektor  $\mathbf{z}_j$ 

```

Algoritmus 14.3: Algoritmus hľadania váhového vektora.**Obrázok 14.10:** Oblasť riešení perceptrónovej funkcie.

Podobnou účelovou funkciou je kvadratická funkcia

$$O_Q = \sum_{\mathbf{z} \in \mathcal{Z}_i} (\mathbf{u}^T \mathbf{z})^2, \quad (14.29)$$

ktorá má na rozdiel od O_P spojitý gradient. Problémom tejto funkcie je, že je možno až príliš hladká a gradientná metóda môže konvergovať na hranici oblasti riešení (pre $\mathbf{a} = \mathbf{0}$). Oblasť riešení je znázornená na obrázku 14.10(a).

Druhým problémom je, že dlhšie vektory vzoriek $\mathbf{z}$ vplývajú na hodnotu O_Q viac, ako kratšie.

Oba tieto problémy rieši funkcia obsahujúca normalizačný člen $\|\mathbf{z}\|^2$ a člen ϵ určujúci okraj v oblasti riešení

$$O_R = \frac{1}{2} \sum_{\mathbf{z} \in \mathcal{Z}_i} \frac{(\mathbf{u}^T \mathbf{z} - \epsilon)^2}{\|\mathbf{z}\|^2}. \quad (14.30)$$

Vplyv týchto členov na oblasť riešení je znázornený na obrázku 14.10(b). V tomto prípade $\mathcal{Z}_i$ je množina vzoriek, pre ktoré platí $\mathbf{u}^T \mathbf{z} \leq \epsilon$. Gradient tejto funkcie je

$$\nabla O_R = \sum_{\mathbf{z} \in \mathcal{Z}_i} \frac{\mathbf{u}^T \mathbf{z} - \epsilon}{\|\mathbf{z}\|^2} \mathbf{z}. \quad (14.31)$$

Pri použití účelovej funkcie (14.30) nastáva nesprávna klasifikácia, ak $\mathbf{u}^T \mathbf{z} \leq \epsilon$. Algoritmus 14.3 sa preto zmení na algoritmus 14.4.

```

inicializuj  $\mathbf{u}_0 = 0$ ,  $i=0$ ,  $k=0$ 
opakuj
   $k=(k+1) \bmod N$ 
  ak  $\mathbf{u}^T \mathbf{z}_k \leq \epsilon$ 
     $\mathbf{z} = \mathbf{z}_k$ 
     $\mathbf{u}_{i+1} = \mathbf{u}_i + \eta(i) \frac{\epsilon - \mathbf{u}^T \mathbf{z}}{\|\mathbf{z}\|^2} \mathbf{z}$ 
     $i=i+1$ 
  koniec ak
kým  $\mathbf{u}^T \mathbf{z}_j > \epsilon$  pre všetky  $\mathbf{z}_j$ 

```

Algoritmus 14.4: Algoritmus hľadania váhového vektora s okrajom.

14.4.5 Klasifikácia do K tried

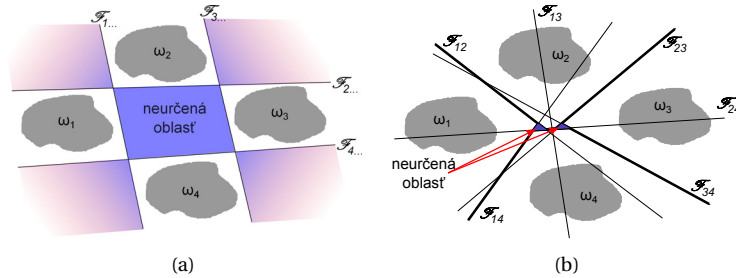
Pri klasifikácii do viacerých tried môžeme postupovať dvomi spôsobmi.

1. Pre každú klasifikačnú triedu vytvoríme rozhodovaciu funkciu, ktorá oddelí vzorky tejto triedy od vzoriek ostatných tried. V tomto prípade máme K rozhodovacích funkcií.
2. Pre každú dvojicu rôznych klasifikačných tried vytvoríme rozhodovaciu funkciu oddelujúcu vzorky týchto dvoch tried bez ohľadu na vzorky ostatných tried. V tomto prípade máme $K(K-1)/2$ rozhodovacích funkcií.

Oba tieto prístupy môžu viesť k situácii, kde časť priestoru príznakov nie je zaradená do žiadnej klasifikačnej triedy (pozri obrázok 14.11).

Tento problém vyrieši vytvorenie množiny K rozhodovacích funkcií $\{f_k(\mathbf{x})\}_{k=1}^K$ a zaradenie vzorky $\mathbf{x}$ do tej klasifikačnej triedy, ktorej hodnota po dosadení je najväčšia

$$\mathbf{x} \in \omega_i \Leftrightarrow \forall_{j \neq i} f_i(\mathbf{x}) > f_j(\mathbf{x}). \quad (14.32)$$



Obrázok 14.11: Klasifikácia do viacerých tried (a) oddelením jednej triedy od ostatných a (b) oddelením dvojíc tried.

Takto vytvorený klasifikátor sa nazýva lineárny stroj. Tento stroj delí priestor na K konvexných regiónov. Ak sa dva regióny $\mathcal{R}_i$ a $\mathcal{R}_j$ dotýkajú, potom hranica medzi týmito regiónmi leží na nadrovine $\mathcal{F}_{ij}$ danej rovnicou $f_i(\mathbf{x}) = f_j(\mathbf{x})$, a teda platí

$$\mathbf{w}_i^T \mathbf{x} + b_i = \mathbf{w}_j^T \mathbf{x} + b_j, \quad (14.33)$$

$$(\mathbf{w}_i - \mathbf{w}_j)^T \mathbf{x} + b_i - b_j = 0. \quad (14.34)$$

Podľa (14.18) je vzdialenosť vzorky $\mathbf{x}$ od $\mathcal{F}_{ij}$ daná

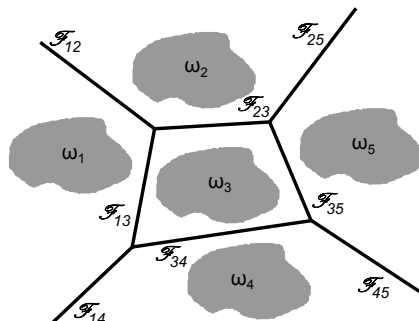
$$\frac{f_i(\mathbf{x}) - f_j(\mathbf{x})}{\|\mathbf{w}_i - \mathbf{w}_j\|}. \quad (14.35)$$

Vidíme, že v prípade lineárneho stroja nie sú dôležité samostatné váhové vektory, ale ich rozdiel. Pre K klasifikačných tried existuje $K(K-1)/2$ nadrovín $\mathcal{F}_{ij}$, ale nie všetky regióny musia byť susedné, preto hraníc medzi regiónmi je väčšinou menej (pozri obrázok 14.12).

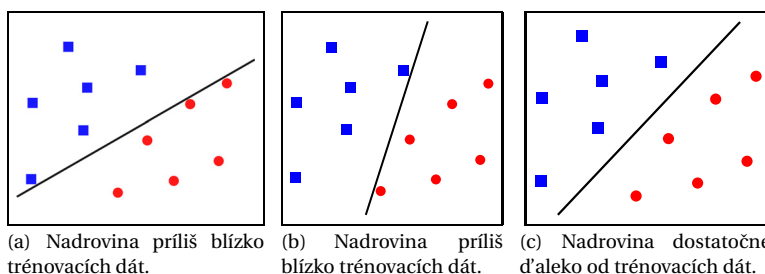
V prípade lineárneho stroja, ak by bola každá trieda reprezentovaná jednou vzorkou, klasifikácia je rovnaká ako pri technike jedného najbližšieho suseda (pozri kap. 14.2, obrázok 14.5).

14.5 Mechanizmy podporných vektorov

Lineárny klasifikátor z predchádzajúcej kapitoly vyberie niektorý z váhových vektorov ležiacich v oblasti riešení. Výsledkom môže byť rozdeľujúca nadrovina, ktorá síce dobre klasifikuje tréningové dáta, ale pri testovacích dátach dosahuje veľkú chybu (pozri obr. 14.13). Čiastočným rieše-



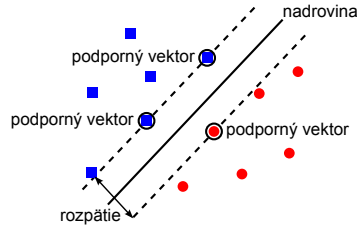
Obrázok 14.12: Lineárny stroj pre 5 klasifikačných tried.



Obrázok 14.13: Nadrovina určená perceptrónom môže byť príliš blízko tréningovým dátam.

ním je použité účelovej funkcie O_R , avšak optimálna hodnota parametra c nie je určená.

Tento problém riešia mechanizmy podporných vektorov (support vector machines - SVM). SVM určujú rozdeľujúcu nadrovinu tak, aby testovacie dáta mali čo najväčšiu voľnosť. Snažia sa nájsť nadrovinu čo najviac vzdialenú od tréningových vzoriek (pozri obr. 14.13(c)). Podpornými vektormi sú potom tie vzorky každej triedy, ktoré sú najbližšie k rozdeľujúcej nadrovine. Vzdialenosť medzi podpornými vektormi rôznych tried sa nazýva rozpätie (margin, pozri obr. 14.14). Našou úlohou je nájsť takú nadrovinu reprezentovanú váhovým vektorom, ktorá túto vzdialenosť maximalizuje. Rovnica nadroviny má tvar (14.13). Bez ujmy na všeobecnosti môžeme predpokladať, že pre podporné vektory platí $\mathbf{w}^T \mathbf{x} + b = \pm 1$ (podľa príslušnosti k triede ω_1 , resp. ω_2). Potom podľa vzťahu (14.18) je



Obrázok 14.14: Podpornými vektormi sú tie vzorky každej triedy, ktoré sú najbližšie k rozdeľujúcej nadrovine. Vzdialenosť medzi podpornými vektormi rôznych tried sa nazýva rozpätie (margin).

veľkosť rozpätia rovná

$$\frac{2}{\|\mathbf{w}\|}. \quad (14.36)$$

Naša optimalizačná úloha je teda

$$\mathbf{w}^* = \operatorname{argmax} \frac{2}{\|\mathbf{w}\|}. \quad (14.37)$$

To je ekvivalentné k minimalizačnej úlohe

$$(\mathbf{w}^*, b^*) = \operatorname{argmin} \frac{1}{2} \|\mathbf{w}\|^2, \quad (14.38)$$

aby

$$\mathbf{w}^T \mathbf{x}_i + b \geq 1 \text{ pre } \mathbf{x} \in \omega_1, \quad (14.39)$$

$$\mathbf{w}^T \mathbf{x}_i + b \leq -1 \text{ pre } \mathbf{x} \in \omega_2. \quad (14.40)$$

Podmienky (14.39) a (14.40) môžeme vyjadriť jednou podmienkou

$$(\mathbf{w}^T \mathbf{x}_i + b) k_i \geq 1, \quad (14.41)$$

kde

$$k_i = \begin{cases} 1 & \text{pre } \mathbf{x} \in \omega_1, \\ -1 & \text{pre } \mathbf{x} \in \omega_2. \end{cases} \quad (14.42)$$

Minimalizačnú úlohu (14.38) vyriešime pomocou Lagrangeovej metódy

neurčitých koeficientov

$$\begin{aligned} L(\mathbf{w}, b, \alpha_i) &= \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{i=1}^N \alpha_i ((\mathbf{w}^T \mathbf{x}_i + b) k_i - 1) \\ &= \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{i=1}^N \alpha_i (\mathbf{w}^T \mathbf{x}_i + b) k_i + \sum_{i=1}^N \alpha_i, \end{aligned} \quad (14.43)$$

aby

$$\alpha_i \geq 0. \quad (14.44)$$

Riešenie hľadáme pomocou derivácií Lagrangianu

$$\frac{\partial L}{\partial \mathbf{w}} = \mathbf{w} - \sum_{i=1}^N \alpha_i k_i \mathbf{x}_i \equiv 0 \quad (14.45)$$

$$\frac{\partial L}{\partial b} = - \sum_{i=1}^N \alpha_i k_i \equiv 0 \quad (14.46)$$

$$(\mathbf{w}^T \mathbf{x}_i + b) k_i - 1 \geq 0 \quad (14.47)$$

$$\alpha_i \geq 0 \quad (14.48)$$

$$\alpha_i ((\mathbf{w}^T \mathbf{x}_i + b) k_i - 1) = 0 \quad (14.49)$$

Podmienky (14.45) – (14.49) sa nazývajú Kuhn-Tuckerove (KT) podmienky. Z (14.45) vidíme, že váhový vektor vypočítame ako

$$\mathbf{w} = \sum_{i=1}^N \alpha_i k_i \mathbf{x}_i. \quad (14.50)$$

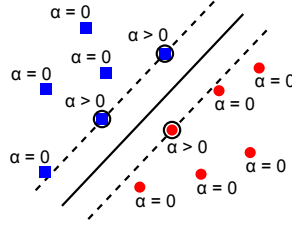
Ďalej zo vzťahov (14.47) – (14.49) vyplýva

$$\alpha_i = 0 \Leftrightarrow (\mathbf{w}^T \mathbf{x}_i + b) k_i - 1 > 1, \quad (14.51)$$

$$\alpha_i > 0 \Leftrightarrow (\mathbf{w}^T \mathbf{x}_i + b) k_i - 1 = 1, \quad (14.52)$$

čiže pre body ležiace vnútri oblasti ω_1 alebo ω_2 sú Lagrangeove multiplikátory nulové (vzťah (14.51)) a pre body ležiace na okraji oblasti (podporné vektory) sú multiplikátory nenulové (vzťah (14.52)). Táto situácia je znázornená na obrázku 14.15. Preto váhový vektor ovplyvňujú len podporné vektory $\mathbf{x}_i \in SV$

$$\mathbf{w} = \sum_{\mathbf{x}_i \in SV} \alpha_i k_i \mathbf{x}_i. \quad (14.53)$$



Obrázok 14.15: Lagrangeove multiplikátory sú nenulové pre podporné vektory.

Aby sme našli hodnoty α_i , vyriešime duálny problém. Do rovnice (14.43) dosadíme vzťah (14.45) a s využitím vzťahu (14.46) dostaneme duálnu maximalizačnú úlohu k minimalizačnej úlohe (14.38).

$$\alpha^* = \operatorname{argmax} \left(\sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j k_i k_j \mathbf{x}_i^T \mathbf{x}_j \right), \quad (14.54)$$

aby

$$\sum_{i=1}^N \alpha_i k_i = 0, \quad (14.55)$$

$$\alpha_i \geq 0. \quad (14.56)$$

Ku klasifikácii potom pristupujeme nasledovne

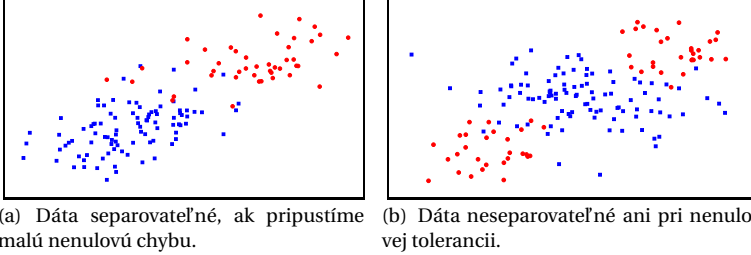
1. Úlohu (14.54) vyriešime metódami kvadratického programovania.
2. Určíme množinu podporných vektorov $\mathbf{x}_i \in SV$, pre ktoré $\alpha_i > 0$.
3. Pomocou vzťahu (14.47) určíme hodnotu

$$b = \frac{1}{|SV|} \sum_{\mathbf{x}_j \in SV} \left(k_j - \sum_{\mathbf{x}_i \in SV} \alpha_i k_i \mathbf{x}_i^T \mathbf{x}_j \right). \quad (14.57)$$

4. Dosadením (14.53) do (14.12) zistíme hodnotu klasifikačnej funkcie

$$f(\mathbf{x}) = \sum_{\mathbf{x}_i \in SV} \alpha_i k_i \mathbf{x}_i^T \mathbf{x} + b. \quad (14.58)$$

V nasledovnej kapitole sa pozrieme, ako pristupovať k neseparovateľným dátam.



Obrázok 14.16: Dva typy lineárne neseparovateľných dát.

14.5.1 Neseparovateľné dáta

Neseparovateľnosť dát môže mať viac príčin. V prípade, že rozdelenie pravdepodobnosti dát v jednotlivých triedach sa prekrýva, dáta nie sú lineárne separovateľné s nulovou chybou (obrázok 14.16(a)). Ak zoslabíme podmienku nulovej tolerancie chyby, môže použiť modifikovaný SVM klasifikátor s voľným rozpätím.

Druhým typom dát sú dáta, ktoré nie sú lineárne separovateľné ani pri nenulovej tolerancii chyby (obrázok 14.16(b)). Tento problém rieši ne-lineárny SVM klasifikátor.

SVM s voľným rozpätím

Podmienku nulovej chyby zoslabíme zavedením druhotných premenných ξ_i , ktoré určujú voľnosť chybnéj klasifikácie. S týmito premennými sa minimalizačná úloha (14.38) zmení na

$$(\mathbf{w}^*, b^*, \xi_i^*) = \arg \min \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^N \xi_i, \quad (14.59)$$

aby

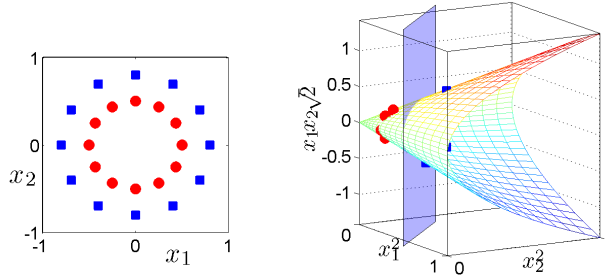
$$(\mathbf{w}^T \mathbf{x}_i + b) k_i \geq 1 - \xi_i, \quad (14.60)$$

kde konštanta C určuje penalizáciu za chybnú klasifikáciu a ovplyvňuje veľkosť rozpätia.

Pri riešení tohto problému opäť prejdeme cez Lagrangián k duálnemu problému (14.54). Táto duálna úloha sa nemení, mení sa len podmienka (14.56) na

$$0 \leq \alpha_i \leq C. \quad (14.61)$$

Vhodnú hodnotu C nájdeme vzájomnou validáciou (pozri kap. 16).



Obrázok 14.17: Dáta na obrázku vľavo nie sú v 2D priestore lineárne separovateľné. Po prechode do 3D priestoru už lineárne separovateľné sú, ako vidno na obrázku vpravo.

Nelineárny SVM

Druhou možnosťou riešenia pre lineárne neseparovateľné dáta je prejsť do viacrozmerného priestoru. Na obrázku 14.17 vľavo vidíme príklad dát, ktoré sú separovateľné, ale nie lineárne. Ak by sme tieto dáta previedli do trojrozmerného priestoru predpisom

$$\phi: \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \rightarrow \begin{pmatrix} x_1^2 \\ x_2^2 \\ \sqrt{2}x_1x_2 \end{pmatrix}, \quad (14.62)$$

dáta by boli separovateľné. Táto situácia je znázornená na obrázku 14.17 vpravo.

Duálna maximalizačná úloha (14.54) pre transformované dáta je nasledovná

$$\alpha^* = \operatorname{argmax} \left(\sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j k_i k_j \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j) \right), \quad (14.63)$$

aby

$$\sum_{i=1}^N \alpha_i k_i = 0, \quad (14.64)$$

$$\alpha_i \geq 0. \quad (14.65)$$

Po vyriešení tejto úlohy určíme pomocou vzťahu (14.47) hodnotu

$$b = \frac{1}{|SV|} \sum_{\mathbf{x}_j \in SV} \left(k_j - \sum_{\mathbf{x}_i \in SV} \alpha_i k_i \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j) \right). \quad (14.66)$$

Hodnota klasifikačnej funkcie je

$$f(\mathbf{x}) = \sum_{\mathbf{x}_i \in SV} \alpha_i k_i \phi(\mathbf{x}_i)^T \phi(\mathbf{x}) + b. \quad (14.67)$$

Vidíme, že vo všetkých vzťahoch vystupuje vektor $\phi(\mathbf{x}_i)$ vždy len v skalárnom súčine s iným transformovaným vektorom. Táto skutočnosť sa dá využiť tak, že nemusíme určiť ako presne vyzerá transformácia ϕ , ale stačí určiť ako vyzerá skalárny súčin v transformovanom priestore

$$K(\mathbf{x}, \mathbf{y}) = \phi(\mathbf{x})^T \phi(\mathbf{y}). \quad (14.68)$$

Úloha (14.63) sa zmení na

$$\alpha^* = \arg \max \left(\sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j k_i k_j K(\mathbf{x}_i, \mathbf{x}_j) \right). \quad (14.69)$$

Hodnotu b vypočítame ako

$$b = \frac{1}{|SV|} \sum_{\mathbf{x}_j \in SV} \left(k_j - \sum_{\mathbf{x}_i \in SV} \alpha_i k_i K(\mathbf{x}_i, \mathbf{x}_j) \right) \quad (14.70)$$

a hodnotu klasifikačnej funkcie ako

$$f(\mathbf{x}) = \sum_{\mathbf{x}_i \in SV} \alpha_i k_i K(\mathbf{x}_i, \mathbf{x}) + b. \quad (14.71)$$

Takže namiesto transformácie každého príznakového vektora do viacrozmerného priestoru počítame hodnotu funkcie $K(\cdot)$ pôvodných vektorov. Tento výpočtový trik sa nazýva kernelový trik a funkcia $K(\cdot)$ je kernelová funkcia alebo jadro.

Príklad 14.1: Určime funkciu $K(\cdot)$ pre transformačnú funkciu zo vzťahu (14.62).

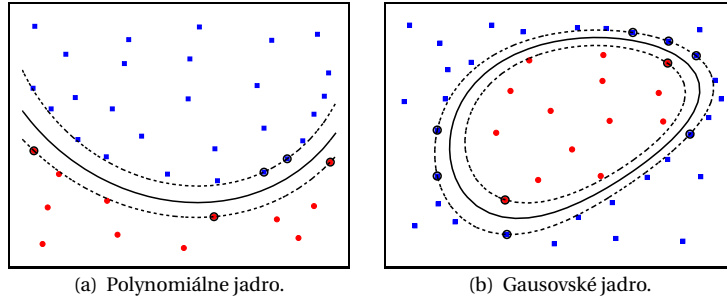
$$K(\mathbf{x}, \mathbf{y}) = (x_1^2, x_2^2, \sqrt{2}x_1x_2) \cdot \begin{pmatrix} y_1^2 \\ y_2^2 \\ \sqrt{2}y_1y_2 \end{pmatrix} = \quad (14.72)$$

$$= x_1^2 y_1^2 + x_2^2 y_2^2 + 2x_1 y_1 x_2 y_2 = \quad (14.73)$$

$$= (x_1 y_1 + x_2 y_2)^2 = \quad (14.74)$$

$$= (\mathbf{x}^T \mathbf{y})^2 \quad (14.75)$$

■



Obrázok 14.18: Klasifikácia lineárne neseparovateľných dát pomocou nelineárnych SVM s rôznymi jadrami.

Na to, aby funkcia $K(\cdot)$ bola platnou kernelovou funkciou, musí existovať funkcia $\phi(\cdot)$ taká, aby platil vzťah (14.68). Podľa Mercerovej podmienky, každá symetrická, kladne semidefinitná funkcia je platnou kernelovou funkciou. Funkcia je symetrická, ak

$$K(\mathbf{x}_i, \mathbf{x}_j) = K(\mathbf{x}_j, \mathbf{x}_i) \quad (14.76)$$

a kladne semidefinitná ak

$$\forall \mathbf{a} \in \mathbb{R}^N, \mathbf{a}^T \mathbf{K} \mathbf{a} \geq 0, \quad (14.77)$$

kde $\mathbf{K}_{N \times N}$ je Grammova matica s prvkami $k_{ij} = K(\mathbf{x}_i, \mathbf{x}_j)$.

Medzi najpoužívanéjšie jadrá patria polynomiálne jadrá

$$K(\mathbf{x}, \mathbf{y}) = (a + \mathbf{x}^T \mathbf{y})^p \quad (14.78)$$

a Gaussovské jadrá

$$K(\mathbf{x}, \mathbf{y}) = e^{-\frac{\|\mathbf{x} - \mathbf{y}\|^2}{2\sigma^2}}. \quad (14.79)$$

Obrázok 14.18 ukazuje použitie týchto jadier na klasifikáciu lineárne neseparovateľných dát.

14.6 Neurónové siete

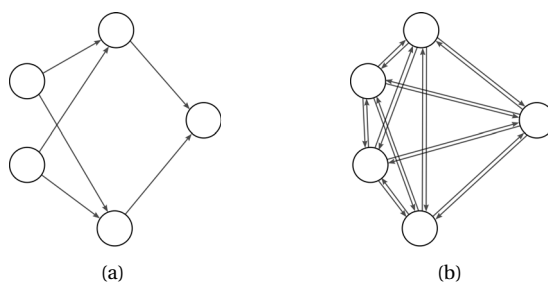
Neurónové siete (artificial neural networks – ANN) sú siete pospájaných uzlov (neurónov), medzi ktorými sa presúva informácia. Každý uzol vyšle ďalej informáciu na základe vstupu, ktorý dostane. Postup je opísaný algoritmom 14.5.


```

pre uzol U
  výstupy z uzlov predchádzajúcej vrstvy tvoria
    vstupy do uzla U
  tieto vstupy sa prenasobia váhami a spočítajú
  na túto sumu sa aplikuje nelineárna aktivačná
    funkcia  $\phi$ 
  tento výstup sa pošle dalej

```

Algoritmus 14.5: Postup prenosu informácie medzi uzlami v neurónovej sieti.



Obrázok 14.19: Rôzne typy neurónových sietí. (a) Dopredná sieť. (b) Spätná sieť (Hopfieldova).

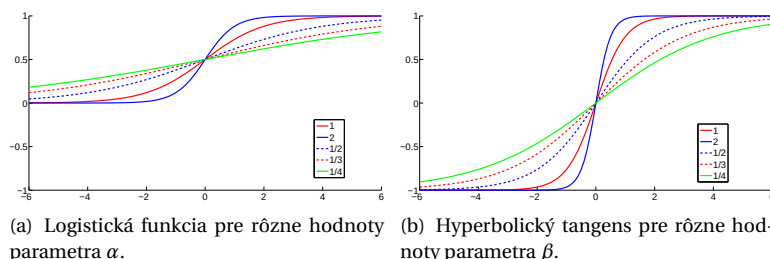
Existuje viacero druhov neurónových sietí. Ak informácia v sieti postupuje len jedným smerom od vstupov k výstupom, sieť tvorí orientovaný acyklický graf. Túto sieť nazývame dopredná (obr. 14.19(a)). V tejto kapitole sa budeme hlbšie zaoberať len základným typom doprednej siete – viacvrstvovým perceptrónom.

Každá sieť, ktorá netvorí acyklický graf, sa nazýva spätná. Príkladom takejto siete je Hopfieldova sieť (obr. 14.19(b)), ktorá tvorí kompletný graf a váhy sú symetrické $w(i \rightarrow j) = w(j \rightarrow i)$.

Medzi neurónové siete patria aj neriadené metódy. Z nich sa budeme v kapitole 15.3 zaoberať Kohonenovou sieťou.

14.6.1 Aktivačné funkcie

Existuje niekoľko často používaných aktivačných funkcií. Môže to byť napríklad prahová funkcia, ktorá mapuje kladné vstupy na hodnotu jedna

**Obrázok 14.20:** Používané aktivačné S-funkcie.

a záporné na hodnotu nula alebo mínus jedna

$$\phi(x) = \begin{cases} 1 & \text{ak } x \geq 0, \\ 0 & \text{inak.} \end{cases} \quad (14.80)$$

Táto funkcia však nie je diferencovateľná, čo potom robí problém pri tréningu siete (viac v kapitole 14.6.3).

Lepšou voľbou aktivačnej funkcie je logistická funkcia (pre $\alpha = 1$ nazývaná sigmoida)

$$\phi(\alpha, x) = \frac{1}{1 + e^{-\alpha x}}, \quad (14.81)$$

alebo hyperbolický tangens

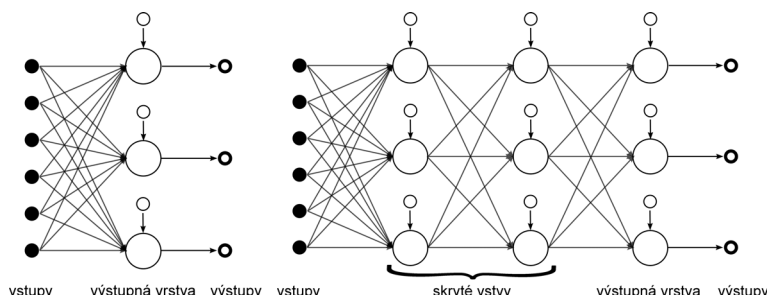
$$\phi(\beta, x) = \tanh(\beta x) = \frac{e^{\beta x} - e^{-\beta x}}{e^{\beta x} + e^{-\beta x}}. \quad (14.82)$$

Obe tieto funkcie majú graf pripomínajúci tvar S, budeme ich preto nazývať S-funkcie. Presný tvar závisí od parametrov α a β (obr. 14.20).

V problémoch rozpoznávania sa väčšinou v skrytých vrstvách používajú S-funkcie a vo výstupnej vrstve lineárna aktivačná funkcia. Nelineárna funkcia v skrytých vrstvách umožňuje sieti zachytiť nelineárne vzťahy medzi vstupom a výstupom siete.

14.6.2 Viacvrstvový perceptrón

Základným typom neurónovej siete je dopredná sieť nazývaná aj viacvrstvový perceptrón (multilayer perceptron – MLP). Táto neurónová sieť má niekoľko vrstiev – vstupnú, výstupnú a nula alebo viac skrytých vrstiev. Schematické znázornenie jednotlivých vrstiev neurónovej siete je na



Obrázok 14.21: Príklad jednovrstvovej a viacvrstvovej siete.

obrázku 14.21. Vstupnú vrstvu tvoria priamo dáta z vektorov príznačkov. Z výstupnej vrstvy vychádzajú údaje, na základe ktorých sú vstupné dáta klasifikované.

14.6.3 Učenie neurónovej siete

Podstatou učenia neurónovej siete je určenie váhovej matice (váhových vektorov v jednotlivých vrstvách). Hodnoty váhovej matice, ktoré minimalizujú zvolené optimalizačné kritérium, určujeme pomocou gradientovej metódy. Gradientová metóda hľadá minimum funkcie pomocou derivácie. Jednoduchý gradientný algoritmus pre neurónové siete je spätné šírenie (propagácia) chyby.

Učenie začína inicializáciou náhodnej váhovej matice. Vektory príznačkov z trénovacej množiny postupne prichádzajú na vstup a „prebublajú“ sieťou na koniec. Výstupný vektor sa potom porovná požadovanou klasifikáciou a určí sa hodnota kritéria E . Následne sa chyba spätne šíri a upravujú sa hodnoty váhovej matice.

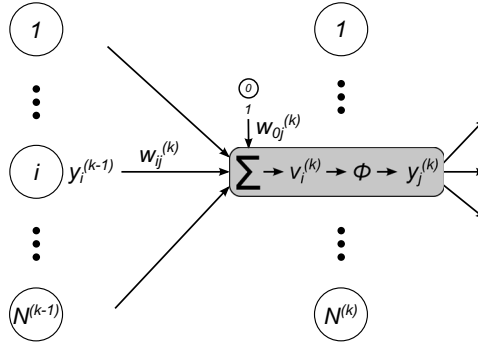
Na určenie jednotlivých váh použijeme gradientnú metódu (pozri aj vzťah (14.21))

$$\mathbf{W} \leftarrow \mathbf{W} + \Delta \mathbf{W}, \quad (14.83)$$

kde

$$\Delta \mathbf{W} = -\eta \nabla E|_{\mathbf{W}}. \quad (14.84)$$

Označme $w_{ij}^{(k)}$ váhu od i -teho neurónu v $(k-1)$ -vej vrstve do j -teho neurónu v k -tej vrstve (pozri obrázok 14.22).



Obrázok 14.22: Presun medzi vrstvami siete. Výstupy z vrstvy $k - 1$ prenášobené váhou vstupujú do neurónov vrstvy k , kde sa spočítajú a na tento súčet sa aplikuje aktivačná funkcia.

Potom pre konkrétnu váhu w_{ij}

$$\Delta w_{ij}^{(k)} = -\eta \frac{\partial E}{\partial w_{ij}^{(k)}}. \quad (14.85)$$

Ďalej označme $y_j^{(k)}$ výstup j -teho neurónu v k -tej vrstve. Potom vstup do tohto neurónu tvorí

$$v_j^{(k)} = \sum_{i=1}^{N^{(k-1)}} w_{ij}^{(k)} y_i^{(k-1)} + w_{0j}^{(k)}, \quad (14.86)$$

kde $N^{(k-1)}$ je počet neurónov v $(k-1)$ -vej vrstve a $w_{0j}^{(k)}$ je prah. Na tento vstup aplikujeme aktivačnú funkciu ϕ , a teda výstupom tohto neurónu je

$$y_j^{(k)} = \phi^{(k)} \left(v_j^{(k)} \right). \quad (14.87)$$

Vzťah (14.85) môžeme pomocou derivácie zloženej funkcie teraz napísať ako

$$\Delta w_{ij}^{(k)} = -\eta \frac{\partial E}{\partial v_j^{(k)}} \frac{\partial v_j^{(k)}}{\partial w_{ij}^{(k)}}. \quad (14.88)$$

Pre poslednú parciálnu deriváciu vo vzťahu (14.88) platí

$$\frac{\partial v_j^{(k)}}{\partial w_{ij}^{(k)}} = y_i^{(k-1)}. \quad (14.89)$$

Označme

$$\delta_j^{(k)} = -\frac{\partial E}{\partial v_j^{(k)}} = \frac{\partial E}{\partial y_j^{(k)}} \frac{\partial y_j^{(k)}}{\partial v_j^{(k)}}, \quad (14.90)$$

potom

$$\Delta w_{ij}^{(k)} = \eta \delta_j^{(k)} y_i^{(k-1)}. \quad (14.91)$$

Problémom zostáva vyčíslit' $\delta_j^{(k)}$. Rozoberieme si situáciu pre viac-vrstvové siete. V takomto prípade musíme rozlišovať medzi šírením chyby vo výstupnej vrstve a v skrytých vrstvách.

Šírenie chyby vo výstupnej vrstve

Vyčíslime jednotlivé parciálne derivácie vo vzťahu (14.90). Pre výstupnú vrstvu $k = m$ platí, že

$$\delta_j^{(m)} = -\frac{\partial E}{\partial v_j^{(m)}} = \frac{\partial E}{\partial y_j^{(m)}} \frac{\partial y_j^{(m)}}{\partial v_j^{(m)}}. \quad (14.92)$$

Potom

$$\frac{\partial y_j^{(m)}}{\partial v_j^{(m)}} = \phi'(v_j^{(m)}). \quad (14.93)$$

Vo výstupnej vrstve je chyba neurónu daná chybou, akú jeho výstup nadobúda voči očakávanému výstupu. Často používaným optimalizačným kritériom je stredná kvadratická chyba (MSE)

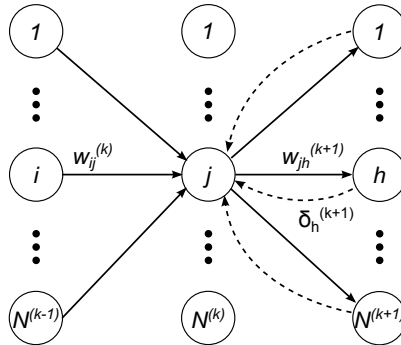
$$E(\mathbf{W}) = \frac{1}{2} \sum_{j=1}^M (y_j^{(m)} - o_j)^2, \quad (14.94)$$

kde pre daný vstup a váhy $\mathbf{W}$, o_j je očakávaný výstup neurónovej siete v uzle j a M je počet neurónov vo výstupnej vrstve. Potom

$$\frac{\partial E}{\partial y_j^{(m)}} = y_j^{(m)} - o_j, \quad (14.95)$$

čiže

$$\delta_j^{(m)} = (o_j - y_j^{(m)}) \phi'(v_j^{(m)}). \quad (14.96)$$



Obrázok 14.23: Šírenie chyby spätne medzi vrstvami.

Šírenie chyby v skrytých vrstvách

Pre k -tu vrstvu platí

$$\frac{\partial y_j^{(k)}}{\partial v_j^{(k)}} = \phi'(v_j^{(k)}). \quad (14.97)$$

Chyba j -teho neurónu v k -tej skrytej vrstve sa propaguje zo všetkých neurónov $(k+1)$ -vej vrstvy (obr. 14.23). Preto s využitím vzťahov (14.90) a (14.86)

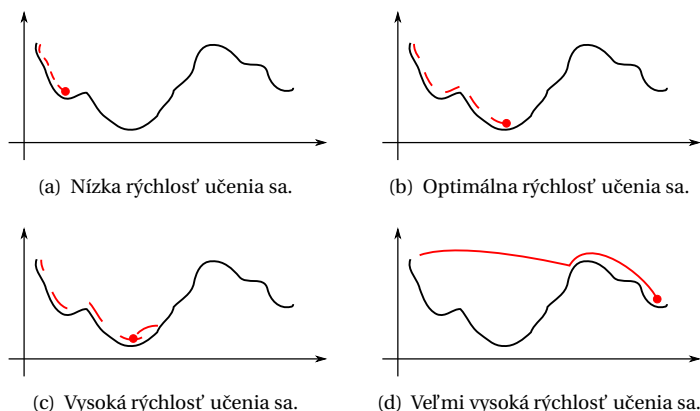
$$\begin{aligned} \frac{\partial E}{\partial y_j^{(k)}} &= \sum_{h=1}^{N^{(k+1)}} \frac{\partial E}{\partial v_h^{(k+1)}} \frac{\partial v_h^{(k+1)}}{\partial y_j^{(k)}} = \\ &= - \sum_{h=1}^{N^{(k+1)}} \delta_h^{(k+1)} w_{jh}^{(k+1)}. \end{aligned} \quad (14.98)$$

Z toho vyplýva, že

$$\delta_j^{(k)} = \phi'(v_j^{(k)}) \sum_{h=1}^{N^{(k+1)}} \delta_h^{(k+1)} w_{jh}^{(k+1)}. \quad (14.99)$$

14.6.4 Rýchlosť učenia sa neurónovej siete

Na rýchlosti učenia sa závisí, či sieť pri tréňovaní dosiahne globálne minimum chybovej funkcie. Obrázok 14.24 ilustruje rôzne prípady voľby rýchlosti učenia sa. Chybová funkcia je znázornená čiernou a krok učenia sa červenou farbou. Veľmi nízka rýchlosť učenia sa (obr. 14.24(a))



Obrázok 14.24: Rýchlosť učenia sa neurónovej siete. Chybová funkcia je znázorená čiernou a krok (rýchlosť) učenia sa červenou farbou. Pri nízkej rýchlosti môže riešenie uviaznuť v lokálnom minime. Pri optimálnej rýchlosti dosiahneme globálne minimum. Pri vysokej rýchlosti tiež, ale riešenie môže predtým oscilovať. Pri veľmi vysokej rýchlosti učenia sa môže riešenie divergovať.

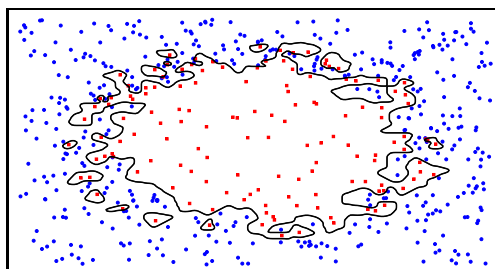
znamená pomalý tréning, ktorý môže skončiť v lokálnom minime, z ktorého sa už nedostane. Pri optimálnej rýchlosti (obr. 14.24(b)) je tréning síce tiež pomalý, ale dosiahne globálne optimum. Vysoká rýchlosť učenia sa (obr. 14.24(c) a 14.24(d)) zabezpečí rýchlejší tréning, ale môže proces hľadania môže oscilovať a pri veľmi vysokých rýchlostiach až divergovať.

Veľmi často sa používa premenlivá rýchlosť učenia sa daná heuristikou, že váhy ANN majú byť malé. Na začiatku procesu nastavíme rýchlosť učenia sa na určitú hodnotu a vypočítame nové váhy. Určíme chybu, akú sieť dosiahne pri týchto nových váhach.

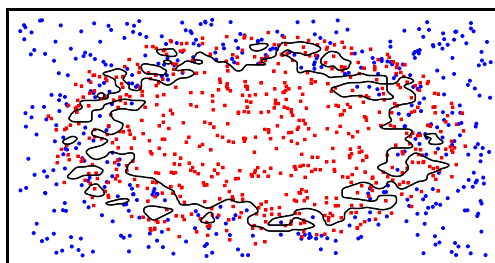
Ak nová chyba prevýši starú o stanovený prah, ideme zlým smerom. Preto sa nové váhy zahodia a rýchlosť učenia sa sa zníži.

Malé zvýšenie chyby povolíme kvôli možnosti dostať sa z lokálneho minima. V prípade, že nová chyba je vyššia ako stará o menej než stanovený prah, váhy sa upravia.

Ak by nová chyba bola nižšia než stará, tak smerujeme k minimu, môžeme upraviť váhy a zvýšiť rýchlosť učenia sa.



(a) Trénovacie dáta.



(b) Testovacie dáta.

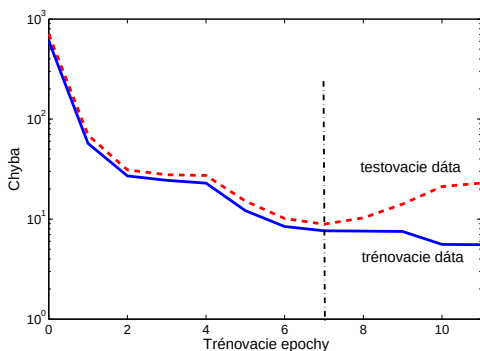
Obrázok 14.25: Ukážka preučeného klasifikátora. Chyba dosiahnutá na trénovacích dátach je síce nízka, ale chyba na testovacích dátach je veľmi vysoká.

14.6.5 Ukončenie učenia sa neurónovej siete

Jednou z možností kedy ukončiť tréning je nastaviť maximálny počet iterácií. Tento prístup neberie vôbec do úvahy, ako sa vyvíja dosiahnutá chyba siete, a preto môže skončiť ďaleko od optimálneho výsledku.

Zdalo by sa, že ideálnym kritériom zastavenia je, keď dosiahneme globálne minimum chybovej funkcie. Problémom je veľmi dlhý čas tréningu, pričom dosiahnutie globálneho optima nie je zaručené. Keď z požiadavky globálneho minima zľavíme, zastavíme učenie, keď chyba klesne pod stanovený prah. Čo to ale znamená z hľadiska kvality klasifikátora?

Kvalitu klasifikátora hodnotíme podľa toho, ako dobre klasifikuje neznáme dáta. Nízka chyba dosiahnutá pri tréningu nezaručuje správnu klasifikáciu neznámych dát. Klasifikátor sa totiž môže dostať do stavu preučenia. Prípadne, keď trénovacia množina nezodpovedala rozdeleniu reálnych dát, niektoré vzťahy medzi dátami sa klasifikátor nemusel „naučiť“.



Obrázok 14.26: Chyba pri učení ANN dosiahnutá na trénovacích a testovacích dátach. Vidíme, že koniec tréningu by mal nastať v epoche 7, keď chyba na testovacích dátach začala narastať.

Problém preučenia ilustrujeme na obrázku 14.25. Vľavo vidíme množinu trénovacích dát, a výsledný klasifikátor, ktorý sa snažil minimalizovať dosiahnutú chybu. Na pravom obrázku sú rozhodovacie hranice preučeného klasifikátora aplikované na testovacie dáta. Vidíme, že chyba, ktorú tento klasifikátor na týchto dátach dosiahne je veľmi veľká.

Riešením tohto problému je rozdeliť množinu dát na tréningovú a testovaciu a počas tréningu vyhodnocovať chybu na oboch množinách. Typický priebeh týchto chýb je znázornený na obrázku 14.26. Na začiatku tréningu chyba oboch množín klesá až do okamihu, keď klasifikátor vstúpi do stavu preučenia. Tréning ukončíme, keď sa chyba v testovacej množine začne zvyšovať.

Spôsobom, ako rozdeliť množinu na tréningovú a testovaciu časť, sa venujeme v kapitole 16.

Neriadené klasifikátory

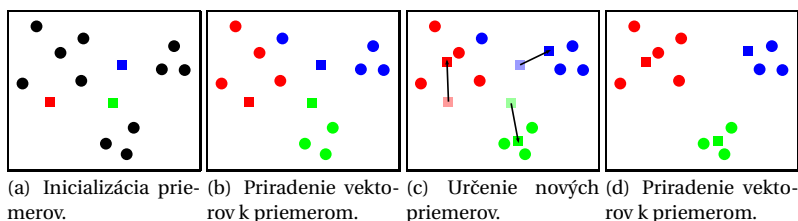
Neriadené metódy klasifikácie používame, keď k vstupným dátam neexistuje informácia o tom, do ktorej klasifikačnej triedy patria. Tieto metódy sa snažia odhaliť vnútornú štruktúru dát a rozdeliť dáta do zhlukov zodpovedajúcich jednotlivým klasifikačným triedam. V oblasti počítačového videnia sa používajú v klasifikácii objektov pri hľadaní reprezentatívnych vektorov tried, pri kvatovaní farieb na určenie palety alebo pri segmentácii obrazu na oddelenie objektov. Zhlučovací metódy sa dajú podľa typu výsledného rozdelenia priestoru rozdeliť nasledovne.

Nehierarchické metódy sú najpoužívanějšími metódami, ktoré vytvárajú rozdelenie priestoru na konečný počet zhlukov. Zhluky sa tvoria na základe určených metrík (vzdialenosť, podobnosť, ...) a postupne sa iteratívne zlepšujú, kým nie je dosiahnutá stabilita zhlukov. Príkladom je metóda K priemerov (kap. 15.1).

Hierarchické metódy vytvoria hierarchickú stromovú štruktúru rozdelení priestoru. Používajú sa prístupy zhora nadol (rozdeľujúci prístup – zhluky sa postupne delia) a zdola nahor (aglomeratívny prístup – zhluky sa postupne spájajú). Hierarchické metódy sú bližšie opísané v kapitole 15.2.

15.1 Metóda K priemerov

Najpoužívanjšou a najjednoduchšou metódou hľadania zhlukov v dátach patrí metóda K priemerov. Jej základnou myšlienkou je určenie zclu-



Obrázok 15.1: Prvé 4 kroky algoritmu K priemerov (alg. 15.1).

kov a ich priemerov. Počet zhlukov K je daný vopred. Je to iteratívna metóda, kde zhluky a ich priemery sa postupne upravujú (algoritmus 15.1).

Inicializácia priemerov
 Iterovanie, kým nenastane konvergencia

1. Priradenie vektorov k jednotlivým priemerom
2. Určenie nových priemerov

Algoritmus 15.1: Základný algoritmus metódy K priemerov.

Prvým krokom je inicializácia priemerov $\mu_1^{(0)}, \mu_2^{(0)}, \dots, \mu_K^{(0)}$. V klasickom algoritme sa zoberie prvých K vektorov $\mathbf{x}_i$, alebo sa priemery určia náhodne (rovnomerne z oboru hodnôt dát).

V iteratívnej časti algoritmu striedame dva kroky. Najprv sa priradia vektory do zhlukov určených priermi. Vektor $\mathbf{x}_i$ priradíme do zhliku $C_k^{(t)}$ s priemerom $\mu_k^{(t)}$, ak pre všetky $j \in \{1, \dots, K\}$

$$\|\mathbf{x}_i - \mu_k^{(t)}\| \leq \|\mathbf{x}_i - \mu_j^{(t)}\|. \quad (15.1)$$

Po priradení všetkých vektorov vypočítame nové priemery

$$\mu_k^{(t+1)} = \frac{1}{|C_k^{(t)}|} \sum_{\mathbf{x} \in C_k^{(t)}} \mathbf{x}. \quad (15.2)$$

Ilustrácia inicializácie a prvých troch krokov algoritmu je na obrázku 15.1.

Algoritmus K priemerov konverguje k lokálnemu optimu, globálne optimum nie je zaručené. Účelová funkcia je

$$O = \sum_{k=1}^K \sum_{\mathbf{x} \in C_k} \|\mathbf{x} - \mu_k\|. \quad (15.3)$$

Na dosiahnutie globálneho minima účelovej funkcie návod neexistuje. Klasifikácia závisí na počte zhlukov K a na inicializácii priemerov. Na určenie „správneho“ počtu zhlukov môžeme použiť vzájomnú validáciu (kap. 16). Na určenie počiatkových priemerov existujú rôzne heuristiky, ktoré využívajú znalosti o dátach (napríklad rozdelenie pravdepodobnosti).

15.1.1 Metóda C priemerov

Vylepšením metódy K priemerov je použitie neurčitých (fuzzy) zhlukov. Základný algoritmus je rovnaký ako pri metóde K priemerov (alg. 15.1), avšak vektor $\mathbf{x}_i$ nie je striktnie priradený zhľuku C_k , ale pomocou funkcie členstva $R_k(\mathbf{x}_i) \in \langle 0, 1 \rangle$ je určená pravdepodobnosť priradenia vektora $\mathbf{x}_i$ do zhľuku C_k . Platí, že

$$\sum_{k=1}^K R_k(\mathbf{x}_i) = 1. \quad (15.4)$$

Účelová funkcia v tomto prípade je

$$O = \sum_{k=1}^K \sum_{\mathbf{x} \in C_k} R_k(\mathbf{x}) \|\mathbf{x} - \boldsymbol{\mu}_k\|. \quad (15.5)$$

Pri iterovaní algoritmu sa vektorom zmení funkcia členstva

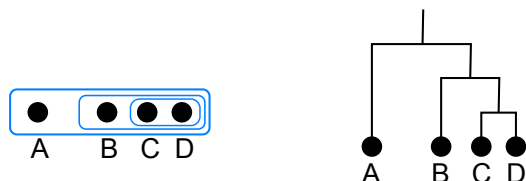
$$R_k^{(t+1)}(\mathbf{x}_i) = \left(\sum_{j=1}^C \left(\frac{\|\mathbf{x}_i - \boldsymbol{\mu}_k^{(t)}\|}{\|\mathbf{x}_i - \boldsymbol{\mu}_j^{(t)}\|} \right)^{\frac{2}{m-1}} \right)^{-1}, \quad (15.6)$$

kde $m \in \langle 1, \infty \rangle$ je parameter určujúci váhu, akou vektory prispievajú k určeniu členstva a priemeru. V literatúre sa uvádza $\langle 1, 1.5 \rangle$ ako vhodný interval na výber m . V prípade, že vektor $\mathbf{x}_i$ kolide s priemerom $\boldsymbol{\mu}_j$, tak

$$R_k^{(t+1)}(\mathbf{x}_i) = \begin{cases} 1 & \text{ak } k = j, \\ 0 & \text{inak.} \end{cases} \quad (15.7)$$

Následne sa vypočítajú nové priemery

$$\boldsymbol{\mu}_k^{(t+1)} = \frac{\sum_{i=1}^N \left(R_k^{(t+1)}(\mathbf{x}_i) \right)^m \mathbf{x}_i}{\left(\sum_{i=1}^N R_k^{(t+1)}(\mathbf{x}_i) \right)^m}. \quad (15.8)$$



Obrázok 15.2: Postupnosť vytvárania zhľukov vyjadrená dendrogramom.

15.2 Hierarchické zhľukovanie

Hierarchické zhľukovacie metódy sa delia podľa prístupu na

Aglomeratívne: Pri tomto type zhľukovania začíname s počtom zhľukov rovným počtu dát. Každý vektor príznakov tvorí zhľuk. Postupne spájame najbližšie zhľuky a tým vytvárame stromovú štruktúru zvanú dendrogram (obr. 15.2). Na konci postupu je jediný zhľuk, obsahujúci všetky vektory príznakov. V ďalšom texte opíšeme metódy spájania zhľukov.

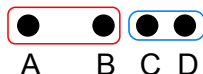
Rozdeľujúce: Tento prístup je opačný – začíname s jediným zhľukom a postupne ho delíme. Nevýhodou tohto prístupu je, že hľadanie optimálneho rozdelenia množiny na dve podmnožiny má vysokú zložitosť (treba preskúmať $2^{n-1} - 1$ možností, kde n je mohutnosť množiny). Preto sa tieto metódy využívajú len pre malé množiny dát.

Nevýhodou oboch prístupov je, že v každom kroku hľadajú lokálne optimálne rozdelenie. V ďalších krokoch sa zhľuky nedajú preskúpiť, aj keby to rozdelenie v tomto kroku zlepšilo.

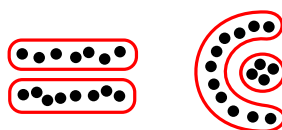
Pri aglomeratívnom prístupe musíme v každom kroku určiť dva najpodobnejšie zhľuky. Po každom vytvorení nového zhľuku musíme prepočítať podobnosť s ostatnými zhľukmi, ktorých vzájomné podobnosti zostávajú nezmenené. Podobnosť sa určuje vzdialenosťou medzi zhľukmi. Vzdialenosť môže byť definovaná rôzne. Nech $S(C_i, C_j)$ je podobnosť zhľukov a $d(\cdot)$ zvolená metrika (Euklidovská, Hammingova, Manhattan, ...). Potom najpoužívanejšie metódy na určenie podobnosti sú

Centroidná: Podobnosť je definovaná vzdialenosťou priemerov zhľukov

$$S(C_i, C_j) = d(\mu_i, \mu_j). \quad (15.9)$$



Obrázok 15.3: Pri metóde maximálnej vzdialenosti bude pre body z obrázka 15.2 v druhom kroku vytvorený „nepravý“ zhluk. Vidíme, že body B, C a D tvoria prirodzený zhluk, avšak metóda maximálnej vzdialenosti vytvorí zhluk naznačený červenou farbou.



Obrázok 15.4: Metóda minimálnej vzdialenosti identifikuje zhluky, ktoré kompletne linkovanie ani cetroidná metóda neodhalia.

Maximálna: Spočítajú sa vzájomné vzdialenosti všetkých prvkov zhlu-
kov a vyberie sa najväčšia

$$S(C_i, C_j) = \max_{\mathbf{x} \in C_i, \mathbf{y} \in C_j} d(\mathbf{x}, \mathbf{y}). \quad (15.10)$$

Nazýva sa aj metóda najvzdialenějšího suseda (úplné linkovanie). Vytvára skoro sférické zhluky približne rovnakej veľkosti. Zabráňuje vzniku zret'azených zhlukov. Výskyt odľahlých hodnôt môže viesť k vzniku „nepravých“ zhlukov (obr. 15.3).

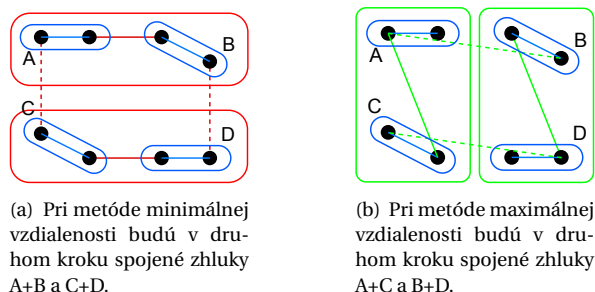
Minimálna: Podobná metóda ako predchádzajúca, ale vyberie sa naj-
menšia vzdialenosť

$$S(C_i, C_j) = \min_{\mathbf{x} \in C_i, \mathbf{y} \in C_j} d(\mathbf{x}, \mathbf{y}). \quad (15.11)$$

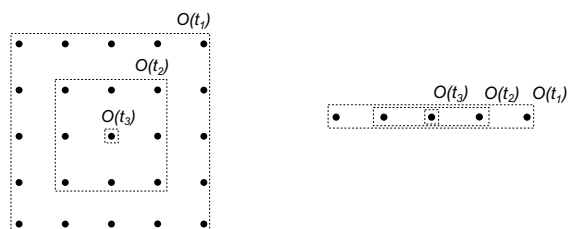
Je to metóda najbližšieho suseda (jednoduché linkovanie). Nevý-
hodou je, že sa môžu zret'aziť zhluky, ktoré sú na vzdialenejších
koncoch výrazne nepodobné. Na druhej strane dokáže identifikovať
zhluky, ktoré kompletne linkovanie ani cetroidná metóda ne-
odhalia (obr. 15.4).

Priemerná: V tomto prípade sa spočíta nevážený priemer vzdialeností

$$S(C_i, C_j) = \frac{1}{|C_i||C_j|} \sum_{\mathbf{x} \in C_i, \mathbf{y} \in C_j} d(\mathbf{x}, \mathbf{y}). \quad (15.12)$$



Obrázok 15.5: Porovnanie zhlukov pri metóde minimálnej a maximálnej vzdialenosti.



Obrázok 15.6: Usporiadanie neurónov SOM do mriežky alebo lineárne.

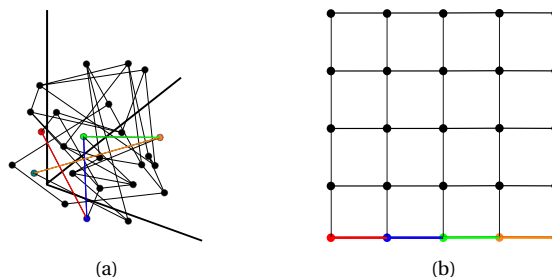
Táto metóda nie je príliš citlivá na odchýlky. Odľahlé hodnoty nemajú výrazný vplyv na výsledné rozdelenie priestoru. Existuje aj metóda využívajúca vážený priemer, kde sa veľkosť (mohutnosť) zhľuku berie ako váha.

Použitie rôznych metód vedie k rôznym výsledným rozdeleniam. Na obrázku 15.5 je porovnanie rozdelení pre minimálnu a maximálnu metódu.

15.3 Samoorganizujúce sa mapy

Samoorganizujúce sa mapy (SOM, alebo Kohonenove siete) sú typom neriadených neurálnych sietí, v ktorých sú neuróny v určitom topologickom vzťahu. Pri ANN (kap. 14.6) sú neuróny vo vrstvách nezávislé a váhy sa upravujú bez ohľadu na susedné neuróny.

SOM sú jednovrstvové siete, kde sú neuróny umiestnené v mriežke (väčšinou 1D alebo 2D, pozri obr. 15.6). Každý neurón teda obsahuje in-



Obrázok 15.7: Dva priestory SOM. V priestore váh (a) môžu byť na začiatku trénovania siete váhy topologicky blízkyh susedov ďaleko od seba. Farebné spojnice ilustrujú polohu uzlov jedného riadka mriežky SOM v priestore váh (b).

formáciu o svojom topologickom okolí. Vstupné dáta sú porovnávané s váhovým vektorom každého neurónu a sieť je trénovaná algoritmom „víťaz berie všetko“.

Inicializácia

Učenie siete:

1. Vstup vektora príznakov
2. Určenie víťazného neurónu
3. Víťaz berie všetko

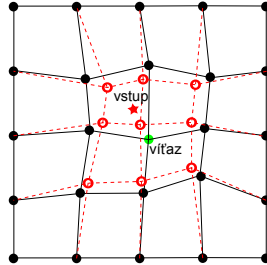
Algoritmus 15.2: Základný algoritmus učenia sa SOM.

Základný postup pri učení SOM je opísaný v algoritme 15.2. Prvým krokom je inicializácia váhových vektorov. Tieto inicializujeme náhodnými hodnotami z daného intervalu (podľa riešeného problému). Váhové vektory majú rovnaký rozmer ako vektory príznakov. Pri SOM musíme rozlišovať medzi topologickým vzťahom dvoch neurónov v mriežke a vzťahom ich váhových vektorov v Euklidovskom priestore (obr. 15.7), a to hlavne pri 2D vektore príznakov.

Vo fáze učenia sa siete prichádzajú na vstup jednotlivé vektory príznakov $\mathbf{x}$ a určí sa víťazný neurón i^* , ktorého váhový vektor je v zmysle Euklidovskej vzdialenosti najbližšie $\mathbf{x}$

$$i^* = \underset{i}{\operatorname{argmin}} \|\mathbf{x} - \mathbf{w}_i\|. \quad (15.13)$$

Ďalším krokom je úprava váhového vektora víťazného neurónu, ako aj neurónov v jeho topologickom okolí (obr. 15.8). Tieto váhové vektory



Obrázok 15.8: Úprava váhového vektora víťazného neurónu a neurónov v jeho topologickom okolí.

sa posunú smerom k $\mathbf{x}$ podľa vzťahu

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + \alpha(t)O(i, i^*, t)\|\mathbf{x} - \mathbf{w}_i(t)\|, \quad (15.14)$$

kde $O(i, i^*, t)$ je funkcia určujúca okolie víťazného neurónu a $\alpha(t) \in (0, 1)$ je rýchlosť učenia. Funkcia okolia aj rýchlosť učenia sa s časom zmenšujú, čím zaručia ukončenie učenia. Na začiatku procesu sa rýchlo mení dostatočne veľké okolie víťazného neurónu. Tu sa náhodne inicializované váhové vektory dostanú do približne výsledného usporiadania a potom s postupne zmenšujúcim sa okolím a pomalšou zmenou vektorov sa dolad'ujú presné hodnoty. V učení teda rozpoznávame fázu usporiadavania a fázu dolad'ovania.

Okolie môže byť pravouhlé (ak sú neuróny usporiadané v pravouhlej mriežke), definované funkciou

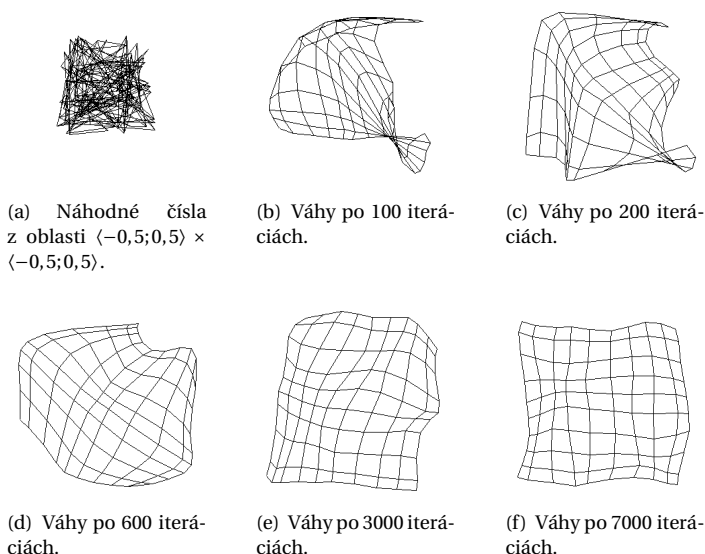
$$O(i, i^*, t) = \begin{cases} 1 & \text{ak } \|\mathbf{w}_{i^*}(t) - \mathbf{w}_i(t)\| \leq \sigma(t), \\ 0 & \text{inak} \end{cases} \quad (15.15)$$

alebo gaussovské, definované ako

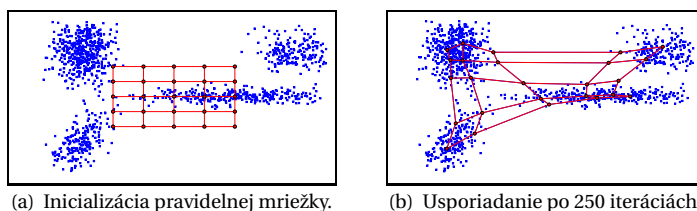
$$O(i, i^*, t) = \exp\left(-\frac{\|\mathbf{w}_{i^*}(t) - \mathbf{w}_i(t)\|^2}{\sigma(t)}\right), \quad (15.16)$$

pričom hodnota parametra $\sigma(t)$ klesá so zvyšujúcim sa t , a tým sa okolie v čase zmenšuje. Pri pravouhlom okolí sa všetky body okolia menia s rovnakou váhou, pri gaussovskom sú body na okraji okolia menej významne posunuté.

Príklad tréningu siete je na obrázku 15.9. Mriežka je štvorcová s rozmermi 10×10 . Vstupnými vektormi sú rovnomerne rozdelené náhodné čísla z oblasti $\langle -1, 1 \rangle \times \langle -1, 1 \rangle$. V prvej etape sú váhy náhodne inicializované na hodnoty v oblasti $\langle -0, 5; 0, 5 \rangle \times \langle -0, 5; 0, 5 \rangle$. V ďalších etapách sa



Obrázok 15.9: Tréning SOM na rovnomerne rozdelených dátach. (a) Inicializácia. (b), (c) a (d) Fáza usporiadavania. (e) a (f) Fáza dolaďovania.



Obrázok 15.10: SOM tréning pre nerovnomerne rozdelené dáta. Body mriežky (váhové vektory) sa presúvajú smerom k zhlukom dát.

váhové vektory usporiadavajú (etapy 200 – 600), potom nastáva fáza dolaďovania. Vidíme, že váhové vektory sa upravia tak, aby odhalili štruktúru tréningových dát.

V prípade nerovnomerne rozdelených vstupných dát sa váhové vektory presunú smerom do jednotlivých zhlukov dát, ako vidno na obrázku 15.10.



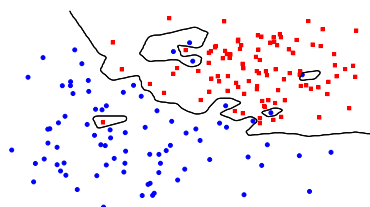
Hodnotenie kvality klasifikácie

Aby sme vedeli určiť kvalitu klasifikácie, musíme definovať a vedieť, ako sa klasifikátor správa pri dátach, s ktorými sa doteraz nestretol. Chybou je keď má klasifikátor 100% účinnosť na dátach, na ktorých bol natrénovaný a na neznámych dátach zlyhá. Takémuto stavu sa hovorí preučenie. Klasifikátor sa naučil „naspamäť“ určitú množinu dát. Preto pri určovaní kvality klasifikátora používame dve množiny dát – *trénovacie* a *testovacie* dáta. Trénovacie dáta sa používajú na učenie (trénovanie) klasifikátora. Testovacie dáta na určenie chybovosti (error rate) už natrénovaného klasifikátora.

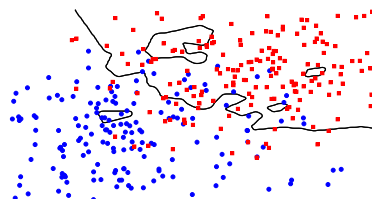
Aby sme mohli určiť chybu, musíme aj pri testovacích dátach poznať triedu, do ktorej dáta patria. Preto trénovaciu aj testovaciu množinu dát vyrobíme z pôvodnej množiny dát (s určenou klasifikačnou triedou).

Problém môže nastať, ak máme napríklad málo dát, alebo sa pri rozdelení dostanú do trénovacej množiny dáta jedného typu a do testovacej dáta druhého. Preto existujú metódy ako vybrať vzorky do testovacej množiny. Medzi ne patria napríklad metódy vzájomnej validácie alebo bootstrap.

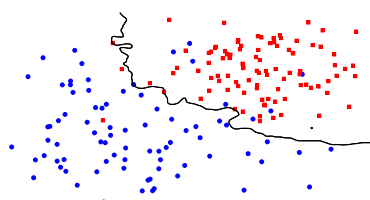
Okrem odhadu chyby klasifikátora sa metódy vzájomnej validácie používajú aj na určenie optimálnych parametrov klasifikátorov. Pri použití jednotlivých možných hodnôt skúmaného parametra odhadneme chybu výsledného parametra metódou vzájomnej validácie a následne vyberieme tú hodnotu parametra, ktorá dáva najmenšiu odhadnutú chybu. Príkladom môže byť určenie vstupnej hodnoty K v metóde K najbližších susedov (obr. 16.1), alebo v metóde K priemerov.



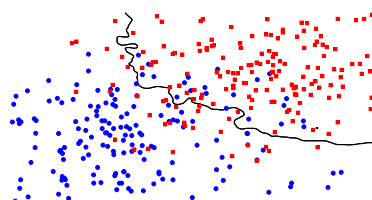
(a) $K = 1$, tréning, dosiahnutá chyba 0



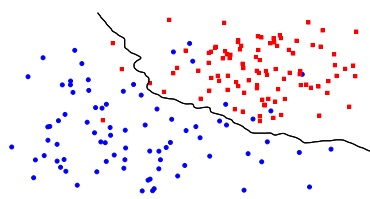
(b) $K = 1$, testovanie, dosiahnutá chyba 0,1667



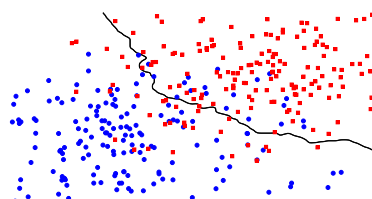
(c) $K = 7$, tréning, dosiahnutá chyba 0,0565



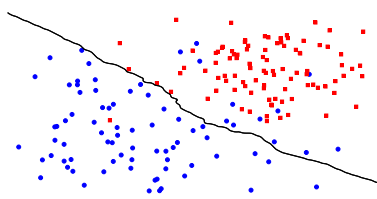
(d) $K = 7$, testovanie, dosiahnutá chyba 0,1637



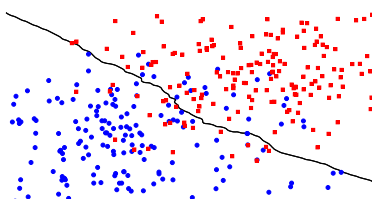
(e) $K = 15$, tréning, dosiahnutá chyba 0,0565



(f) $K = 15$, testovanie, dosiahnutá chyba 0,1488

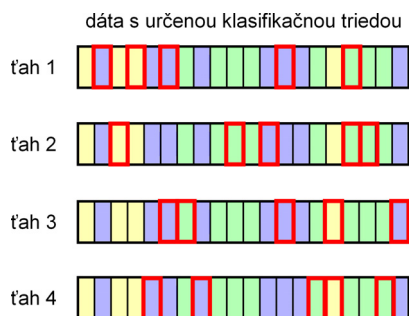


(g) $K = 51$, tréning, dosiahnutá chyba 0,0734



(h) $K = 51$, testovanie, dosiahnutá chyba 0,1339

Obrázok 16.1: Dosiahnutá chyba pri trénovaní a testovaní metódy K najbližších susedov.



Obrázok 16.2: Náhodné prevzorkovanie pre $K = 4$. Červeno označené vzorky boli vybrané do testovacej množiny.

16.1 Vzájomná validácia

Pri vzájomnej validácii sa rozdelenie pôvodnej množiny na tréningovú a testovaciu vykoná K -krát za sebou. Pre každé rozdelenie sa zistí dosiahnutá chyba E_i . Odhadnutá chyba E klasifikátora je potom priemerom jednotlivých zistených chýb

$$E = \frac{1}{K} \sum_{i=1}^K E_i. \quad (16.1)$$

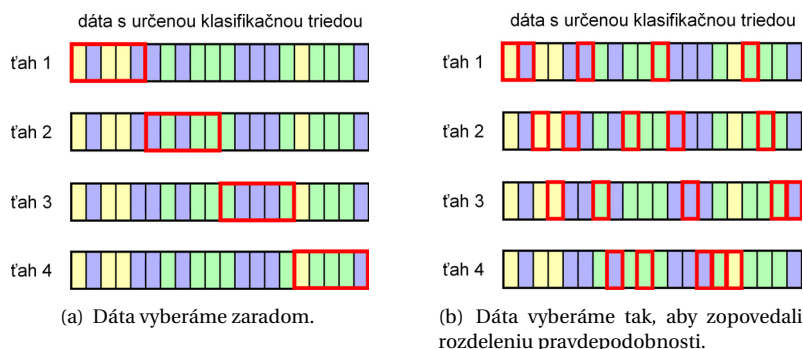
Medzi metódy vzájomnej validácie patria náhodné prevzorkovanie, K -násobná vzájomná validácia, vrstvená K -násobná vzájomná validácia a vzájomná validácia s vynechaním jednej vzorky.

16.1.1 Náhodné prevzorkovanie

Ako názov metódy napovedá, dáta sa z pôvodnej množiny do testovacej vyberajú náhodne, bez vrátenia naspäť. Pri každom z K výberov (ťahov) sa vytvorí testovacia množina M vzoriek, ktorá v jednotlivých výberoch môže obsahovať aj rovnaké vzorky a niektoré vzorky sa do testovacej množiny vôbec nedostanú. Táto situácia je znázornená na obrázku (16.2).

16.1.2 K -násobná vzájomná validácia

Aby sme zabránili opakovaniu sa vzoriek v jednotlivých testovacích množinách, rozdelíme pôvodnú množinu na K rovnako veľkých podmnožín. Každá z týchto podmnožín je raz použitá ako testovacia množina. V tomto



Obrázok 16.3: Obyčajná a vrstvená 4-násobná vzájomná validácia. Červeno označené vzorky boli vybrané do testovacej množiny.

pripade je každá vzorka použitá na tréovanie aj testovanie, ako vidíme na obrázku 16.3(a).

16.1.3 Vrstvená K-násobná vzájomná validácia

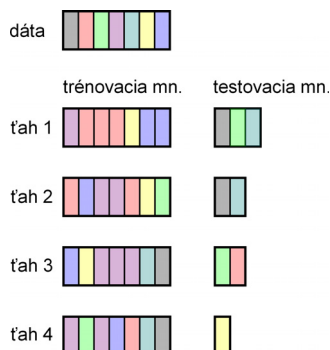
Aby sme zabezpečili, že tréovacie a testovacie množiny sú vyvážené, rozdelíme množinu dát X tak, aby podmnožiny zachovávali distribúciu jednotlivých tried. Podiel (pravdepodobnosť) vzoriek jednotlivých tried je rovnaký v pôvodnej množine X aj v každej z jej podmnožín $X_1, \dots, X_K$ tak, ako je znázornené na obrázku 16.3(b).

16.1.4 Vzájomná validácia s vynechaním jednej vzorky

Ak máme k dispozícii málo dát, snažíme sa tréovať klasifikátor na čo najväčšej množine, aby zachytil všetky vzťahy medzi dátami. V tomto prípade testujeme na jedinej vzorke. Táto situácia zodpovedá K-násobnej vzájomnej validácii, kde K je rovné počtu vzoriek.

16.2 Bootstrap

Aj pri metóde bootstrap delíme množinu dát K krát a počítame priemernú dosiahnutú chybu. Rozdiel je v tom, že pri tejto metóde sa v jednom výber môže v tréovacej množine jedna vzorka vyskytovať viackrát. Do tréovacej množiny vyberáme vzorky náhodne s opakovaním. V každom výbere je veľkosť tréovacej množiny rovná veľkosti pôvodnej množiny.



Obrázok 16.4: Možná realizácia výberu vzoriek do trénovacej a testovacej množiny metódou bootstrap. Vzorky, ktoré sa nedostanú do trénovacej množiny tvoria testovaciu množinu.

žiny. Vzorky, ktoré sa nedostanú do trénovacej množiny sa použijú na testovanie. Veľkosť testovacej množiny sa tak v jednotlivých výberoch mení. Príklad možnej realizácie metódy je na obrázku 16.4.

Pri tejto metóde sa pravdepodobnosť, že vzorka skončí v testovacej množine, rovná 0,368. Pozrime sa na jednu konkrétnu vzorku. Pravdepodobnosť, že ju nevyberieme do trénovacej množiny je $1 - 1/N$. Keďže výber opakujeme N -krát, pravdepodobnosť, že vzorku nevyberieme ani raz je

$$\left(1 - \frac{1}{N}\right)^N \approx e^{-1} = 0,368. \quad (16.2)$$

Klasifikátor teda trénujeme na 63,2 % dát. Preto sa tejto metóde hovorí aj 0,632 bootstrap.

16.3 Chybová funkcia klasifikátora

Chyby klasifikátora sa delia na dva typy. Povedzme, že sa v obraze snažíme detegovať konkrétny objekt. Ak sa tento objekt v obraze vyskytuje a náš klasifikátor ho deteguje, dosiahne správne pozitívny výsledok (true positive – TP). Ak by ho nedetegoval, dopustí sa chyby typu II a jeho výsledok je nesprávne negatívny (false negative – FN). V prípade, že objekt sa v obraze nevyskytuje a klasifikátor ho deteguje, nastáva chyba typu I a výsledok je nesprávne pozitívny (false positive – FP). V prípade, že klasifikátor určí, že objekt nie je prítomný, dostávame správne negatívny vý-

sledok (true negative – TN). Tabuľka 16.1 dáva názorný prehľad možných výsledkov klasifikátora.

Tabuľka 16.1: Typy a identifikácia možných výsledkov klasifikátora.

Skutočnosť objekt sa	Rozhodnutie – objekt sa	
	vyskytuje	nevyskytuje
vyskytuje	Správna klasifikácia Správne pozitívny výsl.	Chyba typu II Nesprávne negatívny výsl.
nevyskytuje	Chyba typu I Nesprávne pozitívny výsl.	Správna klasifikácia Správne negatívny výsl.

Niektoré chyby klasifikácie môžu byť závažnejšie (majú horšie následky) – napríklad neodhalenie smrteľnej choroby v medicíne, ktorá by sa dala ľahko liečiť liekmi, ktoré nemajú vedľajšie účinky. Teda neklasifikovanie choroby v prípade chorého človeka má horšie následky ako nasadenie liekov u zdravého človeka. Číselne môžeme „stratu“ v prípade správnej klasifikácie vyjadriť nulou a v prípade nesprávnej klasifikácie nejakou nenulovou hodnotou. Toto ohodnotenie sa nazýva funkcia straty

$$L(s, a) : S \times A \rightarrow \mathbb{R}_0^+, \quad (16.3)$$

kde S je priestor stavov (skutočných klasifikačných tried) a A je priestor akcií (výstupov klasifikátora). Najčastejšie používanou funkciou straty je 0-1 funkcia $L(s, a) = 1 - \delta_{sa}$, kde δ_{sa} je Kroneckerova delta (tab. 16.2(a)). V niektorých aplikáciách (detekcia spamu, medicína atď.) sa používa nesymetrická funkcia straty (tab. 16.2(b)).

Tabuľka 16.2: Ohodnotenie chýb klasifikátora.

Skutočnosť	Rozhodnutie	
	zdravý	chorý
zdravý	0	1
chorý	1000	0

(a) Rôzne hodnotené chyby.

Skutočnosť	Rozhodnutie	
	nie tvár	tvár
nie tvár	0	1
tvár	1	0

(b) Rovnako hodnotené chyby.

Pomocou funkcie straty $L(\cdot)$ môžeme vyjadriť očakávanú podmienenú stratu pri klasifikácii vzorky $\mathbf{x}$ patriacej triede ω nasledovne

$$R_\omega(\mathbf{x}) = \sum_{\psi \in A} L(\omega, \psi) P(\psi | \mathbf{x}). \quad (16.4)$$

Tabuľka 16.3: Príklad matice zámen pre klasifikáciu číslíc.

čísllica	klasifikácia										
	1	2	3	4	5	6	7	8	9	0	R
1	87	0	0	0	1	0	0	0	0	0	0
2	0	88	1	0	0	0	0	0	1	1	1
3	0	0	75	1	0	0	0	10	4	0	3
4	0	0	0	79	0	0	0	0	0	0	0
5	0	0	0	0	79	6	0	0	0	4	1
6	0	0	0	0	8	80	1	0	0	2	0
7	0	1	0	0	0	0	83	0	0	0	0
8	0	0	15	0	0	1	0	65	7	0	0
9	0	0	4	0	0	0	0	10	71	0	1
0	0	0	0	1	0	1	0	0	0	90	1

Celková očakávaná podmienená strata pri použití tohto klasifikátora je

$$R = \sum_{\omega \in S} \int_{\mathcal{R}_{\omega}} R_{\omega}(\mathbf{x}) d\mathbf{x} = \sum_{\omega \in S} \sum_{\psi \in A} \int_{\mathcal{R}_{\omega}} L(\omega, \psi) P(\psi | \mathbf{x}) d\mathbf{x}, \quad (16.5)$$

kde $\mathcal{R}_{\omega}$ je rozhodovacia oblasť triedy ω .

16.4 Metódy porovnávania klasifikátorov

Okrem vyčíslenia očakávanej straty môžeme klasifikátory porovnávať na základe iných vlastností. Jednou z nich je matica zámen (alebo kontingenčná tabuľka). Z matice zámen môžeme pre binárne klasifikátory (klasifikujú do dvoch tried) vypočítať niekoľko koeficientov kvality klasifikácie.

16.4.1 Matica zámen

Matica zámen je kontingnečnou tabuľkou, do ktorej zapisujeme počty vzoriek klasifikované do jednotlivých tried. Príkladom matice zámen pri klasifikácii číslíc je tabuľka 16.3. V tejto matici je špeciálny stĺpec označený R, ktorý dovoľuje klasifikátoru „nevyjadriť sa“ (reject option), ak hodnota klasifikačného kritéria nepresiahla stanovaný prah klasifikácie pre žiadnu triedu.

Z tejto matice môžeme vidieť, že klasifikátor nie je úplne dobre natreňovaný, lebo si zamieňa čísllice 3, 8 a 9 a tiež 5 a 6. Mohlo by sa zdať,

že klasifikátor rozpoznáva rovnako dobre číslice 4 a 5, lebo majú na diagonále rovnaký počet vzoriek. Problémom tejto matice však je, že počet vzoriek v jednotlivých triedach nie je konštantný. V takomto prípade sa niekedy do matice zámeny dáva miesto absolútnych počtov klasifikovaných vzoriek, ich percentuálny podiel v jednotlivých triedach. Mnohorozmerné matice zámen sa lepšie vizualizujú pomocou teplotných máp, ako vidno na obrázku 16.5.

Pre binárne klasifikátory, ktoré klasifikujú dáta do dvoch tried, zodpovedajú údaje v matici zámen jednotlivým typom klasifikácie TP , TN , FP , FN .

16.4.2 Koeficienty kvality klasifikácie

Z hodnôt TP , TN , FP , FN môžeme vypočítať nasledovné koeficienty kvality klasifikácie.

Citlivost' testu

Citlivosť klasifikátora je daná pomerom správne pozitívnych výsledkov k počtu hľadaných objektov skutočne sa vyskytujúcich v dátach

$$TPR = \frac{TP}{TP + FN}. \quad (16.6)$$

Táto hodnota je ovplyvňovaná hodnotou FN . Čím vyšší počet nesprávne negatívnych výsledkov, tým nižšia citlivosť klasifikátora. Znamená to, že klasifikátor nevie odhaliť všetky výskyty objektu. Ideálny klasifikátor nemá nesprávne negatívne výsledky a dosahuje 100% citlivosť.

Citlivosť sa nazýva aj senzitivita, miera skutočnej pozitivity (true positive rate – TPR) a v oblasti získavania informácií úplnosť alebo záchyt (recall).

Výlučnosť testu

Výlučnosť klasifikátora sa meria pomerom skutočne negatívnych výsledkov k počtu vzoriek neobsahujúcich hľadaný objekt

$$TNR = \frac{TN}{TN + FP}. \quad (16.7)$$

Je to miera ako dobre dokáže klasifikátor určiť negatívne vzorky. Čím viac nesprávne pozitívnych, tým nižšia hodnota TNR . Vysoká výlučnosť sa vyžaduje napríklad pri testovaní výrobkov na výrobnéj linke, kde sa výrobky, ktoré neprejdú testom, zahadzujú. Pri nízkej výlučnosti je medzi zahodenými veľké množstvo fungujúcich výrobkov. Ideálny klasifikátor má 100% výlučnosť.

Výlučnosť sa nazýva aj špecificita alebo miera skutočnej negativity (true negative rate – TNR).

S výlučnosťou klasifikátora súvisí aj miera nesprávnej pozitivity

$$FPR = \frac{FP}{FP + TN} = 1 - TNR. \quad (16.8)$$

Prediktívna hodnota

V medicíne sa používa prediktívna hodnota testu, čo je vlastne posteriórna pravdepodobnosť. Pozitívna prediktívna hodnota (PPH) testu je pravdepodobnosť, že pacient skutočne má danú chorobu, keď test vyšiel pozitívne. Čím menej je nesprávne pozitívnych výsledkov, tým viac sa hodnota PPH blíži k jednotke

$$PPH = \frac{TP}{TP + FP}. \quad (16.9)$$

PPH sa v oblasti získavania informácií nazýva presnosť (precision).

Negatívna prediktívna hodnota testu určuje pravdepodobnosť, že pacient je skutočne zdravý, ak test vyšiel negatívne. Čím menej nesprávne

negatívnych výsledkov, tým vyššia hodnota NPH

$$NPH = \frac{TN}{TN + FN}. \quad (16.10)$$

Prediktívna hodnota závisí od rozdelenia pravdepodobnosti choroby v skúmanej populácii (apriórna pravdepodobnosť).

Účinnosť

Počet správnych výsledkov klasifikácie (pozitívnych alebo negatívnych) v pomere k počtu všetkých skúmaných vzoriek nazývame účinnosť. Čím menej nesprávnych výsledkov, tým vyššia účinnosť

$$Eff = \frac{TP + TN}{TP + FP + TN + FN}. \quad (16.11)$$

Účinnosť voláme aj efektívnosť alebo správnosť (accuracy).

F₁ skóre

V oblasti získavania informácií sa na hodnotenie kvality dolovacích metód používa vážený harmonický priemer úplnosti a presnosti nazývaný F skóre

$$F_{\beta} = \frac{(1 + \beta^2)(PPH \cdot TPR)}{\beta^2 PPH + TPR}. \quad (16.12)$$

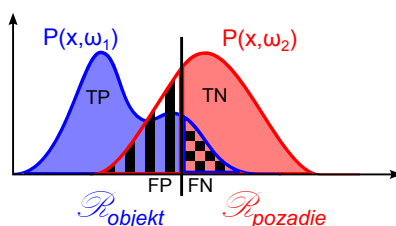
Najčastejšie sa používa F₁ skóre, kde váha presnosti a úplnosti je rovnaká ($\beta = 1$)

$$F_1 = \frac{2PPH \cdot TPR}{PPH + TPR}. \quad (16.13)$$

16.4.3 ROC krivky

Pôvodne z oblasti rádiológie sa do oblasti klasifikácie dostal pojem ROC krivka. ROC je z anglického Receiver Operating Characteristics – hodnota opisujúca prevádzkovú charakteristiku prijímača (radaru).

ROC krivka opisuje vzťah citlivosti a miery nesprávnej pozitivity, kde na x -ovej osi je FPR a na y -ovej TPR . ROC krivka je vhodná pri binárnych klasifikátoroch, ktoré klasifikačné kritérium vyhodnocujú voči nejakému prahu, prípadne obsahujú nejaký premenný parameter, ktorého optimálna hodnota sa dá určiť práve z analýzy ROC krivky.



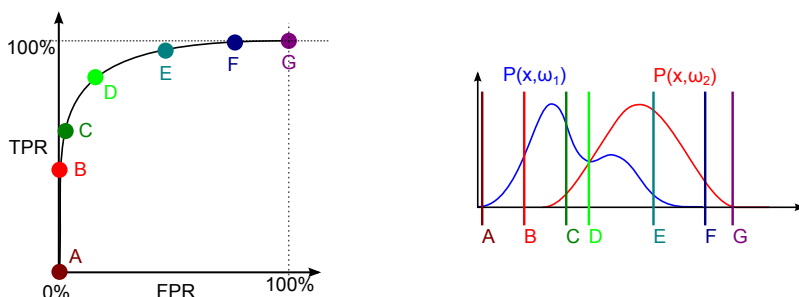
Obrázok 16.6: Rozdelenie pravdepodobnosti s vyznačeným klasifikačným prahom a určenými hodnotami TP , TN , FP , FN .

Vytvorenie ROC krivky

Ako príklad si uveďme klasifikátor, ktorý zaradí vzorku do klasifikačnej triedy na základe hodnoty jediného príznaku. Ak $x < h$, tak vzorka patrí do triedy ω_1 , inak do ω_2 . Parametrom tohto klasifikátora je hodnota prahu h . Nech trieda ω_1 reprezentuje objekt a trieda ω_2 pozadie obrazu. Potom obrázok 16.6 ilustruje hodnoty TP , TN , FP , FN pri danom rozdelení pravdepodobnosti.

Keď meníme prah h , menia sa hodnoty FPR a TPR . Pre každú hodnotu prahu zanesieme bod (FPR , TPR) do grafu, ako je vidno na obrázku 16.7. Pri $h = 0$ sú všetky vzorky klasifikované ako pozadie (negatívne). To znamená, že $FP = 0$ a aj $TP = 0$ a teda aj $FPR = 0$ a $TPR = 0$ (bod A na obrázku). Postupným zvyšovaním prahu sa zvyšuje hodnota TPR , pričom FPR môže byť nulová, prípadne skoro nulová (bod B na obrázku). Pri určitej hodnote prahu sa začne zvyšovať aj hodnota FPR (bod C na obrázku). V bode D na obrázku je prah Bayesovho klasifikátora (porovnaj s obrázkom 14.1(b)). Od bodu E je hodnota FPR približne 100 %, čiže všetky objekty máme správne klasifikované. Posledným bodom ROC krivky je bod (100 %, 100 %), keď všetky vzorky sú klasifikované ako objekt.

V praxi nemôžeme vyhodnotiť všetky možné hodnoty parametra klasifikátora (napríklad, keď zmena parametra vyžaduje nové tréningové klasifikátora) a výsledná krivka je schodovitá. Preto sa ROC krivka generuje pomocou vzájomnej validácie, keď v niekoľkých behoch vygenerujeme ROC krivky a spriemerujeme ich. Prípadne použijeme rôzne štatistické metódy odhadov rozdelenia pravdepodobnosti, z ktorých potom vygenerujeme ROC krivku.



Obrázok 16.7: Body ROC krivky zodpovedajú rôznym hodnotám klasifikačného prahu.

Výber parametra klasifikátora pomocou ROC krivky

Oblasť medzi bodom A a C na obrázku 16.7 je oblasť prísneho klasifikátora, kde veľmi málo vzoriek objektov je klasifikovaných ako objekt. Na druhej strane v tejto oblasti nemáme skoro žiadne nesprávne pozitívne klasifikácie.

Oblasť medzi bodmi E a G je oblasť nedbanlivého klasifikátora, kde všetky vzorky objektov klasifikujeme správne, ale za cenu veľmi vysokých hodnôt FP .

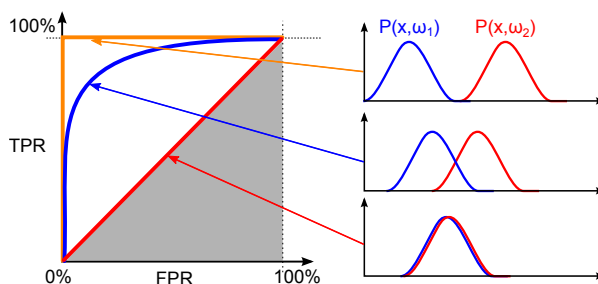
Oblasť optimálneho klasifikátora leží medzi bodmi C a E.

Parameter výsledného klasifikátora potom môžeme vybrať podľa toho, čo je cieľom klasifikácie. Ak je náš klasifikátor jedným z prvých v rade za sebou idúcich klasifikátorov, môžeme ho skonštruovať ako nedbalý, niekde v okolí bodu E, pričom sa spoľahneme na to, že nasledujúce klasifikátory postupne znížia FPR .

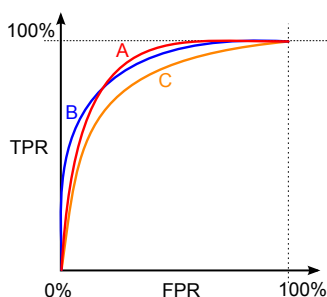
Ako najlepší sa často vyberá klasifikátor, ktorý leží na ROC krivke čo najbližšie k ľavému hornému rohu, a teda maximalizuje citlivosť a zároveň výlučnosť. Ďalším používaným bodom je bod s rovnakou hodnotou citlivosti a výlučnosti. Tento leží na prieniku ROC krivky s druhou diagonálou.

Priebeh ROC krivky

Tvar ROC krivky závisí od toho, ako dobre sa dajú odseparovať vzorky jednotlivých tried. Na obrázku 16.8 sú uvedené tri základné typy ROC kriviek. V prvej situácii vieme nájsť ideálny prah, keď klasifikátor správne identifikuje všetky vzorky objektu aj pozadia. Táto situácia je v reálnych aplikáciách veľmi zriedkavá. Príkladom typickej ROC krivky je krivka pre



Obrázok 16.8: Tvar ROC krivky zodpovedá rozdeleniu pravdepodobnosti.



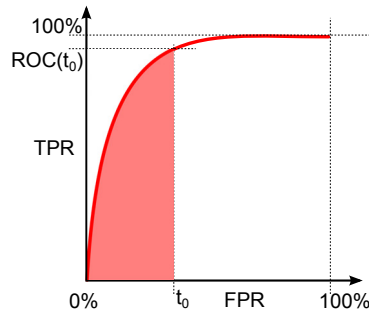
Obrázok 16.9: Porovnanie troch klasifikátorov pomocou ROC kriviek.

rozdelenie v strede, kedy sa pri klasifikácii snažíme minimalizovať dosiahnutú chybu. V najhoršom prípade, keď rozdelenia pravdepodobnosti oboch tried splývajú, ROC krivka splýva s diagonálou a kvalita klasifikácie je rovnaká, ako keby sme si hádzali mincou.¹

Porovnanie kvality klasifikátora pomocou ROC kriviek

Na obrázku 16.9 je porovnanie výsledkov rôznych klasifikátorov. Problémom je určiť, ktorý z klasifikátorov je lepší. Vidíme, že klasifikátor C dosahuje horšie výsledky než klasifikátory A a B. Medzi týmito dvomi však nie je jeden jednoznačne lepší. Pri nízkych hodnotách *FPR* sa javí ako lepší klasifikátor B, ale klasifikátor A dosiahne vyššiu citlivosť, pre vyššie hodnoty *FPR*. Jednou z možností, ako určiť, ktorý z A a B je lepší, je spočítať obsah plochy pod krivkou.

¹Ak by ROC krivka nášho klasifikátora ležala v šedej oblasti, smutne sklopíme hlavu a ideme vrátiť titul.



Obrázok 16.10: Obsah oblasti pod ROC krivkou obmedzená maximálne tolerovanou hodnotou $FPR = t_0$.

AUC hodnota

Integrál ROC krivky vyjadruje plochu pod ROC krivkou. Tomu zodpovedá aj anglické označenie *AUC* (area under the curve). Hodnota *AUC* sa pohybuje v intervale $(0, 1)$, pričom hodnotu 1 dosahuje ideálny klasifikátor a 0,5 náhodný. Kvalita klasifikátora sa pomocou hodnoty *AUC* dá vyjadriť pomocou stupnice v tabuľkách 16.4. Závisí od typu riešeného problému, ktorú tabuľku použijeme.

Tabuľka 16.4: Dve používané tabuľky *AUC* hodnôt.

Rozpätie	Stupnica	Rozpätie	Stupnica
0,5 – 0,6	zlyhávajúci	0,5 – 0,75	priemerný
0,6 – 0,7	slabý	0,75 – 0,92	dobrý
0,7 – 0,8	priemerný	0,92 – 0,97	veľmi dobrý
0,8 – 0,9	dobrý	0,97 – 1,0	vynikajúci
0,9 – 1,0	vynikajúci		

V niektorých aplikáciách nemá zmysel uvažovať parametre klasifikátora, pri ktorých sa dosahuje vysoká hodnota *FPR*. V tomto prípade sa nepočíta integrál na celom obore $(0, 1)$, ale len po nejakú maximálne tolerovanú hodnotu $FPR = t_0$, ako je znázornené na obrázku 16.10.

Štatistické porovnanie

Štatisticky sa rozdielnosť dvoch klasifikátorov pomocou hornoty AUC dá porovnať použitím nasledovnej testovacej štatistiky

$$CR = \frac{AUC_1 - AUC_2}{\text{Std}(AUC_1 - AUC_2)}, \quad (16.14)$$

ktorá má $(0, 1)$ normálne rozdelenie. Hypotéza H_0 je, že rozdiel medzi hodnotami AUC je nulový.

16.4.4 McNemarov test

Majme klasifikátory $K1$ a $K2$, ktorých kvalitu chceme porovnať. Na výpočet testovacej štatistiky McNemarovho testu pre oba klasifikátory určíme nasledovné hodnoty

- n_{00} : počet vzoriek správne klasifikovaných oboma klasifikátormi,
- n_{01} : počet vzoriek správne klasifikovaných $K1$ ale nesprávne $K2$,
- n_{10} : počet vzoriek nesprávne klasifikovaných $K1$ ale správne $K2$,
- n_{11} : počet vzoriek nesprávne klasifikovaných oboma klasifikátormi.

Následne vypočítame testovaciu štatistiku

$$M = \frac{(|n_{01} - n_{10}| - 1)^2}{n_{01} + n_{10}}, \quad (16.15)$$

ktorá má χ^2 rozdelenie s jedným stupňom voľnosti. To znamená, že ak $M > 3,84$, potom s 95% spoľahlivosťou môžeme zamietnuť nulovú hypotézu, že oba klasifikátory majú rovnakú chybovosť.





Zoznam použitej literatúry

- [Ami02] Y. Amit. *2D object detection and recognition: models, algorithms, and networks*. The MIT Press, Cambridge, Massachusetts, 2002.
- [Bis06] C. M. Bishop. *Pattern Recognition and Machine Learning*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006.
- [Bra97] A. P. Bradley. „The use of the area under the ROC curve in the evaluation of machine learning algorithms“. *Pattern Recognition* 30 (1997), str. 1145–1159.
- [CC00] L. d. F. Costa a R. Cesar. *Shape Analysis and Classification: Theory and Practice*. Boca Raton, FL, USA: CRC Press, Inc., 2000.
- [CDB86] R. L. Cannon, J. V. Dave a J. Bezdek. „Efficient Implementation of the Fuzzy c-Means Clustering Algorithms“. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 8.2 (1986), str. 248–255.
- [DG06] J. Davis a M. Goadrich. „The Relationship Between Precision-Recall and ROC Curves“. *Proceedings of the 23rd International Conference on Machine Learning, ICML '06*. Pittsburgh, Pennsylvania: ACM, 2006, str. 233–240.
- [DK09] F. De la Torre a T. Kanade. „Discriminative Cluster Analysis“. *Theory and Novel Applications of Machine Learning*. Ed. M. Er a Y. Zhou. In-teh, Croatia, 2009.

- [DZ00] B. Dawant a A. Zijdenbos. „Image segmentation“. *Handbook of Medical Imaging*. Ed. J. M. Fitzpatrick a M. Sonka. Zv. 2. SPIE Press, Bellingham, WA, 2000. Kap. 2.
- [Far97] I. Farkaš. „Samoorganizujúce sa mapy“. *Úvod do teórie neurónových sietí*. Ed. V. Kvasnička a kol. IRIS, Bratislava, 1997, str. 142–189.
- [Faw06] T. Fawcett. „An Introduction to ROC Analysis“. *Pattern Recogn. Lett.* 27.8 (jún 2006), str. 861–874.
- [FP02] D. A. Forsyth a J. Ponce. *Computer Vision: A Modern Approach*. Prentice Hall Professional Technical Reference, 2002.
- [FP91] D. Fisher a M. Pazzani. „Computational Models of Concept Induction“. *Concept Formation: Experience and Knowledge in Unsupervised Learning*. Ed. D. Fisher, M. Pazzani a P. Langley. San Mateo, CA: Morgan Kaufmann, 1991, str. 3–44.
- [Kru+13] R. Kruse a kol. *Computational Intelligence – A Methodological Introduction*. Texts in Computer Science. Springer, 2013.
- [MH84] B. J. McNeil a J. A. Hanley. „Statistical Approaches to the Analysis of ROC curves“. *Medical Decision Making* 4.2 (1984), str. 136–149.
- [MM01] W. Martinez a A. Martinez. *Computational Statistics Handbook with MATLAB*. Chapman & Hall/CRC Computer Science & Data Analysis. Taylor & Francis, 2001.
- [MR10] O. Maimon a L. Rokach, ed. *Data Mining and Knowledge Discovery Handbook*. Druhé vyd. Springer, 2010.
- [NA08] M. Nixon a A. Aguado. *Feature extraction and image processing*. Academic Press. Academic, 2008.
- [Náv+02] P. Návrat a kol. *Umelá inteligencia*. Vydavateľstvo STU, Bratislava, 2002.
- [Ora+13] M. Oravec a kol. *Metódy strojového učenia na extrakciu príznakov a rozpoznávanie vzorov*. Zv. II. FELIA s.r.o., 2013.
- [PWG02] S. W. Perry, H. S. Wong a L. Guan. *Adaptive Image Processing: A Computational Intelligence Perspective*. Boca Raton, FL, USA: CRC Press, Inc., 2002.
- [Qui86] J. R. Quinlan. „Induction of Decision Trees“. *Mach. Learn.* 1.1 (mar. 1986), str. 81–106.
- [SA96] P. Sinčák a G. Andrejková. *Neurónové siete (Inžiniersky prístup)*. Zv. I. ELFA Press, Košice, 1996.

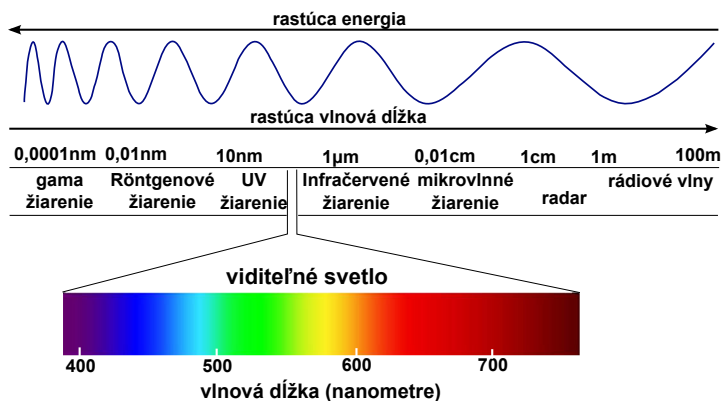
- [SS01] L. G. Shapiro a G. C. Stockman. *Computer Vision*. Prentice Hall, 2001.
- [Sze10] R. Szeliski. *Computer Vision: Algorithms and Applications*. Springer London, 2010.
- [TK10] S. Theodoridis a K. Koutroumbas. *Pattern Recognition*. Štvrté vyd. Academic Press, 2010.

ČASŤ IV

FARBA A VIZUÁLNE VNÍMANIE

Svetlo a farba

Osvetlenie je dôležitou a nepostrádateľnou zložkou životného prostredia. Služi materiálnym aj kultúrnym potrebám, jeho úroveň je významným ukazovateľom úrovne spoločnosti [CIE13].



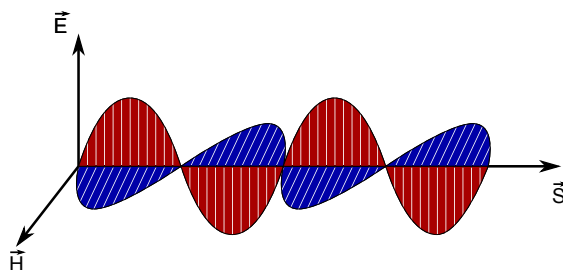
Obrázok 17.1: Elektromagnetické vlny. Viditeľné spektrum spolu s menami neviditeľných vln.

Podľa Maxwellovej teórie elektromagnetických vln má svetlo charakter elektromagnetického vlnenia s rozličnými dĺžkami. Fakt, že dokážeme vnímať farby, spôsobuje schopnosť oka reagovať na elektromagnetické vlny určitej dĺžky. Len veľmi úzka časť elektromagnetického spektra (vi-

ditel'ná časť spektra) sa nazýva **svetlo**. Nachádza sa v rozsahu vlnových dĺžok medzi 360 nm a 780 nm. Viditeľné farby v tomto rozsahu sa nazývajú **spektrálne, čisté farby**. Teda čisté farby zodpovedajú jednotlivým vlnovým dĺžkam.

Presnejšie povedané tento rozsah vlnových dĺžok je viditeľným svetlom pre človeka. Niektoré druhy živočíchov vnímajú iné rozsahy. Napríklad včely ho majú posunutý smerom ku kratším vlnovým dĺžkam, a teda môžu „vidieť“ ultrafialové žiarenie. Naopak niektoré plazy vnímajú ako viditeľné aj to, čo je pre človeka už neviditeľné infračervené žiarenie. Rozsah vnímaných vlnových dĺžok je daný predovšetkým tým, že v oblasti viditeľného svetla je maximum elektromagnetického žiarenia dopadajúceho na zemský povrch zo Slnka. Na obrázku 17.1 je zobrazený interval farieb zodpovedajúci rôznym vlnovým dĺžkam spolu s menami neviditeľných vĺn.

Svetelné vlnenie je priečne vlnenie, pri ktorom intenzita elektrického poľa $\vec{E}$ a intenzita magnetického poľa $\vec{H}$ sú kolmé na smer postupu vlnenia a zároveň vektory $\vec{E}$ a $\vec{H}$ sú na seba kolmé. Elektromagnetická vlna je znázornená na obrázku 17.2. Svetelné vlnenie má periodickú povahu a môžeme ho rozložiť na zložky so sínusovým priebehom monochromatického žiarenia. Monochromatické žiarenie sa prejavuje vždy jednou určitou farbou, preto sa nazýva aj jednofarebné svetlo alebo spektrálna farba. **Spektrálna krivka** je graf intenzity v závislosti od vlnovej dĺžky.



Obrázok 17.2: Elektromagnetická vlna.

Elektromagnetické vlnenie, a teda aj svetlo, charakterizujú tri základné vlastnosti:

- amplitúda (svietivosť),
- frekvencia (alebo farba),
- polarizácia (uhol vlnenia).

Svetelné vlnenie môžeme charakterizovať vlnovými dĺžkami λ alebo frekvenciami f . Vzťah medzi nimi je daný ako pomer

$$\lambda = \frac{v}{f}, \quad (17.1)$$

kde v je fázová rýchlosť. Elektromagnetické vlny sa šíria dokonca aj vo vákuu. Vtedy je vzťah medzi frekvenciou a vlnovou dĺžkou (λ) elektromagnetického žiarenia

$$f\lambda = c, \quad (17.2)$$

kde c je rýchlosť svetla vo vákuu.

Energia E vlny súvisí s jej frekvenciou podľa vzťahu

$$E = hf, \quad (17.3)$$

kde h je Planckova konštanta, nazvaná podľa nemeckého fyzika Maxa Plancka.

Existujú optické javy, ktoré sa nedajú opísať pomocou Maxwellovej elektromagnetickej teórie. Možno ich však vysvetliť pomocou kvantového modelu svetla. Podľa kvantovej teórie sa svetlo skladá z mnohých diskretných svetelných kvánt – **fotónov**. Energia fotónu sa vyjadruje v jednotkách elektrovolt (eV). Fotón ako častica elektromagnetického poľa je daná okrem energie aj svojou hmotnosťou a hybnosťou. Pokojová hmotnosť fotónu je nulová, inak by sa nemohol pohybovať rýchlosťou svetla.

Kvôli dualite častice a vlnenia má svetlo vlastnosti aj vlnové aj časticové.

17.1 Veličiny opisujúce svetlo

Elektromagnetické vlnenie je nositeľom energie. Dôležitým faktorom popri veľkosti **hustoty energie žiarenia** je aj jej **priestorové rozloženie**. Vlnenie sa môže líšiť tiež svojím spektrálnym zložením. Aby bolo možné elektromagnetické žiarenie využiť, je potrebné jeho energetické charakteristiky poznať a vedieť ich merať. Týmto meraním sa zaoberá radiometria.

Vlastnosti svetla, bez ohľadu na detektor a jeho vlastnosti, opisujú radiometrické veličiny. Ak chceme merať vlastnosti svetla podľa jeho účinku na ľudské oko, používame fotometrické veličiny.

17.1.1 Rádiometrické veličiny

Rádiometrické veličiny sa niekedy nazývajú aj energetické, preto sa na ich označenie používa index e . V nasledujúcom texte si uvedieme základné rádiometrické veličiny.

Žiarivý tok

Žiarivý tok Φ_e je energia vyžiarená telesom za jednotku času (resp. výkon vyžiarený, absorbovaný, alebo prenášaný vlnením)

$$\Phi_e = \frac{dE}{dt}, \quad \text{jednotka } W.$$

Intenzita vyžarovania

Intenzita vyžarovania M_e je energia vyžiarená jednotkovou plochou za jednotku času

$$M = \frac{d\Phi_e}{dS}, \quad \text{jednotka } W.m^{-2}.$$

Žiarivosť

Žiarivosť I_e je energia vyžiarená bodovým zdrojom v danom smere do malého priestorového uhla

$$I_e = \frac{d\Phi_e}{d\Omega}, \quad \text{jednotka } W.sr^{-1},$$

kde sr je označenie pre steradián, jednotku priestorového uhla.

Doteraz definované veličiny opisujú žiarenie ako celok, bez ohľadu na to, aké vlnové dĺžky sú v ňom obsiahnuté.

Spektrálna hustota žiarivého toku

Spektrálna hustota žiarivého toku $\Phi_{e\lambda}$ je energia vyžiarená za jednotku času, ktorá pripadá na interval vlnových dĺžok $d\lambda$

$$\Phi_{e\lambda} = \frac{d\Phi_e}{d\lambda}.$$

Celkový žiarivý tok

Celkový žiarivý tok získame integrovaním cez celý interval vlnových dĺžok (λ_1, λ_2) , v ktorom zdroj emituje žiarenie

$$\Phi_e = \int_{\lambda_1}^{\lambda_2} \Phi_{e\lambda} d\lambda.$$

17.1.2 Fotometrické veličiny

Všetky fotometrické veličiny a jednotky sa definovali na základe fyzikálnych veličín a jednotiek prostredníctvom **kriviek spektrálnej citlivosti ľudského oka** $V(\lambda)$ pre denné (fotopické) a $V'(\lambda)$ pre nočné (skotopické) videnie (viď kapitolu 18).

Svietivosť

Fotometrická veličina charakterizujúca bodový zdroj je svietivosť. Je to veličina, ktorá vyjadruje schopnosť približne bodového zdroja vyvolať v danom smere zrakový vnem. Jej jednotkou je **kandela** (cd), ktorá patrí medzi základné jednotky sústavy SI a od nej sa odvodzujú jednotky ostatných fotometrických veličín. Je analógiou k žiarivosti a zapisuje sa

$$I = \frac{d\Phi}{d\Omega}, \quad \text{jednotka cd (kandela).}$$

Svetelný tok

Odvozené fotometrické jednotky sú lúmen a lux. **Lúmen** (lm) je jednotkou **svetelného toku**. Keď bodový zdroj svetla vysiela svetelné žiarenie v smere určenom uhlom α do malého priestorového uhla, v ktorého vrchole sa nachádza, hovoríme o svetelnom toku

$$d\Phi = I^\alpha d\Omega, \quad \text{jednotka cd.sr, tj. lúmen (lm).}$$

Osvetlenie

Ak bodový zdroj so svietivosťou I vysiela svetelný tok $d\Phi$ do malého priestorového uhla $d\Omega$, pri kolmom dopade na malú plochu dS spôsobuje na nej **osvetlenie**

$$E = \frac{d\Phi}{dS}, \quad \text{jednotka lm.m}^{-2}, \text{ tj. lux.}$$

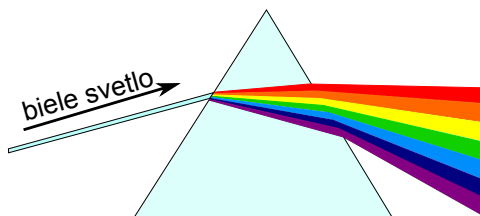
Jednotkou osvetlenia je **lux**.

Jas

Plošné svetelné zdroje charakterizuje **jas** L . Jas vyjadrujeme ako podiel svetivosti plošného elementu v danom smere k priemetu plošného elementu do roviny kolmej na uvažovaný smer.

17.2 Funkcie spektrálnej hustoty

V 17. storočí urobil Isaac Newton radikálny objav, keď rozložil biele svetlo na spojité spektrum farieb pri prechode svetla cez sklenený hranol (viď obrázok 17.3) a vyvrátil tak predchádzajúcu teóriu, že farby vznikajú zmiešavaním svetla a tmy (temnoty).



Obrázok 17.3: Rozklad bieleho svetla na spektrum farieb po prechode cez sklenený hranol.

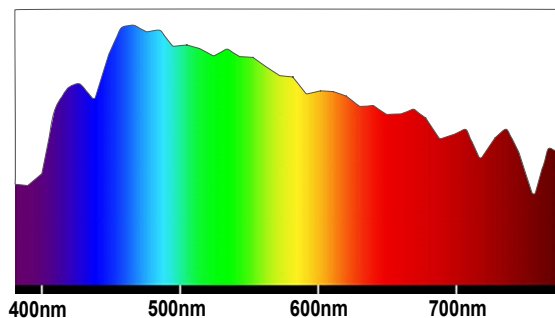
V súčasnosti pod pojmom spektrum rozumieme rozklad žiarenia podľa vlnových dĺžok alebo frekvencií. Spektrum môže byť spojité alebo nespojité v závislosti od skupenstva látkových objektov vysielajúcich žiarenie. Nespojité spektrum sa ešte rozdeľuje na čiarové a pásové spektrum.

Spektrum charakterizujeme pomocou pomerného spektrálneho zloženia žiarenia, ktoré zobrazuje závislosť pomernej spektrálnej hustoty žiarivého toku od vlnovej dĺžky.

Svetlo vyžarované zo zdroja vo väčšine prípadov obsahuje žiarenie viacerých vlnových dĺžok. Spektrálna krivka intenzity žiarenia pre takúto zdroj svetla je znázornená na obr. 17.4. Energia svetla sa vypočíta ako plocha pod spektrálnou krivkou.

17.3 Geometrické vlastnosti svetla

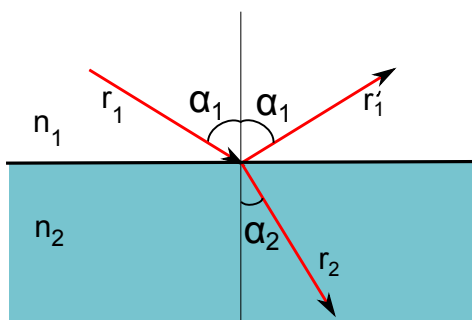
Základným pojmom geometrickej optiky je svetelný lúč. Svetelný lúč dopadajúci na povrch látky sa čiastočne odrazí, čiastočne prejde látkou a čiastočne ho látka pohltí. Dopadajúci lúč sa odrazí podľa zákonov odrazu.



Obrázok 17.4: Spektrálna krivka intenzity žiarenia pre svetlo s viacerými vlnovými dĺžkami.

Svetlo riadia tieto zákony geometrickej optiky

1. **Zákon priamočiareho šírenia svetla:** V homogénnom prostredí sa svetlo pohybuje po priamkach (svetelné lúče).
2. **Zákon odrazu:** Odrazený lúč leží v rovine dopadu určenej normálovým vektorom dopadu v danom bode dopadu a uhol medzi normálou a oboma lúčmi je komplementárny.
3. **Zákon lomu (Snell–Descartes):** Lúč prechádzajúci z jedného prostredia do druhého mení na rozhraní týchto prostredí svoj smer.



Obrázok 17.5: Čiastočný odraz a lom svetelného lúča dopadajúceho na rozhranie dvoch priesvitných materiálov.

Podľa zákona lomu, ak r_1 je lúč dopadajúci na rozhranie dvoch priesvitných materiálov s indexami lomu n_1 a n_2 (viď obr. 17.5), a r_2 je lomený

lúč, potom r_1 , r_2 a normála na rozhranie sú koplanárne a pre uhly α_1 a α_2 medzi normálou a dvomi lúčmi platí vzťah

$$n_1 \cdot \sin \alpha_1 = n_2 \cdot \sin \alpha_2. \quad (17.4)$$

Proces formovania obrazu pozostáva z troch častí: osvetlenie, odrazivosť materiálu a pozorovanie. Materiály v scéne interagujú s prichádzajúcim svetlom a spôsobujú jeho reflexiu, odraz. Farba objektu je daná spektrom dopadajúceho svetla a svetlom absorbovaným a/alebo odrazeným (vlastnosťami objektu). V prípade matných materiálov použijeme Lambertov model odrazu. Hovoríme, že intenzita odrazeného svetla od povrchu je nezávislá od uhla pozorovania.

17.4 Žiarenie čierneho telesa

Každé teleso s teplotou vyššou ako je absolútna nula (0 K tj. $-273,15^\circ\text{C}$) vyžaruje do okolia elektromagnetické žiarenie. Vlnová dĺžka, na ktorej toto teleso najviac vyžaruje, závisí od jeho teploty. Čím väčšia je teplota telesa, tým kratšiu vlnovú dĺžku má žiarenie, ktoré vysielá do priestoru. Táto vlnová dĺžka sa dá jednoducho vypočítať podľa Wienovho zákona posunu, ktorý hovorí, že vlnová dĺžka λ_{max} , na ktorú pripadá maximum vyžarovania čierneho telesa, je nepriamo úmerná termodynamickej teplote T

$$\lambda_{max} = \frac{b}{T},$$

kde $b = 2,898 \cdot 10^{-3} \text{ m K}$ je konštanta a T je teplota povrchu telesa v Kelvinoch.

Schopnosť telesa vysielat' elektromagnetické žiarenie úzko súvisí s jeho schopnosťou pohlcovať žiarenie, pretože teleso pri konštantnej teplote je v termodynamickej rovnováhe so svojím okolím, teda získava pohlcovaním energie od okolia rovnaké množstvo energie, ako do okolia odovzdáva.

Príklad: Železnú tyč v tme pri izbovej teplote okolo 20°C (293 K) neuvídime. Do okolia vyžaruje elektromagnetické žiarenie s vlnovou dĺžkou okolo $10 \mu\text{m}$, čo je vlnová dĺžka neviditeľného, infračerveného žiarenia. Keď ju však začneme zahrievať, začne sa jej teplota zvyšovať a s ňou sa bude skracovať vlnová dĺžka žiarenia, ktorá pri určitej teplote začne patriť do viditeľnej časti spektra, na ktorú sú citlivé naše oči. Rozžeravenú tyč začne byť vidno a nadobudne červenú farbu.

Ako sme už spomenuli vyššie, väčšina telies, ktoré poznáme, časť žiarenia na ne dopadajúceho odrazí, časť pohltí, a časť žiarenia nimi prejde. Aby bolo možné jednoducho pracovať so žiarením telies, bol zavedený fyzikálny model – **absolútne čierne teleso**. Je to idealizované teleso, ktoré žiadne žiarenie neodrazí ani cezeň žiadne neprejde, ale dokonale pohltí všetku energiu svetla, ktoré naň dopadá a tú vyšle v podobe tepelného žiarenia. V praxi sa ako čierne teleso správa napríklad dutina so začíernenými vnútornými stenami a malým otvorom, z ktorého žiarenie vystupuje.

Ako sa mení teplota, odtieň farby žiarenia sleduje čiaru teplotných žiaričov, tzv. **Planckovu krivku** (*Planckian locus*). Planckova krivka opisuje spektrálne rozdelenie intenzity žiarenia čierneho telesa pri vyznačených teplotách. Túto závislosť popísal v roku 1900 Max Planck (1858 – 1947), ktorý získal za myšlienku existencie kvánt energií v roku 1918 Nobelovu cenu za fyziku.

S rastúcou teplotou sa vlnová dĺžka maximálneho vyžarovania skracuje a tiež sa zväčšuje množstvo vyžarovanej energie. To znamená, že ak sa povie, že nejaké žiarenie zodpovedá žiareniu čierneho telesa s teplotou x Kelvinov, tak spektrálna krivka tohto žiarenia bude Planckova krivka príslušnej vlnovej dĺžky.

Opísali sme parameter farby svetla zodpovedajúci vyžarovaniu čierneho telesa, a to parameter teploty. Môžeme teda hovoriť o farebnej teplote svetelného zdroja. Je to teplota čierneho telesa, ktorého farba žiarenia sa naň najviac podobá. Svetlo určitej farebnej teploty má farbu tepelného žiarenia vydávaného čiernym telesom, zahriatym na túto teplotu. Veľa prirodzených zdrojov svetla môže byť aproximovaných žiarením ideálneho čierneho telesa.



L'udský vizuálny systém

Zrak je pre človeka zo všetkých zmyslov považovaný za najdôležitejší. Pomocou zraku vníma človek asi 80 % všetkých informácií zo svojho okolia. Zmyslovým orgánom zraku je oko, ktorým človek vníma vizuálne obrazy okolitého sveta.

Fakt, že dokážeme vnímať farby, spôsobuje schopnosť oka reagovať na elektromagnetické vlny určitej dĺžky. Svetlo, teda elektromagnetické žiarenie z okolitého sveta, sa cez našu optickú sústavu „oko“ prenáša na receptorové bunky (tyčinky a čapíky) na sietnici, kde vyvolávajú chemickú reakciu. Tieto potom prenášajú vyvolané vzruchy pomocou nervových vlákien do mozgu, ktorý ich spracuje na vnímaný obraz.

V rôznych častiach oka nastáva pohlcovanie rôznych vlnových dĺžok žiarenia. Tento fakt sa dá využiť napríklad pri operáciách oka, ale predovšetkým chráni oko pred poškodením. Žiarenie je pohlcované pomocou rozkladu proteínov a iných látok, čím sa spotrebovávajú energia. Vyššie dávky nevhodného žiarenia spôsobujú slzenie, zvýšenie teploty a tlaku v oku, zápal a podobne.

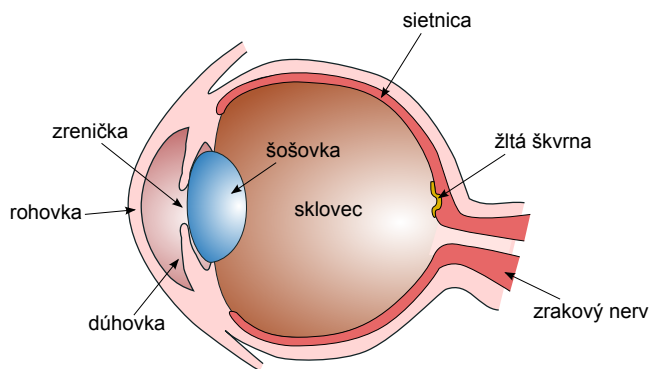
- 100 – 315 nm – absorbuje sa prevažne v rohovke, zvyšok sa rozptýli v sklovci,
- 315 – 400 nm – absorbuje sa prevažne v šošovke za pomoci premeny proteínov,
- 400 – 780 nm – viditeľné svetlo je oko schopné v priebehu 0,25 sekundy zredukovať pomocou zreničky na znesiteľné množstvo, ale na kratších vlnových dĺžkach už nedokáže tak rýchle zareagovať,

- 780 – 1400 nm – prechádza cez šošovku a dopadá na sietnicu, kde môže spôsobiť aj vážne poškodenie.
- viac ako 1400 nm – je absorbované v rohovke. Spôsobuje silné slzenie a zvyšovanie teploty a tlaku sklovca.

18.1 Štruktúra oka

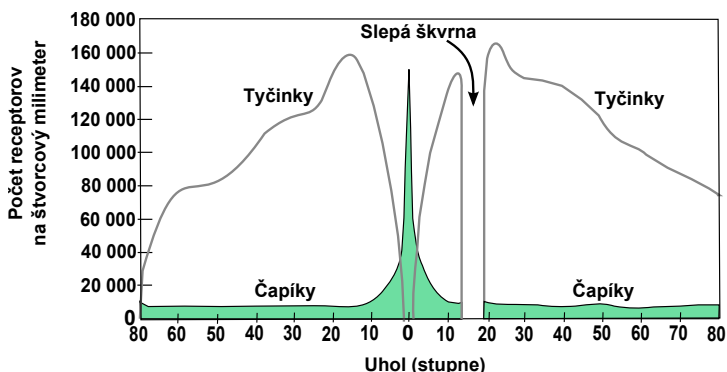
Svetlo vstupuje do oka cez **šošovku** (*lens*). Zmenou tvaru šošovky sa mení ohnisková vzdialenosť, zabezpečuje sa ostré premietnutie obrazu na sietnicu. Toto prispôsobenie šošovky sa nazýva akomodácia.

Pred šošovkou sa nachádza kruhovitý sval nazývaný **dúhovka** (*iris*). Najdôležitejšou úlohou dúhovky je regulácia množstva svetla vstupujúceho do oka. Dúhovka má centrálny okrúhly pohyblivý otvor – **zreničku** (*pupil*), ktorá sa môže zužovať alebo rozširovať. Vo vnútri oka sa nachádza svetlocitlivá membrána – **sietnica** (*retina*). Sietnica je vlastným orgánom zraku a tvorí vnútornú vrstvu oka. Pokrýva celú dutinovú stranu očnej gule. Zachytáva obraz a zrakovým nervom ho posiela do zrakových centier v mozgu. Je tvorená z dvoch druhov svetlocitlivých receptorov (fotoreceptorov) – **čapíkov** (*cones*) a **tyčíniek** (*rods*). Štruktúra oka je znázornená na obrázku 18.1.



Obrázok 18.1: Prierezný rez okom zobrazujúci jeho štruktúru.

V strede sietnice, približne oproti zreničke oka, leží žltá škvrna kruhovitého až eliptického tvaru o priemere asi 3 mm. V jej strede je preliačina, ktorá sa nazýva **fovea centralis**. Väčšina čapíkov je sústredených vo vnútri tejto malej oblasti. Toto miesto leží vo vrchole optickej osi oka a predstavuje miesto najostrejšieho videnia, pretože sa do neho premieta



Obrázok 18.2: Distribúcia tyčínok a čapíkov v sietnici. Čapíky sú väčšinou sústredené v centre sietnice, v žltej škvrne. Priestorové rozlišovanie je preto výrazne slabšie na perifériách.

centrálny lúč. V oblasti žltej škvrny sú nakopené len čapíky, ktoré tu majú najväčšiu hustotu. Tyčinky sa nachádzajú zväčša na periférii sietnice. Distribúcia tyčínok a čapíkov v sietnici je znázornená na obrázku 18.2.

Tyčinky a čapíky

Na sietnici je neuniformné rozloženie čapíkov a tyčínok. Vo fovei pri maximálnej hustote čapíkov je ich 150 tis. – 200 tis./mm².

Tyčinky obsahujú látku citlivú na achromatické svetlo, nazývanú **rodopsin**. Na základe fotochemických reakcií očného rhodopsinu je vnímaná intenzita svetla. Čapíky obsahujú obdobnú látku **iodopsin**, ktorý je zložený z viacerých látok citlivých na svetlo modrej, zelenej a červenej farby. Tieto organické farbivá senzibilizujú čapíky a tyčinky.

Čapíky slúžia na farebné videnie v jasných svetelných podmienkach, pri svetlosti > 1cd/m², keď sú dobre viditeľné a rozpoznateľné farby – denné videnie (photopic vision). Tyčinky neumožňujú farebné videnie, sú použité pri veľmi slabom svetle, približne pri svetlosti < 0,01cd/m², keď sú viditeľné iba šedotónové obrysy objektov – nočné videnie (scotopic vision). Tyčinky sú približne 25-krát citlivejšie než čapíky. Na súmrakovom videní (mesopic vision) sa podieľajú čiastočne tyčinky aj čapíky, aktivuje sa pri 0,01 – 1cd/m², keď farby nie sú jasne viditeľné.

18.2 Procesy prebiehajúce na sietnici

Počas histórie skúmania farebného videnia sa formovali dve základné teórie a to **trichromatická teória** a **teória oponentných procesov**, obe ešte pred objavením svetlocitlivých buniek v oku.

Trichromatická teória sa spája s menami Th. Younga a H. Helmholtza. Nezávisle na sebe dospeli k záveru, že všetky farby spektra môžeme vytvoriť zmiešaním troch základných monochromatických farieb. Young za tieto tri základné farby považoval červenú, žltú a modrú, Helmholtz zase červenú, zelenú a fialovú.

Filozofickým pozadím oponentnej teórie je koncept J. W. Goetheho. Goethe si všimol, že niektoré farby sa k sebe veľmi dobre hodia a niektoré sa spolu takmer nikdy nevyskytujú. Goetheho teória je filozofickým podkladom fyziologických výskumov, ktoré neskôr zrealizoval Ewald Hering a vytvoril teóriu oponentných farieb.

V ľudskom oku sú skutočne práve tri typy farebných receptorov, na ktorých prebiehajú mechanizmy na báze trichromatizmu, no spracovanie informácie z nich, prebiehajúce v hlbších štruktúrach sietnice a mozgovej kôre, má charakter oponentných procesov.

18.2.1 Trichromatické procesy

Ľudský vizuálny systém je založený na troch typoch čapíkov uložených na sietnici oka. Prvý typ čapíkov reaguje na červenú časť farebného spektra a teda na dlhé vlny. Druhý typ reaguje na zelenú časť spektra s vlnami strednej dĺžky. A tretí typ reaguje na modrú časť farebného spektra, kde sa nachádzajú krátke vlny. Maximálna odozva prvého typu čapíkov je pri vlnovej dĺžke 559 nm, druhého typu pri 531 nm a tretieho typu 419 nm. Na základe niektorých zistení môžu mať niektoré ženy tetrachromatický vizuálny systém – štyri typy čapíkov. Pomer počtu čapíkov citlivých na dlhé L, stredné M a krátke S vlny je $L : M : S = 60 : 30 : 5$. Tento pomer sa mení v závislosti od človeka, pomer $L : M$ môže byť u ľudí s normálnym videním od 2,7 : 1 až po 16,5 : 1.

Vnem z modrého kanálu ovplyvňuje aj tvorbu melatonínu a teda biologické hodiny: spánok/bdenie. Dynamický rozsah ľudského vizuálneho systému pokrýva rozsah intenzity svetla približne 10^{12} . Rozlišujeme adaptáciu na svetlo a na tmu. Adaptáciu na svetlo môžeme merať v sekundách, avšak adaptácia na tmu trvá minúty. Pri chromatickej adaptácii sa samostatne prispôsobujú jednotlivé primárne farebné kanály (L, M, S). Svetlocitlivé receptory merajú množstvo svetla vstupujúceho do retiny. Nech $S_i(\lambda)$, $i \in \{r, g, b\}$ je krivka odozvy jednotlivých typov čapíkov citli-

vých na červenú, zelenú a modrú časť spektra. $E(\lambda)$ je (irradiance) svetelné žiarenie s vlnovou dĺžkou λ dopadajúce na receptor. Potom celkové množstvo merané jedným receptorom môžeme vypočítať

$$Q_i = \int S_i(\lambda) E(\lambda) d\lambda. \quad (18.1)$$

Pri pozorovaní matného povrchu je žiarivosť E , dopadajúca na receptor, vypočítaná ako produkt medzi odrazenou energiou objektu R a množstvom energie L dopadajúceho na objekt

$$E(\lambda) = R(\lambda) L(\lambda), \quad (18.2)$$

takže celkovú energiu môžeme prepísať ako

$$Q_i = \int S_i(\lambda) R(\lambda) L(\lambda) d\lambda. \quad (18.3)$$

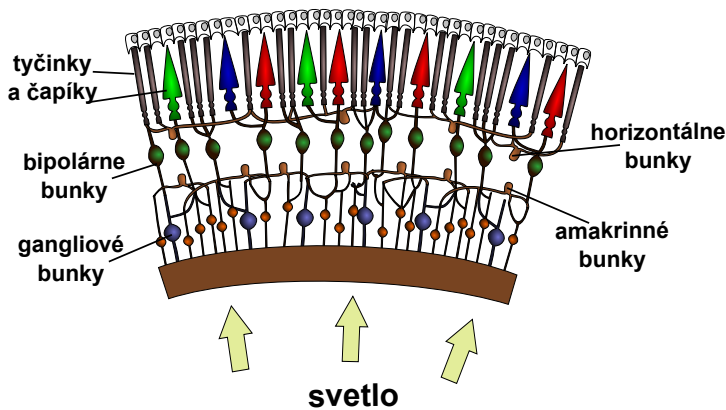
18.2.2 Oponentné procesy v sietnici

Trichomatická teória farebného videnia nie je schopná vysvetliť fenomén oponentných farieb. Táto skutočnosť viedla Ewalda Heringa k navrhnutiu oponentného princípu kódovania farieb. Oponentné procesy prebiehajú v sietnici.

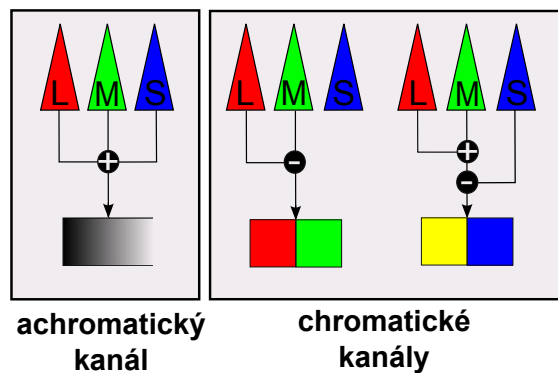
Sietnica (obr. 18.3) je komplikovaný viacúrovňový orgán. V zadnej časti sietnice sú uložené čapíky a tyčinky. **Gangliové bunky** (*ganglion retinal cells*) sa nachádzajú v prednej časti sietnice a v strednej časti medzi týmito dvomi úrovňami sú **bipolárne bunky** (*bipolar cells*). V tejto štruktúre sa nachádzajú aj **horizontálne** a **amakrinné bunky**, ktoré slúžia na horizontálnu výmenu informácií. Zaujímavé je, že prichádzajúce svetlo musí najprv prejsť cez úroveň v opačnom poradí, aby dosiahlo úroveň, kde sa nachádzajú fotoreceptory.

Odozva vychádzajúca z fotosenzitívnych receptorov je kombinovaná v gangliových bunkách sietnice, kde sa formujú **tri oponentské kanály** (obr. 18.4). Jeden achromatický čierno-biely a dva chromatické červeno-zelený a žltomodrý. Z gangliovej bunky sa pošle pulzový signál cez optický nerv do vizuálnej časti mozgovej kôry, kde je spracované samotné vnímanie farieb.

V retine je 100 až 126 miliónov fotoreceptorov, z toho približne 95 až 120 miliónov tyčiniek a 5 až 6 miliónov čapíkov. Keďže na prenos informácie existuje len asi 1 milión gangliových buniek, je zrejmé, že táto musí byť nejakým spôsobom komprimovaná. Ak by bol každý fotoreceptor spojený s individuálnou bunkou, neurónové vlákno by muselo mať



Obrázok 18.3: Štruktúra sietnice pozostávajúca z troch úrovní. Svetlo musí prejsť najprv pomedzi gangliové a bipolárne bunky, aby sa dostalo k fotosenzitívnym čapíkom a tyčinkám. Informácia potom prechádza v opačnom poradí naspäť do gangliových buniek a odtiaľ cez optický nerv do mozgu. Informácia sa vymieňa aj postranne cez horizontálne a amakrinné bunky.



Obrázok 18.4: Tri oponentské procesy pozostávajúce z jedeného achromatického kanála a dvoch chromatických. Chromatické kanály kombinujú červeno-zelenú a žltomodrú informáciu.

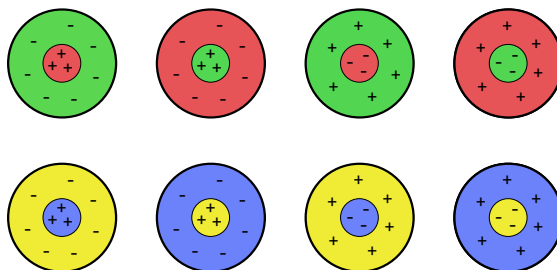
veľkú hrúbku. A teda je zrejmé, že predtým, ako sa zraková informácia z jednotlivých čapíkov pošle do mozgu, dochádza už v sietnici na gangliových bunkách k jej veľkému predspracovaniu. V mieste, kde optický nerv opúšťa oko, sa nenachádzajú žiadne fotoreceptory. Toto miesto sa nazýva **slepá škvrna**.

18.2.3 Gangliové bunky a receptívne polia

Bipolárne bunky prenášajú signály od receptorových buniek ku gangliovým bunkám. Axóny (nervové výbežky) gangliových buniek formujú očný nerv. Jedna gangliová bunka je spojená s viacerými fotoreceptormi cez bipolárne bunky. Gangliové bunky prijímajú vstup okrem bipolárnych aj z amakrinných buniek.

Všetky receptory, ktoré ovplyvňujú výstup jednej gangliovej bunky, patria oblasti, ktorá sa nazýva **receptívne pole gangliovej bunky**. Receptívne polia susedných sietnicových gangliových buniek sa výrazne prekrývajú.

Väčšina gangliových buniek má približne kruhové receptívne pole. To znamená, že malý kruh buniek na sietnici vysiela svoje signály do jednej gangliovej bunky. Tento kruh sa ešte delí na vnútorný (centrum) a vonkajší (okolie). Gangliové bunky vedia rozlíšiť, či svetlo dopadá do centra alebo do okolia a podľa toho reagujú. Keď svetlo dopadá do centra projekčného poľa bunky, jedna skupina gangliových buniek sa aktivizuje, tieto sa nazývajú bunky s **aktivizovaným centrom** (*on-center cells*). Keď svetlo dopadá na plochu ich okolia, sú inhibované. Bunky s **neaktívnym centrom** (*off-center cells*) sú inhibované, keď svetlo dopadá na ich centrum. Avšak, keď na centrum svetlo prestane dopadať, na chvíľu sa aktivizujú. Aktívne sú vtedy, keď svetlo dopadá na ich okolie [Dob05].



Obrázok 18.5: Receptívne polia sietnicových gangliových oponentných buniek s aktívnym a s neaktívnym centrom. Štyri červeno-zelené a štyri žltomodré typy oponentských buniek.

Ako možno vidieť na obrázku 18.5 receptívne polia buniek môžu byť červeno-zelené alebo žltomodré.

Z predchádzajúceho vyplýva, že vnem „červeno-zelenej“ farby neexistuje, pretože ak je bunka excitovaná červeným svetlom, zapne sa červený kanál a zároveň na druhom čapíku je signál zo zeleného kanála in-

hibovaný. Obidva kanály nepracujú synchronne ale skôr antagonisticky, buď vnímame predmet červený, alebo zelený.

Takisto existencia buniek citlivých na modrú a žltú umožňuje vysvetliť, prečo nikdy nevidíme farbu, ktorá by bola žltkasto modrá. Buď totiž vidíme modrú, alebo vidíme žltú.

18.3 Vizualný kortex

Spracovanie vizualnej informácie sa deje v mozgovej kôre (cortex). Tvar objektu, jeho farba, smer pohybu sa spracováva v rôznych častiach vizualnej časti mozgovej kôry. Tieto informácie sa delia na dve cesty – „čo“ a „kde“. Tvar a farba sa spracováva pomocou „čo“ cesty a priestorová informácia sa spracováva pomocou „kde“ cesty, ktorá taktiež rieši pohyb. Toto delenie sa začína už v sietnici.

Ďalšie delenie sietnicových gangliových buniek je na M bunky (magno – veľké) a P bunky (parvo – malé). Veľké bunky sú citlivé na veľké objekty a vedia sledovať rýchle zmeny v podnetoch. Nesú najmä informáciu o hrubých črtách predmetov a pohybe. Malých buniek je viac, majú menšie receptívne polia, nesú najmä informáciu o farbe a detailoch. V obidvoch skupinách sa vyskytujú ako off-center, tak on-center bunky [Dob05]. Asi 80 % sietnicových gangliových buniek sú P bunky. Reagujú na vysoké frekvencie a sú citlivé na rôzne vlnové dĺžky. Asi 10 % buniek je M buniek. Tie reagujú na nízke frekvencie a nie sú citlivé na vlnové dĺžky. Zvyšných 10 % buniek je tvorených asi ôsmimi ďalšími typmi buniek [Ebn07].

Zaujímavosť: Treba si uvedomiť, že obraz na našej sietnici sa neustále mení, pretože naše oči neustále vykonávajú sáky, tj. malé pohyby, ktorými sa zameriavajú stále na inú časť okolia. Ak by sme sa pozerali bez pohybu očí stále na jedno a to isté miesto, bunky, ktoré sú zodpovedné za prenos zrakovej informácie, by sa unavili, a obraz by sa nám pomaly vytratil. Tieto bunky potrebujú zmenu, aby boli aktívne. Vyskúšajte si napríklad uprene sa pozerat' na bodku na čistom liste papiera. Ak sa pozeráte skutočne sústredene, po čase vám bodka zmizne [Dob05].

Zraková kôra sa delí na viacero oblastí. Okrem primárnej zrakovej oblasti označovanej ako V1, rozoznávame ešte ďalšie dôležité oblasti ako V2, V3, V4 a V5 (MT), ktoré prispievajú k ďalšiemu spracovaniu zrakovej informácie.

Informáciu o pohybe prenáša najmä aktivita M-buniek. Prechádza oblasťou V2 do oblasti MT a ďalej do parietálnej kôry. Informáciu o tvare nesie najmä aktivita buniek P. Z oblasti V2 prechádza do oblasti V4 a ďalej do temporálnej kôry.

Zo sietnice je informácia o farbe (z troch typov čapíkov, z ktorých je každý citlivý na iné vlnové dĺžky; typ S na krátke, M na stredné a L na dlhé) vedená do kôry P-bunkami, o ktorých sme hovorili vyššie. Avšak tieto bunky zároveň vedú aj informáciu o jase, respektíve kontraste. Farba sa v P-bunkách spracováva podobným princípom, na akom fungujú „on-center“ a „off-center“ neuróny. Jeden typ P-buniek mieša signály z L a M čapíkov, druhý zasa signály z S čapíkov oproti kombinácii signálu z L a M čapíkov.

Vznik obrazov na sietnici sprevádza jav iradiácie. Predmety, ktoré majú vyšší jas ako okolie, sa zdajú byť o niečo väčšie ako rovnako veľké predmety nižšieho jasu.

18.4 Poruchy farbocitu

Pri vnímaní farieb môžu nastať rôzne poruchy. Porucha farbocitu môže byť vrodená alebo získaná. Existujú tri druhy vrodených porúch.

Prvým typom je **dichromázia**, pri ktorej jeden druh čapíkov chýba. Podľa toho, ktorý, rozlišujeme

Protanope – chýbajú čapíky citlivé na červenú farbu,

Deutanope – chýbajú čapíky citlivé na zelenú farbu,

Tritanope – chýbajú čapíky citlivé na modrú farbu.

Ľudia trpiaci na protanopiu alebo deutanopiu nerozlišujú červenú a zelenú farbu. Touto poruchou farbocitu je postihnutých približne 8 % mužov a 0,5 % žien.

Ďalším typom poruchy farbocitu je **anomálna trichromázia**. Pri tejto poruche je znížená citlivosť jedného typu čapíkov, čo spôsobuje posun farebného spektra a človek má posunuté vnímanie farieb.

Tretím typom je **monochromázia**, pri ktorej človek nerozlišuje farby vôbec.

Existujú však aj poruchy vnímania farieb, tvarov a pohybu spôsobené poškodením častí mozgu, kde sa táto informácia spracováva. Takou je napríklad **achromatopsia**, keď človek nerozoznáva farby z dôvodu poškodenia mozgu v oblasti V4. Poškodenie určitých častí oblasti V1 a V2 spôsobuje **chromatopsiu**, kedy človek nie je schopný vidieť tvary a formy,

ale farbu rozoznáva. Alebo **akinetopsia** spôsobená poškodením oblasti V5, ktorá spracováva informácie o pohybe a stereoskopickú hĺbku a pri ktorej človek nie je schopný vidieť objekty, ktoré sú v pohybe.

Kolorimetria

Za posledných tristo rokov poznáme viacero pokusov definovať farebné modely. Tie väčšinou vychádzali z aktuálnych vedomostí o farebnom vnímaní človeka. V mnohých prípadoch práve farebné modely mali vysvetliť, ako vnímanie farieb funguje. V roku 1801 Thomas Young prišiel s teóriou, že iba tri primárne farby sú postačujúce na vytvorenie všetkých farieb. Táto teória bola potvrdená v roku 1960, keď boli objavené tri typy fotoreceptorov v sietnici.

Na začiatku 20. storočia bola vytvorená medzinárodná vedecko-technická organizácia zameraná na študijné riešenie základných otázok rozvoja svetelnej techniky - **Medzinárodná komisia pre osvetľovanie** (*Commission Internationale de l'Eclairage* – CIE). Aktuálne združuje viac ako 40 krajín sveta.

Numerickým opisom ľudského vnímania farieb sa v súčasnosti zaoberá **kolorimetria**. Dnešný kolorimetrický aparát vychádza väčšinou z uznesení CIE. Medzi jej najvýznamnejšie odporúčenia patria:

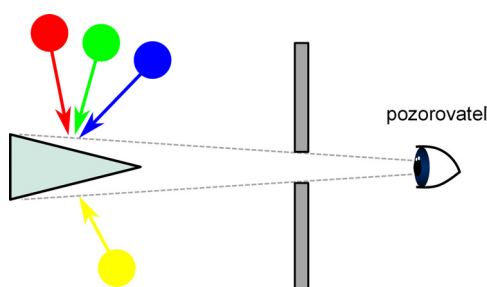
- definícia parametrov štandardného pozorovateľa,
- definícia štandardných iluminantov,
- definícia priestorov farieb a
- diferenčné formuly farby.

Až v roku 1931 bol na základe biologických aspektov vnímania farieb navrhnutý prvý matematicky presný farebný model – CIE XYZ.

Od roku 1993 nadväzuje na prácu CIE Medzinárodné konzorcium pre farbu (International Color Consortium, ICC), a to metodikou farebného manažmentu (color management).

19.1 Grassmanov experiment

V roku 1853 Hermann G. Grassmann urobil nasledujúci experiment. Pozorovateľovi ukázal dva panely, jeden osvetlený nejakou farbou a druhý tromi primárnymi svetelnými zdrojmi červenou (R), zelenou (G) a modrou (B). Pozorovateľ mal za úlohu meniť intenzity troch primárnych zdrojov svetla tak, aby farba vzniknutá ich zmiešaním bola okom nerozoznateľná od danej farby. Experiment je znázornený na obrázku 19.1.



Obrázok 19.1: Grassmanov experiment porovnávania farieb. Pozorovateľ sleduje cez malý otvor panel rozdelný na dve časti, kde dolná časť je osvetlená testovou farbou a horná časť primárnymi svetelnými zdrojmi, ktoré sú zmiešavané aditívne.

Tento experiment predpokladá **metamerizmus farieb**, čo znamená, že rovnaký farebný podnet môže vzniknúť pôsobením farebných podnetov rôzneho spektrálneho zloženia. Takéto podnety potom nazývame, **metamerné farebné podnety**.

V uskutočnenom experimente zistili, že nie pre všetky zvolené farby vieme nájsť vhodné intenzity primárnych zdrojov tak, aby sme úspešne našli „rovnaký“ farebný pár. Preto presunuli červenú farbu na panel spolu s hľadanou farbou. A tak bolo zistené, že pre určitú časť vlnových dĺžok viditeľného spektra je hodnota koeficientov záporná, obzvlášť pre červenú farbu.

Podobné experimenty sa neskôr stali základom kolorimetrie.

19.2 Grassmanov zákon aditívneho miešania farieb

Grassman na základe experimentu formuloval **princíp aditívneho miešania farieb** (*additive mixture*):

„Celková intenzita výsledného zmiešaného svetla sa rovná sume intenzít jednotlivých zmiešavaných svetiel.“ [Gra53]

Ľubovoľnú farbu môžeme vyjadriť ako lineárnu kombináciu troch farieb, pričom žiadna z týchto farieb nesmie byť lineárnou kombináciou zostávajúcich dvoch. Toto sa nazýva **1. Grassmanov zákon miešania farieb** (*Grassman's assumption of additivity*, alebo aj *Grassman's Law*).

Grassmannov predpoklad aditivity je základom systému kolometrie. Napriek tomu, že v posledných desaťročiach bolo opakovane zistené zlyhanie aditivity [VV11; Oic07; WS00; SB59], jeho platnosť je vo všeobecnosti axiomaticky akceptovaná. Cena za zmenu celého CMS (color management system) by bola príliš vysoká.

19.3 Kolorimetrický priestor

Farba Q môže byť namiešaná pomocou R_Q jednotiek červenej, G_Q jednotiek zelenej a B_Q jednotiek modrej farby.

$$Q = R_Q \cdot \mathbf{R} + G_Q \cdot \mathbf{G} + B_Q \cdot \mathbf{B}, \quad (19.1)$$

kde Q je daný **farebný podnet** (*color stimulus*), $\mathbf{R}$, $\mathbf{G}$, $\mathbf{B}$ sú **primárne farebné podnety** (*primary stimuli of unit amounts*) a R_Q , G_Q , B_Q **trichromatické zložky** (*tristimulus values*). Táto rovnica sa nazýva **kolorimetrická rovnica**.

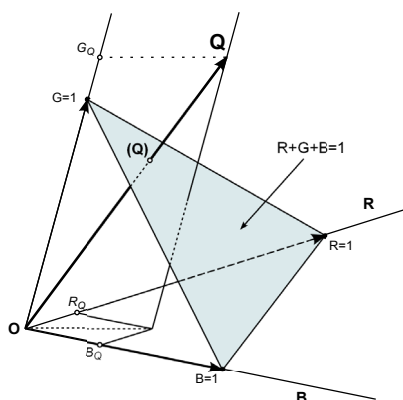
Každý farebný podnet môžeme teda reprezentovať vektorom v 3-dimenzionálnom priestore nazývanom **kolorimetrický priestor** (*tristimulus space*). Osi v tomto priestore sú tvorené farebnými podnetmi $\mathbf{R}$, $\mathbf{G}$ a $\mathbf{B}$. Kolorimetrický priestor $(\mathbf{R}, \mathbf{G}, \mathbf{B})$ spolu s farebným podnetom Q je znázornený na obrázku 19.2.

Vyjadrenie v 3D priestore nie je v praxi vždy vyhovujúce. Lepšie sa Q reprezentuje v 2D priestore. Takáto reprezentácia môže byť napríklad pomocou jednotkovej roviny kde $R + G + B = 1$. Ak hodnoty R , G , B nor-

malizujeme dostaneme chromatické súradnice

$$\begin{aligned} r &= \frac{R}{R+G+B}, \\ g &= \frac{G}{R+G+B}, \\ b &= \frac{B}{R+G+B}. \end{aligned} \quad (19.2)$$

Takáto jednotková rovina tvorí trojuholník – **diagram chromatickosti**, ktorý nazývame aj **Maxwellov farebný trojuholník** (*Maxwell color triangle*).



Obrázok 19.2: (R, G, B) – kolorimetrický priestor. Primárne podnety **R**, **G**, **B** sú reprezentované ako vektory jednotkovej dĺžky so spoločným začiatkom v bode **O**. Farebný podnet **Q** je reprezentovaný ako trichromatický vektor **Q** s komponentami dĺžky R_Q , G_Q , B_Q . Trichromatický vektor pretína jednotkovú rovinu ($R + G + B = 1$) v bode **Q**. Jednotková rovina obsahuje (r, g, b) – chromatický diagram, v ktorom **Q** reprezentuje chromatický bod **Q** [WS00].

Ak testovaný farebný podnet je monochromatický, výsledkom je množina **spektrálnych trichromatických zložiek** (*spectral tristimulus values*). Výpočet (19.1) môžeme urobiť postupne pre všetky spektrálne monochromatické podnety s vlnovou dĺžkou λ . Potom môžeme zapísať

$$Q(\lambda) = R(\lambda) \cdot \mathbf{R} + G(\lambda) \cdot \mathbf{G} + B(\lambda) \cdot \mathbf{B}, \quad (19.3)$$

kde $R(\lambda)$, $G(\lambda)$ a $B(\lambda)$ je množina spektrálnych trichromatických zložiek nameraná vzhľadom na primárne farebné podnety **R**, **G** and **B**.

Nech $\mathbf{Q}$ dostaneme aditívnym miešaním monochromatických podnetov $Q(\lambda)$ v rozsahu farebného spektra 380 – 780 nm s delením stupnice vlnových dĺžok na 1 nm, potom

$$\mathbf{Q} = \sum_{\lambda=380}^{780} Q(\lambda). \quad (19.4)$$

Obzvlášť dôležitá je množina spektrálnych trichromatických zložiek získaná pre všetky monochromatické podnety $Q(\lambda)$ daného farebného stimulu $\mathbf{Q}$, ktoré majú jednotkovú spektrálnu hustotu žiarenia, v každej vlnovej dĺžke λ v celom viditeľnom spektre. Takýto podnet sa nazýva **podnet s izoenergetickým spektrom** a označuje sa $\mathbf{E}$. Trichromatické zložky takejto kolorimetrickej sústavy sa vyjadrujú **trichromatickými členiteľmi**. Kolorimetrická rovnica pre farebný podnet izoenergetického spektra $E(\lambda)$ je nasledová

$$E(\lambda) = \bar{r}(\lambda)\mathbf{R} + \bar{g}(\lambda)\mathbf{G} + \bar{b}(\lambda)\mathbf{B}, \quad (19.5)$$

kde $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, $\bar{b}(\lambda)$ je množina trichromatických členiteľov nameraná vzhľadom na primárne merné podnety $R(\lambda)$, $G(\lambda)$, $B(\lambda)$ a zápis pomocou malého písmenka s čiarkou nad písmenom je štandardné označenie pre vyjadrenie pre špeciálny prípad izoenergetického spektra. Súhrn trichromatických členiteľov pre všetky vlnové dĺžky vytvára **spektrálne krivky trichromatických členiteľov** (*color matching functions – CMF*).

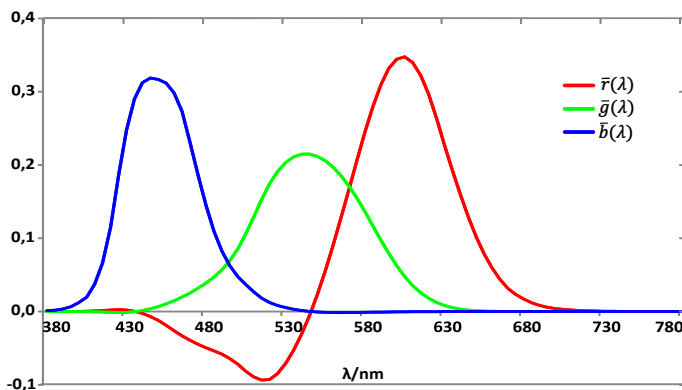
Z obrázka 19.3 je zrejmé, že pre niektoré hodnoty λ potrebujeme záporné množstvo červeného farebného podnetu $\mathbf{R}$. Pre hodnotu $\lambda = 475$ sú kolorimetrické hodnoty farebných podnetov $r(\lambda) = -0,045$, $g(\lambda) = 0,032$ a $b(\lambda) = 0,186$. Takže môžeme zapísať

$$E_{\lambda} = -0,045\mathbf{R} + 0,032\mathbf{G} + 0,186\mathbf{B}. \quad (19.6)$$

Záporné množstvo červeného farebného podnetu $\mathbf{R}$ znamená, že pri porovnávaní farieb potrebujeme pridať $0,045\mathbf{R}$ k E_{λ} , aby sme ju „desaturovali“.

Trichromatické zložky každého podnetu $Q(\lambda)$ vzhľadom na primárne farebné podnety $R(\lambda)$, $G(\lambda)$ a $B(\lambda)$ sa dajú získať zo spektrálnych kriviek trichromatických členiteľov nasledovne:

$$\begin{aligned} R(\lambda) &= \bar{r}(\lambda)Q(\lambda), \\ G(\lambda) &= \bar{g}(\lambda)Q(\lambda), \\ B(\lambda) &= \bar{b}(\lambda)Q(\lambda). \end{aligned} \quad (19.7)$$



Obrázok 19.3: Spektrálne krivky trichromatických členiteľov $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, $\bar{b}(\lambda)$ monochromatického podnetu $E(\lambda)$ v systéme reálnych primárnych podnetov **R**, **G**, **B** s vlnovými dĺžkami $\lambda_R = 700$, $\lambda_G = 546,1$ a $\lambda_B = 435,8$.

Následne z (19.3), (19.4) a (19.7) dostaneme

$$\mathbf{Q} = \left(\sum_{\lambda=380}^{780} \bar{r}(\lambda) Q(\lambda) \right) \mathbf{R} + \left(\sum_{\lambda=380}^{780} \bar{g}(\lambda) Q(\lambda) \right) \mathbf{G} + \left(\sum_{\lambda=380}^{780} \bar{b}(\lambda) Q(\lambda) \right) \mathbf{B} \quad (19.8)$$

Z rovnice vyplýva, že ak máme danú **krivku pomerného spektrálneho zloženia** (*spectral power distribution function – SPD*) nejakého podnetu **Q** a množinu spektrálnych kriviek trichromatických členiteľov nameranú vzhľadom na primárne podnety **R**, **G** a **B**, trichromatické zložky podnetu **Q** vzhľadom na **R**, **G** a **B** môžeme vypočítať ako jednotlivé sumy každej zložky v celom rozsahu vlnových dĺžok.

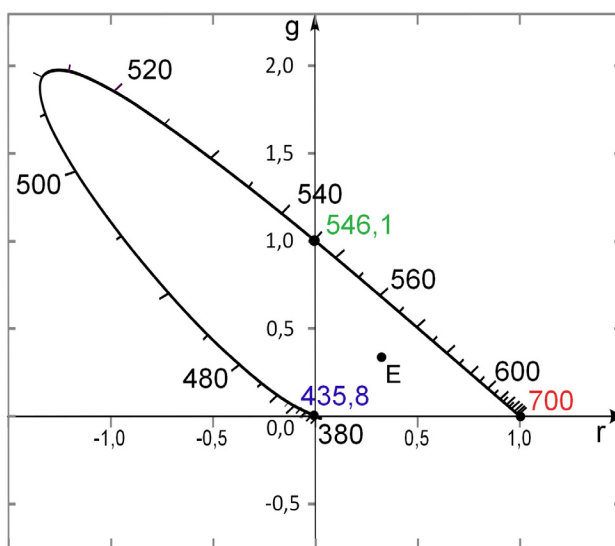
Dva podnety Q_1 a Q_2 s rôznymi SPD vyvolajú rovnaký farebný podnet, tj. sú metamérmi, ak platia nasledujúce rovnice

$$\begin{aligned} \sum_{\lambda=380}^{780} \bar{r}(\lambda) Q_1(\lambda) &= \sum_{\lambda=380}^{780} \bar{r}(\lambda) Q_2(\lambda), \\ \sum_{\lambda=380}^{780} \bar{g}(\lambda) Q_1(\lambda) &= \sum_{\lambda=380}^{780} \bar{g}(\lambda) Q_2(\lambda), \\ \sum_{\lambda=380}^{780} \bar{b}(\lambda) Q_1(\lambda) &= \sum_{\lambda=380}^{780} \bar{b}(\lambda) Q_2(\lambda). \end{aligned} \quad (19.9)$$

Ak aplikujeme rovnice (19.2) na spektrálne trichromatické hodnoty, dostaneme chromatické súradnice

$$\begin{aligned} r(\lambda) &= \frac{\bar{r}(\lambda)}{\bar{r}(\lambda) + \bar{g}(\lambda) + \bar{b}(\lambda)}, \\ g(\lambda) &= \frac{\bar{g}(\lambda)}{\bar{r}(\lambda) + \bar{g}(\lambda) + \bar{b}(\lambda)}, \\ b(\lambda) &= \frac{\bar{b}(\lambda)}{\bar{r}(\lambda) + \bar{g}(\lambda) + \bar{b}(\lambda)}, \end{aligned} \quad (19.10)$$

kde $r(\lambda) + g(\lambda) + b(\lambda) = 1$.



Obrázok 19.4: Chromatický diagram v systéme chromatických súradníc r , g , s vyznačenými vlnovými dĺžkami $\lambda_R = 700$, $\lambda_G = 546,1$ a $\lambda_B = 435,8$ primárnych podnetov **R**, **G**, **B**.

Na obrázku 19.4 je znázornený **rg-diagram chromatickosti**. Tvorí ho pravouhlý trojuholník a spektrálna krivka (locus) monochromatických podnetov od 380 nm do 780 nm. Priamka spájajúca dva konce tejto krivky sa nazýva **priamka sýtych purpurov** (*purple line*). Chromatické body **R**,

G, B monochromatických primárnych podnetov **R**, **G**, **B** sú reprezentované vrcholmi trojuholníka. Chromatický bod E podnetu s izoenergetickým spektrom **E** sa nachádza v strede trojuholníka so súradnicami $r_E = g_E = 1/3$.

Trichromatické zložky podnetu **Q** so spojitým pomerným spektrálnym zložením žiarenia (radiant power distribution) P_λ môžeme zapísať v tvare

$$\begin{aligned} R &= \int_{\lambda=380}^{780} P_\lambda \bar{r}(\lambda) d\lambda, \\ G &= \int_{\lambda=380}^{780} P_\lambda \bar{g}(\lambda) d\lambda, \\ B &= \int_{\lambda=380}^{780} P_\lambda \bar{b}(\lambda) d\lambda. \end{aligned} \quad (19.11)$$

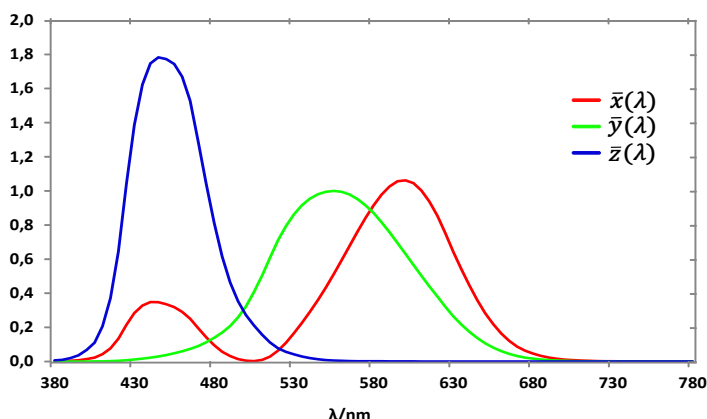
19.4 Štandardný kolorimetrický pozorovateľ

Ako je zrejmé z predchádzajúcej kapitoly, medzi ľuďmi s normálnym farebným videním sú stále významné individuálne rozdiely, preto je potrebné definovať nejaký štandard.

CIE ustanovila dvoch štandardných kolorimetrických pozorovateľov. Štandardný kolorimetrický pozorovateľ je množina spektrálnych kriviek trichromatických členiteľov, ktoré reprezentujú vlastnosti trichromatických členiteľov ľudskej populácie s normálnym farebným videním.

19.4.1 CIE 1931 štandardný kolorimetrický pozorovateľ

Na základe experimentov Guilda a Wrighta [FBH97] definovala v roku 1931 Medzinárodná komisia pre osvetľovanie krivky trichromatických členiteľov $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, a $\bar{b}(\lambda)$, pre vlnové dĺžky od $\lambda = 380$ po 780 nm, s deľením stupnice vlnových dĺžok $\Delta\lambda = 5$ nm a tak ustanovila **CIE 1931 štandardného kolorimetrického pozorovateľa** (*standard colorimetric observer*). Tieto krivky platia pre pozorovanie malých farebných plôch s pozorovacím uhlom 2° . Za základ kolorimetrickej sústavy RGB prijala trojicu merných podnetov s vlnovými dĺžkami 700 nm, $546,1$ nm a $435,8$ nm. Táto sústava je prispôbena tak, aby chromatické súradnice podnetu s izoenergetickým spektrom boli rovnaké. Činitele jasu v tomto systéme sú v pomere $1 : 4,5907 : 0,0601$, čo zodpovedá vyžiarenej energii v pomere $72,0962 : 1,3791 : 1$. Kolorimetrickú sústavu RGB môžeme charakterizo-



Obrázok 19.5: Spektrálne krivky trichromatických členiteľov $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$ podľa CIE 1931 štandardného kolorimetrického pozorovateľa pre monochromatický podnet $E(\lambda)$ v systéme imaginárnych primárnych podnetov **X**, **Y**, **Z**.

vať hodnotami trichromatických členiteľov $\bar{r}(\lambda)$, $\bar{g}(\lambda)$ a $\bar{b}(\lambda)$, ktoré platia pre normálneho kolorimetrického pozorovateľa.

V roku 1971 odporučila CIE hodnoty štandardného kolorimetrického pozorovateľa pre rozšírený interval 360nm – 830nm a s delením stupnice vlnových dĺžok $\Delta\lambda = 1\text{nm}$.

19.4.2 CIE 1964 doplnkový štandardný kolorimetrický pozorovateľ

V praxi sa môžeme stretnúť s prípadom, že viditeľné žiarenie, ktoré vyvolá farebný podnet, vniká do oka pod väčším uhlom ako 2° . Preto bol v roku 1964 ustanovený **CIE 1964 doplnkový štandardný kolorimetrický pozorovateľ** (*supplementary standard colorimetric observer*) pre veľké pozorovacie pole s uhlom 10° a trojica merných podnetov s vlnovými dĺžkami $R = 645,2$ $G = 526,3$ $B = 444,4$ nm.

Podrobný opis vývoja a vlastností CIE štandardného kolorimetrického pozorovateľa je v časti 3.3.3 knihy Colour Science autorov Wyszecki a Stiles [WS00].

Najnovšie špecifikácie pre spektrálne krivky trichromatických členiteľov sú publikované na webstránke CIE¹.

¹<http://www.cie.co.at>

19.5 Kolorimetrická sústava CIE XYZ

Aby všetky viditeľné farby mohli byť definované iba pomocou koeficientov s nezápornými hodnotami, pôvodné spektrálne krivky trichromatických členiteľov boli modifikované na celo-pozitívny systém. V roku 1931 CIE zaviedla tri **imaginárne farebné podnety** (*imaginary primary stimuli*) **X**, **Y** a **Z** a systém CIE XYZ (1931) sa stal základom kolorimetrie. Tieto podnety nie sú fyzicky realizovateľné, existujú iba ako matematická konštrukcia. Výsledné krivky reprezentujúce kolorimetrické členiteľa priemerného pozorovateľa nameraná s ohľadom na imaginárne podnety **X**, **Y**, **Z** sú označené $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$ (obr. 19.5). Kolorimetrická sústava CIE XYZ bola navrhnutá tak, aby členiteľ $\bar{y}(\lambda)$ mal rovnaký priebeh ako krivka citlivosti ľudského oka pri dennom videní.

Projekciou 3-D farebného priestoru do roviny $X + Y + Z = 1$ dostaneme (x, y)-chromatický diagram (obr. 19.6). Projekcia je definovaná ako

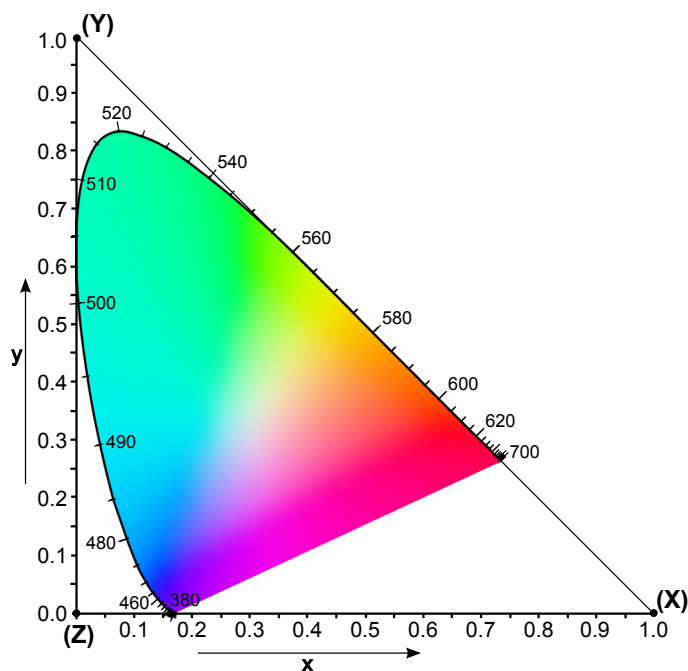
$$\begin{aligned}x &= \frac{X}{X + Y + Z}, \\y &= \frac{Y}{X + Y + Z}, \\z &= \frac{Z}{X + Y + Z} = 1 - x - y.\end{aligned}\tag{19.12}$$

Farba v tomto priestore je určená dvomi súradnicami x, y. Tretia súradnica, z, môže byť tiež definovaná, ale vďaka transformácii je redundantná.

Nové primárne podnety **X**, **Y**, **Z** ležia zvonku mimo domény vektorov reprezentujúcich reálne podnety. Ich reprezentatívne body v diagrame chromatickosti nie sú zahrnuté medzi reálnymi podnetmi. Ležia zvonka chromatického regiónu ohraničeného spektrálnou krivkou a čiarou tých purpurov, ktorá spája dva konce spektrálnej krivky. Trojuholník vytvorený z bodov XYZ úplne ohraničuje spektrálnu krivku a čiaru purpurov. Toto zaisťuje, že chromatické súradnice xyz a korešpondujúce trichromatické hodnoty XYZ akéhokoľvek farebného podnetu nebudú negatívne.

Rovina XZ zobrazuje v kolorimetrickom priestore XYZ farebné podnety s nulovým jasom. Pretína diagram chromatickosti v priamke, ktorá sa nazýva **alychna**. V kolorimetrickej sústave CIE XYZ 1931 diagram chromatickosti nie je rovinou rovnakých jasov.

Hodnoty trichromatických členiteľov, ktoré platia pre štandardného kolorimetrického pozorovateľa, sú použité na stanovenie trichromatických zložiek podnetu so známym spektrálnym zložením žiarenia. Pri ur-



Obrázok 19.6: CIE 1931 chromatický diagram so súradnicovým systémom xy a vyznačenými bodmi kadiel' prechádzajú osi súradnicového systému XYZ . Diagram je ohraničený spektrálnou krivkou s označeniami vlnových dĺžok v nanometroch.

čovaní trichromatických zložiek X , Y , Z rozlišujeme výpočet pre **farebný podnet svetelného zdroja** a **pre farebný podnet objektu**.

Nech Q je farebný podnet svetelného zdroja s krivkou pomerného spektrálneho zloženia $Q(\lambda)$, potom vzťahy pre stanovenie trichromatických zložiek X , Y , Z pre podnet Q sú:

$$\begin{aligned} X &= k \sum_{\lambda=380}^{780} Q(\lambda) \bar{x}(\lambda) \Delta\lambda, \\ Y &= k \sum_{\lambda=380}^{780} Q(\lambda) \bar{y}(\lambda) \Delta\lambda, \\ Z &= k \sum_{\lambda=380}^{780} Q(\lambda) \bar{z}(\lambda) \Delta\lambda, \end{aligned}$$

(19.13)

kde k je normalizačný faktor zvyčajne nastavený na 683 lm/W, preto hodnota Y môže byť použitá priamo na určenie jasu L (*luminance*).

V prípade, keď je žiarenie odrazené od objektu je $Q(\lambda)$ súčinom spektra iluminanta $S(\lambda)$ a spektrálneho činiteľa $\beta(\lambda)$. Spektrálny činiteľ $\beta(\lambda)$ ovplyvňuje výsledné spektrum v dôsledku odrazu alebo prechodu svetla telesom. Vzťahy pre stanovenie trichromatických zložiek X , Y , Z pre farebný podnet objektu dostaneme

$$\begin{aligned} X &= k \sum_{\lambda=380}^{780} S(\lambda)\beta(\lambda)\bar{x}(\lambda)\Delta\lambda, \\ Y &= k \sum_{\lambda=380}^{780} S(\lambda)\beta(\lambda)\bar{y}(\lambda)\Delta\lambda, \\ Z &= k \sum_{\lambda=380}^{780} S(\lambda)\beta(\lambda)\bar{z}(\lambda)\Delta\lambda. \end{aligned} \quad (19.14)$$

CIE vytvorila tento priestor farieb na základe imaginárnych farebných podnetov, a preto nie je žiadne zariadenie schopné zobrazit' tieto farby. Využíva sa zväčša na účely konverzie farieb medzi dvomi rôznymi priestormi. Farebné podnety XYZ reprezentujú abstraktné farby.

CIE 1931 RGB and CIE 1931 XYZ sú prvé matematicky definované priestory farieb.

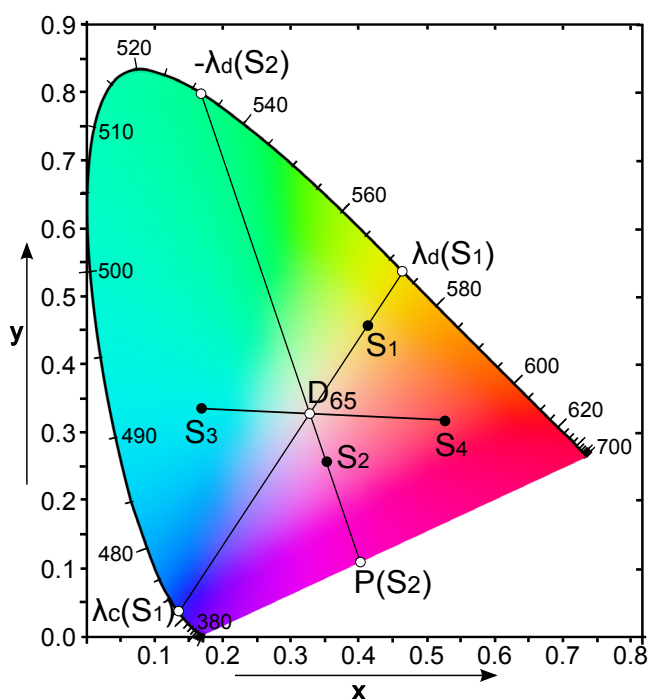
Môžeme pracovať aj s alternatívnym priestorom daným zložkami x , y , Y a označovaným ako **CIE xyY** . Chromatické súradnice x , y špecifikujú odtieň farby a saturáciu a Y zložka reprezentuje jas. Tento priestor je často používaný v praxi na určenie farieb.

19.6 CIE diagram chromatickosti

CIE diagram chromatickosti je diagram v tvare podkovy, po obvode ohraničený krivkou reprezentujúcou čisté monochromatické farby od 380 nm do 780 nm. Priamka spájajúca dva konce tejto krivky sa nazýva **priamka sýtych purpurov** (*purple line*). Bielu farbu dostaneme zložením všetkých farieb a nachádza sa v ťažisku. Na obrázku 19.7 je zobrazený reprezentant bielej farby CIE štandardný iluminant D_{65} . V diagrame chromatickosti sa nachádzajú iba farby s maximálnym jasom. Preto napríklad, hneď farba nie je v CIE diagrame, avšak je to oranžovo-červená so zníženým jasom.

V CIE diagrame určujeme chromatickosť farby pomocou dominantnej vlnovej dĺžky (farebný tón/odtieň) a čistoty farby (sýtosť farby). Sý-

tosť farby, čiže jej saturácia, sa zvyšuje od stredu oblasti bielych farieb smerom k okraju diagramu. Na okraji sa nachádza spektrálna krivka s čistými farbami. Farebný odtieň sa mení pohybom po krivke. Chromatický bod E podnetu s izoenergetickým spektrom E sa nachádza v strede so súradnicami $x_E = y_E = 1/3$.



Obrázok 19.7: CIE 1931 chromatický diagram zobrazujúci vlastnosti dominantných a doplnkových vlnových dĺžok. Dominantná vlnová dĺžka farebného podnetu S_1 , určená vzhľadom k CIE štandardnému iluminantu D_{65} je $\lambda_d(S_1)$. K nej doplnková vlnová dĺžka je $\lambda_c(S_1)$. Farebný podnet S_2 nemá dominantnú vlnovú dĺžku, je určený doplnkovou vlnovou dĺžkou $-\lambda_d(S_2)$. Farby S_3 a S_2 nazývame komplementárne.

Dominantná a doplnková vlnová dĺžka

Dominantná vlnová dĺžka (*dominant wavelength*) farebného podnetu je vlnová dĺžka monochromatického spektrálneho podnetu, ktorý keď zmiešame s nejakým achromatickým podnetom v správnom pomere, zís-

kame požadovaný farebný podnet. V CIE diagrame nájdeme dominantnú vlnovú dĺžku, ak urobíme spojnicu od bieleho svetla cez uvažovaný farebný podnet. Achromatické podnety sa väčšinou udávajú ako CIE štandardný iluminant A alebo D65, ktoré majú neutrálnu farbu a chýba im odtieň farby. Farebný podnet S_1 zobrazený na obrázku 19.7 má dominantnú vlnovú dĺžku, určenú vzhľadom k CIE štandardnému iluminantu D_{65} , $\lambda_d(S_1)$.

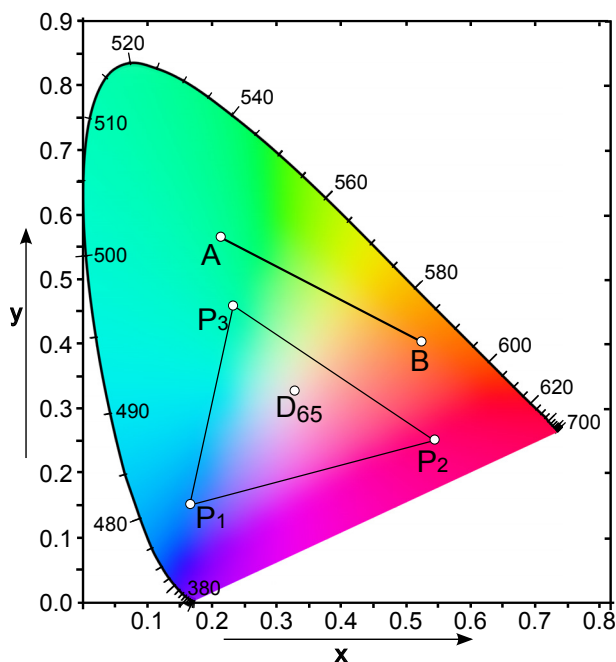
Nie všetky farebné podnety majú dominantnú vlnovú dĺžku. Pre purpurové farby ju nemožno určiť, pre tieto farby definujeme **doplňkovú vlnovú dĺžku** (*complementary wavelength*). Ak túto zmiešame s uvažovaným farebným podnetom v správnom pomere, dostaneme achromatické svetlo. Táto vlnová dĺžka sa zapisuje so záporným znamienkom. Farebný podnet S_2 na obrázku 19.7 nemá dominantnú vlnovú dĺžku. Môžeme ho však definovať, vzhľadom k CIE štandardnému iluminantu D_{65} , pomocou doplnkovej vlnovej dĺžky $-\lambda_d(S_2)$.

Čistota farby

Čistota farby je definovaná v diagrame chromatickosti ako pomer vzdialeností, ktorý udáva posun chromatického bodu daného podnetu od achromatického bodu smerom k spektrálnej krivke alebo k priamke sýtych purpurov. Čistota farby súvisí so saturáciou farby vnímanou za určitých pozorovacích podmienok. Pre bod chromatický S_1 na obrázku 19.7 určíme čistotu farby ako pomer $|D_{65}S_1| : |D_{65}\lambda_d(S_1)|$. Pre určenie čistoty farby bodu S_2 však potrebujeme najprv nájsť bod $P(S_2)$, ktorý leží na priesečníku priamky sýtych purpurov a spojnice bodov S_2 a D_{65} . Čistota farby bodu S_2 je potom pomer vzdialeností $|D_{65}S_2|$ a $|D_{65}P(S_2)|$.

Komplementárne farby

Páry farieb, ktoré ak zmiešame v správnom pomere dávajú achromatické svetlo, sa nazývajú **komplementárne farby**. Na obrázku 19.7 sú komplementárne farby S_3 a S_4 . Významné sú monochromatické komplementárne farby. Komplementárne sú aj farby dané vlnovými dĺžkami $\lambda_d(S_1)$ a $\lambda_c(S_1)$. Pre farby s vlnovými dĺžkami od 493,3 nm po 566,5 nm neexistujú monochromatické komplementárne farby, ich komplementárne farby ležia v oblasti purpurových farieb. Bod na spektrálnej krivke označený ako $-\lambda_d(S_2)$ má komplementárnu farbu $P(S_2)$ ležiacu na priamke purpurov. Achromatické svetlo je väčšinou definované pomocou CIE štandardných iluminantov, ktoré sa nachádzajú približne na Planckovej krivke.



Obrázok 19.8: CIE 1931 chromatický diagram zobrazujúci pákové pravidlo. Farebné podnety na úsečke AB získame aditívnym zložením farieb A a B. Farebné podnety vnútri trojuholníka $P_1P_2P_3$ získame kombináciou základných farieb P_1 , P_2 a P_3 .

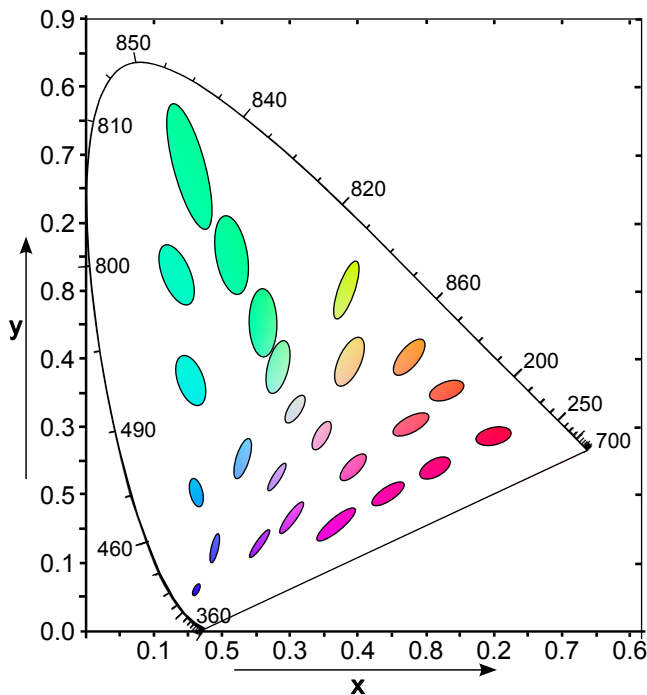
Štandardné biele svetlo (iluminant C) sa nachádza blízko bodu $x = y = z = 1/3$.

Pákové pravidlo

Ak máme v diagrame chromatickosti dva body A a B, potom všetky body úsečky AB, tj. farebné prechody z farby A do B získame aditívnym zložením týchto dvoch farieb (obr. 19.8). Toto nazývame **pákové pravidlo**. Podobne aj v prípade trojuholníka $P_1P_2P_3$, kde vnútorné body trojuholníka získame kombináciou základných farieb P_1 , P_2 a P_3 . Pri pohľade na CIE diagram vidno, že nevieme zostrojiť trojuholník, ktorý by pokrýval celú plochu. Z toho vyplýva, že nie je možné nájsť také 3 základné farby, ktoré by určovali všetky farby. Farebný rozsah, ktorý dostaneme zložením zvolených základných farieb sa nazýva **gamut farieb**. Farebný gamut

môže byť udaný rôznymi polygónmi umiestnenými v diagrame chromatickosti.

19.7 MacAdamove elipsy



Obrázok 19.9: MacAdamove elipsy zobrazené v CIE 1931 chromatickom diagrame, 10 krát zväčšené oproti svojej skutočnej veľkosti.

Chromatický diagram je značne nelineárny. Vzdialenosti v CIE diagrame chromatickosti nezodpovedajú vnímanému rozdielu medzi farbami. Ak zmiešame farby s rovnakou úrovňou jasu v rovnakom pomere, výsledná farba nebude vždy ležať presne v strede medzi týmito dvomi farbami. V 40. rokoch 20. storočia sa David MacAdam zaoberal vizuálnou citlivosťou na rozdiel farieb a jeho závery dnes poznáme pod názvom MacAdamove elipsy. MacAdamove elipsy sú regióny v diagrame chromatickosti, ktoré obsahujú farby ľudským okom nerozlíšiteľné. Teda na

kontúry elipsy sa nachádzajú sotva rozlíšiteľné farby (*just noticeable differences*).

MacAdamove elipsy sú získané štatistickými meraniami vizuálnych vlastností vzorky populácie. Tvar a veľkosť elíps závisí od umiestnenia v CIE diagrame. Na obrázku 19.9 je zobrazených všetkých 25 MacAdamových elíps v pomere 10 : 1 v porovnaní so skutočnými veľkosťami. Vizuálna citlivosť na malé rozdiely vo farbe je dôležitý faktor na určenie presnosti pri porovnávaní farieb najmä v priemysle. Na základe práce MacAdama boli vytvorené nové farebné priestory s cieľom dosiahnuť perceptuálnu jednotnosť. Najvýznačnejšie sú CIELUV a CIELAB, ktoré sú opísané v častiach 20.1.4 a 20.1.5. Aj keď oproti CIE 1931 boli dosiahnuté zlepšenia, MacAdamove elipsy sú takmer kruhové, žiaden z týchto priestorov nie je úplne bez distorzie.

19.8 Štandardné iluminanty

Pri určovaní farby objektov je dôležité svetlo, ktoré tento objekt osvetľuje. Aby sme mohli farbu presne merať, potrebujeme poznať spektrálne zloženie tohto svetla. Medzinárodná komisia pre osvetľovanie definovala množinu štandardných iluminantov používaných v kolorimetrii.

Pod pojmom **iluminant** rozumieme prirodzený svetelný zdroj, ktorý je určený pomerným spektrálnym zložením žiarenia dopadajúceho na objekt, ktorý je pozorovaný pozorovateľom.

CIE štandardný iluminant A

CIE štandardný iluminant A (obr. 19.10) reprezentuje žiarovkové svetlo s farebnou teplotou 2856 K. Bol definovaný CIE v roku 1931.

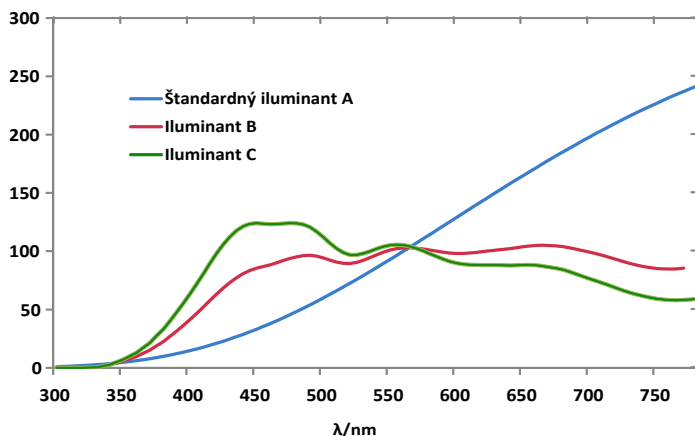
CIE iluminanty B a C

V roku 1931 spolu so štandardným iluminantom A boli definované aj iluminanty B a C (obr. 19.10).

Iluminant B reprezentuje priame slnečné svetlo s korelovanou farebnou teplotou 4874 K.

Iluminant C reprezentuje priemerné denné svetlo s korelovanou farebnou teplotou 6774 K.

Obidva sú však považované za nevhodné pre reprezentovanie prirodzeného denného svetla kvôli ich nepresnosti v oblasti ultrafialového spektra. V súčasnosti sa nepoužívajú a podľa odporúčania CIE z roku 2004 [CIE04], iluminant C nie je štandardom CIE.



Obrázok 19.10: Relatívne spektrálne krivky CIE iluminantov A, B a C.

CIE iluminanty D

Iluminanty D tvoria hlavnú skupinu iluminantov zodpovedajúcich rôznym typom denného svetla (priame slnečné svetlo, slnečné svetlo s rôznymi stupňami oblačnosti, svetlo za jasného dňa, zatiahnutá obloha) (obr. 19.11).

CIE štandardný iluminant D65 reprezentuje denné svetlo v nulovej nadmorskej výške s farbou teploty približne 6504K. Bol odporučený CIE v roku 1964 ako hlavný iluminant pre aproximáciu priemerného denného svetla pre kolorimetriu. CIE ho odporúča použiť vždy, keď je to možné. V prípade ak sa použiť nedá, treba brať do úvahy D50 (5000 K), D55 (5503 K) alebo D75 (7504 K).

D50 svetlo na obzore (*Horizon Light*)

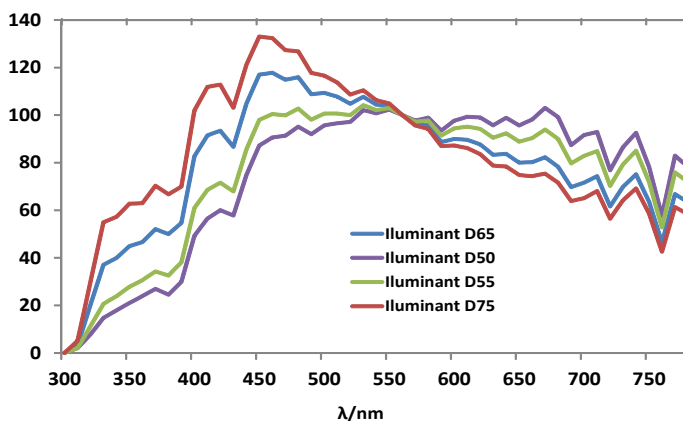
D55 ranné/poobedňajšie svetlo (*Mid-morning / Mid-afternoon Daylight*)

D65 svetlo na poludnie (*Noon Daylight*)

D75 svetlo zo severnej oblohy na severnej pologuli (*North Sky Daylight*)

CIE iluminant E

Iluminant E je teoretický iluminant s izoenergetickým spektrom, čiže jeho pomerné spektrálne zloženie je uniformné cez viditeľné spektrum.



Obrázok 19.11: Relatívne spektrálne krivky CIE iluminantov skupiny D.

Illuminant E nezodpovedá žiareniu čierneho telesa, a preto sa nevyjadruje vo farebnej teplote.

CIE iluminanty F

Illuminanty F je séria zodpovedajúca rôznym fluorescenčným lampám s označeniami F1 až F12.

19.8.1 Štandardné zdroje svetla

CIE rozlišuje medzi iluminantami a zdrojmi svetla. Kým iluminant je daný špecifickým spektrálnym zložením žiarenia dopadajúceho na pozorovaný objekt, pod zdrojom svetla rozumieme **fyzické zariadenie vyžarujúce svetlo**, ako lampa alebo slnko. Iluminanty sú dôležité pre kolorimetrické výpočty, a preto potrebujeme mať aj ich fyzikálne realizácie.

Illuminant nie je vždy presne realizovaný svetelným zdrojom. CIE odporúča nasledovné štandardné zdroje pre aproximáciu iluminantov.

Štandardný zdroj A

Štandardný zdroj A je plynom plnená žiarovka s volfrámovým vláknom (500 W) s farebnou teplotou 2854 K.

Štandardné zdroje D, simulátory denného svetla

Pre iluminanty skupiny D CIE neodporúča žiadne umelé svetelné zdroje, kvôli ich zložitému spektrálnemu zloženiu. CIE však odporúča simulátory denného svetla pre účely kolorimetrie. Kvalita simulátora môže byť meraná pomocou CIE indexu metamerizmu. Podľa knihy [WS00] je vhodným simulátorom iluminantu D65 napríklad xenónová výbojka s filtrom (filtered high-pressure xenon arc lamp – Xe5).

Štandardné zdroje F

V prípade, že treba vybrať iba zopár typických fluorescenčných lúč, CIE odporúča používať prioritne iluminanty F2, F7 a F11.

CIE v technickej správe CIE15:2004, uvádza aj novú množinu fluorescenčných lúč s označením FL3.1 – FL3.15, kde FL3.15 je odporúčaná ako simulátor iluminantu D65.

Modely a priestory farieb

20.1 Homogénne priestory farieb

V priestore XYZ aj v CIE diagrame nezodpovedajú lineárne vzdialenosti na rôznych miestach priestoru subjektívne vnímaným rozdielom vo farbe a naopak. Takáto nehomogénnosť je problémom pre aplikácie, kde je potrebné merať rozdiel farieb. Vývoj formuly pre výpočet rozdielu dvoch farieb (*color difference formulae*), ktorá bola nevyhnutná najmä pre priemysel, bol problematický. Nájsť veličinu, ktorá by uspokojivo opísala rozdiel vnímaný pozorovateľom medzi dvoma danými farebnými podnetmi, spočíva v schopnosti pozorovateľa posúdiť relatívnu veľkosť farebného rozdielu týchto dvoch farebných podnetov. Hodnotenie pozorovateľa sa mení v závislosti od pozorovacích podmienok. Preto boli pre rôzne experimentálne podmienky vyvinuté rôzne formuly.

Boli snahy vyvinúť farebný priestor, ktorý bude vnemovo rovnomerný, čo znamená, že rozdiel farieb medzi dvoma bodmi (dvomi farbami) jednotkovej vzdialenosti v takomto priestore je rovnako vnímaný pre všetky farby.

Tieto vlastnosti sa nepodarilo dosiahnuť v žiadnom z fyzicky a technicky orientovaných priestorov. Treba si ale uvedomiť, že neexistuje farebný priestor, ktorý by nebol skreslený. Preto, ak hovoríme o „homogénnom“ priestore, myslíme „viac-menej“ homogénny priestor. Homogénne priestory farieb sú zaujímavé v rôznych oblastiach počítačového videnia, obzvlášť, ak porovnávame podobné farby. Tieto modely sú nezávislé od

zariadenia. Avšak transformácie do iných priestorov farieb sú výpočtovo veľmi náročné. Toto môže byť problémom najmä vtedy, keď spracovávame video v reálnom čase.

Homogénne priestory farieb majú využitie najmä v tých aplikáciách počítačového videnia, v ktorých sa snažíme priblížiť k ľudskému spôsobu vnímania farieb, ako je výber obrazov z databázy (*image retrieval*), alebo ohodnotenie kvality obrazu (*image quality assessment*). Nie je výhodné použiť ich tam, kde nie je priame spojenie medzi ľudským vizuálnym systémom ako v prípade sterea, alebo sledovania pohybu (*motion tracking*) kvôli veľkej časovej náročnosti pri nelineárnych transformáciách z RGB.

Bolo vyvinutých viacero homogénnych priestorov farieb, v ktorých sa chceli ich autori priblížiť k homogénnosti. Prvý pokus urobil v roku 1935 D. B. Judd. V nasledujúcom texte uvedieme tie priestory, ktoré boli prijaté komisiou CIE. Najvýznamnejšie z nich sú priestory CIELAB a CIELUV. Farebný priestor CIELAB sa používa pri opise farieb matných materiálov a povrchov, kým CIELUV je obvykle vhodný pre opis farby svetla.

20.1.1 CIE 1960 Yuv

V roku 1937 navrhol David MacAdam farebný priestor Yuv, ktorý CIE prijala v roku 1960 ako štandardný homogénny farebný diagram. V literatúre sa môžeme stretnúť s označeniami CIEYUV (1960), CIE–uv (1960) alebo **CIE UCS 1960** (*uniform chromaticity-scale*) diagram. Nové súradnice u a v sú dosiahnuté pomocou lineárnej transformácie

$$\begin{aligned} u &= \frac{2x}{6y - x + 1,5}, \\ v &= \frac{3y}{6y - x + 1,5}, \end{aligned} \quad (20.1)$$

kde x, y sú z priestoru Yxy . Transformácia 20.1 má zabezpečiť, že v novovzniknutom chromatickom diagrame bude vektor jednotkovej veľkosti vnímaný pre všetky farby rovnako. Rozdielová nerovnomernosť je značne, ale nie dostatočne, redukovaná. Tretia súradnica je redundantná.

20.1.2 CIE 1964 $U^*V^*W^*$

Günther Wyszecki na základe CIE 1960 UCS vytvoril farebný priestor CIE 1964 ($U^*V^*W^*$), tiež označovaný skratkou CIEUVW, v ktorom bola nelinearita oproti CIE 1960 UCS lepšie redukovaná. Neskôr ho nahradil CIE

1976 $L^*u^*v^*$. Wyszeckeho priestor je definovaný nasledovne

$$\begin{aligned} U^* &= 13W^*(u - u_0), \\ V^* &= 13W^*(v - v_0), \\ W^* &= 25Y^{(1/3)} - 17, \quad (1 \leq Y \leq 100), \end{aligned} \quad (20.2)$$

kde u, v sa rovná

$$\begin{aligned} u &= \frac{4X}{X + 15Y + 3Z} = \frac{4x}{-2x + 12y + 3}, \\ v &= \frac{6Y}{X + 15Y + 3Z} = \frac{6y}{-2x + 12y + 3}, \end{aligned}$$

a u_0, v_0 sú hodnoty premenných u, v , pre achromatickú farbu umiestnenú v začiatku (U^*, V^*) – systému. Pre farby objektov je postačujúce, ak výber u_0, v_0 zodpovedá iluminantu. X, Y, Z sú trichromatické zložky a x, y sú chromatické súradnice.

20.1.3 CIE 1976 $Y_u'v'$

Na základe neskorších výskumov CIE v roku 1976 modifikovala UCS 1960. Išlo o lineárnu transformáciu priestoru Yuv , kde Y a u zostáva nezmenené a v' súradnica je 1,5 násobkom v . V **CIE 1976 UCS** je rozdielová nelinearita v porovnaní s UCS 1960 mierne redukovaná, ale stále nie dost'. Hodnoty u', v' môžu byť definované pomocou trichromatických zložiek X, Y, Z alebo pomocou chromatických súradníc x, y nasledovne

$$\begin{aligned} u' &= u = \frac{4X}{X + 15Y + 3Z} = \frac{4x}{-2x + 12y + 3}, \\ v' &= 1,5v = \frac{9Y}{X + 15Y + 3Z} = \frac{9y}{-2x + 12y + 3}. \end{aligned} \quad (20.3)$$

20.1.4 CIE 1976 $L^*u^*v^*$

CIE 1976 $L^*u^*v^*$ je prvý takmer homogénny priestor farieb odporúčený CIE, určený pre farebné svetelné zdroje. Je vylepšenou verziou CIE 1964 $U^*V^*W^*$. Je to nelineárny farebný priestor, ale konverzie sú vratné. De-

finovaný je nasledovne

$$\begin{aligned}
 L^* &= \begin{cases} 116 \left(\frac{Y}{Y_n} \right)^{1/3} - 16, & \text{ak } \frac{Y}{Y_n} > 0,008856, \\ 903,3 \left(\frac{Y}{Y_n} \right), & \text{ak } \frac{Y}{Y_n} \leq 0,008856, \end{cases} \\
 u^* &= 13L^*(u' - u'_n), \\
 v^* &= 13L^*(v' - v'_n),
 \end{aligned} \tag{20.4}$$

kde hodnoty u' , v' sú definované podľa 20.3 a hodnoty u'_n , v'_n dostaneme, ak do 20.3 dosadíme za X , Y , Z trichromatické hodnoty X_n , Y_n , Z_n , opisujúce biely referenčný bod. Obvykle sú dané spektrálnym zložením žiarenia jedného zo štandardných CIE iluminantov D65 alebo A. Pri týchto podmienkach sú X_n , Y_n , Z_n trichromatické zložky štandardného iluminanta s Y_n rovným 100. Krivka pre L^* je pre hodnoty menšie ako 0,008856 lineárna. Podľa štandardizácie CIE je najlepšia aproximácia tejto krivky funkcia $x^{0,4}$. Nelineárny vzťah s Y je použitý na priblíženie sa k logaritmickej odozve oka. Hodnoty L^* sa pohybujú v rozmedzí 0 až 100 a relatívny jas Y/Y_n v rozmedzí 0 až 1.

V priestore $L^* u^* v^*$ je pre konštantné L^* zapracovaný (u', v') -chromatický diagram, ktorý je projektívnou transformáciou CIE 1931 (x, y) -chromatického diagramu. Priamky v (x, y) -chromatickom diagrame zostanú po transformácii do (u', v') -diagramu zachované. Toto je dôležitá vlastnosť v prípade, že farebné podnety sú zmiešavané aditívne. Keďže pozícia bodov v (u^*, v^*) -diagrame závisí od hodnoty L^* , nie sú v jednoznačnom vzťahu k chromaticite. Avšak pre konštantné L , priamky z (x, y) -chromatického diagramu alebo z (u', v') -chromatického diagramu zostanú priamkami aj v (u^*, v^*) -diagrame.

Rozdiel farieb

Nech $[L_1, u_1, v_1]$ a $[L_2, u_2, v_2]$ sú dve rôzne farby z $L^* u^* v^*$ priestoru. Potom rozdiel týchto dvoch farieb je jednoducho Euklidovská vzdialenosť

$$\Delta E_{uv}^* = \sqrt{(L_1 - L_2)^2 + (u_1 - u_2)^2 + (v_1 - v_2)^2}. \tag{20.5}$$

Rozdiel ΔE_{uv}^* menší ako 1 je nepostrehnuteľný, rozdiely v rozmedzí 1 až 4 môžu a nemusia byť postrehnuteľné v závislosti od toho, z ktorej časti farebného priestoru pochádzajú. Rozdiel väčší ako 4 je postrehnuteľný.

Polárne súradnice

V praxi je často výhodnejšie použiť polárne súradnice, ktoré lepšie vystihujú vnímanie ľudského oka, ako karteziánske. Sú vyjadrené prostredníctvom oddelených komponentov ako vnímaná svetlosť (lightness L^*), chroma (C^*) a tón farby (hue h_{uv}). Niekedy sa používa aj psychometrické nasýtenie (saturation s_{uv}), čo je vlastne pomer chromy a svetlosti. Aby tento priestor bol viac priaznivý pre používateľa, CIE definovala dva analogické priestory $CIELCH_{uv}$ a $CIELHS_{uv}$. Chroma, tón farby a saturáciu pre $L^* u^* v^*$ vypočítame nasledovne

$$\begin{aligned} C^* &= \left(u^{*2} + v^{*2} \right)^{0,5}, \\ h_{uv} &= \arctan \left(\frac{v^*}{u^*} \right), \\ s_{uv} &= \frac{C^*}{L^*}. \end{aligned} \tag{20.6}$$

Výhodou takejto reprezentácie je, že je oddelená jasová zložka od farebnej, čo môže byť výhodou napríklad pri segmentácii objektov, kde nechceme, aby jas (osvetlenie, tieňovanie) ovplyvňoval výsledok.

20.1.5 CIE 1976 $L^*a^*b^*$

V roku 1976 bol vytvorený aj model CIE $L^* a^* b^*$, ktorý je určený pre farebné objekty s difúznym povrchom. Takýto priestor prvýkrát navrhol Hunter v roku 1948 a ide o ďalšiu transformáciu kolorimetrického priestoru CIEXYZ. V súčasnosti pod skratkou Lab rozumieme priestor CIE 1976 $L^* a^* b^*$ označovaný aj skratkou CIELAB, ktorý je podobný Hunterovmu Lab priestoru. Rozdiel je iba v spôsobe, ako vypočítame koeficienty.

CIE $L^* a^* b^*$ je oponentný trojdimenzionálny farebný priestor, ktorý sa skladá z jasovej zložky L^* a dvoch farebných zložiek a^* a b^* . Hodnoty jasovej zložky L^* sa pohybujú od 0 (čierna) po 100 (biela). Farebné zložky tvoria dve osi, pričom záporné hodnoty osi a^* zodpovedajú zelenej farbe a oproti tomu kladné hodnoty komplementárnej fialovej farbe. Na osi b^* zodpovedajú záporné hodnoty modrej a kladné k nej komplementárnej žltej farbe.

$$\begin{aligned}
L^* &= \begin{cases} 116 \left(\frac{Y}{Y_n} \right)^{1/3} - 16, & \text{ak } \frac{Y}{Y_n} > 0,01, \\ 903,3 \left(\frac{Y}{Y_n} \right), & \text{ak } \frac{Y}{Y_n} \leq 0,008856, \end{cases} \\
a^* &= 500 \left[f \left(\frac{X}{X_n} \right) - f \left(\frac{Y}{Y_n} \right) \right], \\
b^* &= 200 \left[f \left(\frac{Y}{Y_n} \right) - f \left(\frac{Z}{Z_n} \right) \right],
\end{aligned}
\tag{20.7}$$

kde

$$f(t) = \begin{cases} t^{1/3}, & \text{ak } \frac{Y}{Y_n} > 0,01, \\ 7,787t + \frac{16}{116}, & \text{ak } \frac{Y}{Y_n} \leq 0,008856. \end{cases}$$

X_n, Y_n, Z_n sú trichromatické hodnoty opisujúce biely referenčný bod, ktorý reprezentuje perfektne matné biele teleso osvetlené CIE štandardným iluminantom D65, A alebo C. Pri týchto podmienkach sú X_n, Y_n, Z_n trichromatické zložky štandardného iluminanta s Y_n rovným 100. Treba si všimnúť vo vzorci tretiu odmocninu, ktorá bola zavedená z dôvodu získania homogénneho izotropného farebného telesa (*isotropic color solid*).

L^* má škálu rozdelenú približne rovnomerne podľa vnímanej sensitivity. Treba podotknúť, že rozdiely medzi rovnako vzdialenými bodmi na osi Y sú viac postrehnuteľné v tmavších častiach než vo svetlých.

L^* je identická s L^* v CIELUV, ale pre a^* a b^* neexistuje jednoduchý vzťah k zložkám u^* a v^* z CIELUV. Keďže pozícia a^* a b^* závisí od hodnoty L^* , nie je možné ich zakresliť do roviny, a preto na rozdiel od CIELUV, **CIELAB nemá chromatický diagram**. V priestore CIELAB priamky z (x, y)-chromatického diagramu nie sú namapované na priamky ako v CIELUV, ale ich obrazy sú zakrivené.

Polárne súradnice

Podobne ako pre CIELUV, CIE definovala analogický priestor CIELCH vyjadrený pomocou polárnych súradníc, ktoré môžeme vypočítať nasledovne:

$$\begin{aligned}
C^* &= \left(a^{*2} + b^{*2} \right)^{0,5}, \\
h_{ab} &= \arctan \left(\frac{a^*}{b^*} \right).
\end{aligned}
\tag{20.8}$$

Keďže CIELAB nemá chromatický diagram, saturácia nie je v tomto priestore definovaná.

Rozdiel farieb

Formula pre rozdiel farieb je definovaná ako Euklidovská vzdialenosť v 3D priestore CIELAB

$$\Delta E_{ab}^* = \sqrt{\Delta L^{*2} + \Delta a^{*2} + \Delta b^{*2}}. \quad (20.9)$$

20.1.6 CIE CAM02

CIE sa snažila zakomponovať do svojich modelov čoraz viac fenoménov vzhľad farby, aby lepšie modelovali ľudské vnímanie farieb. Ako výsledok takéhoto snaženia vznikol v roku 2002 model CIECAM02.

Model vzhľadu farby (Color appearance model – CAM) modeluje spôsob, akým ľudský vizuálny systém vníma farby objektov pri rôznych svetelných podmienkach a s rôznym pozadím. Farby v obraze, ktorý sledujeme, môžu byť v takomto modeli prispôsobené, aby vyzerali rovnako pri rôznych svetelných podmienkach. Toto umožňuje vyvinúť spôsob ukladania obrazov nezávislý od zariadenia. Napríklad, ak je obraz naskenovaný do počítača a sú známe osvetľovacie podmienky, pomocou CAM ho môžeme transformovať do internej reprezentácie a následne, ak má byť tento obraz zobrazený na monitore alebo vytlačený pri známych svetelných podmienkach, môže byť transformovaný z jeho internej reprezentácie do takej, aby bol vizuálne rovnaký ako pôvodný – skenovaný.

CIECAM02 je momentálne najzložitejší používaný farebný model. Berie do úvahy osvetľujúci zdroj, osvetlenie okolia, pozadie, farebnú škálu a mnoho ďalších vplyvov. Jeho predchodcom bol CIECAM97.

20.2 Hardvérovo orientované modely

V nasledujúcom texte budeme používať pojmy **farebný priestor** a **farebný model**. Preto si najprv vysvetlíme rozdiel medzi nimi. Farebný model je abstraktný matematický model, opisujúci spôsob, ako môžeme farby reprezentovať pomocou n-tíc čísiel. Obvykle sú to tri alebo štyri hodnoty alebo farebné komponenty. Ak je takýto model asociovaný s presným predpisom, ako sú jednotlivé komponenty interpretované, hovoríme o farebnom priestore. Napríklad AdobeRGB a sRGB sú dva rôzne farebné priestory, obidva založené na RGB modeli.

Niektoré farebné priestory, ako ProphotoRGB, obsahujú všetky farby, ktoré sú v Adobe RGB a Adobe RGB obsahuje takmer všetky farby z sRGB. Vo všeobecnosti, ak konvertujeme jeden farebný priestor do druhého, môžeme dostať farby, ktoré sú mimo cieľového priestoru. Takéto farby nemôžu byť reprezentované v novom priestore farieb presne, označujeme ich „farby mimo gamut“.

20.2.1 Lineárny model RGB

Základnými (primárnymi) farbami modelu RGB sú

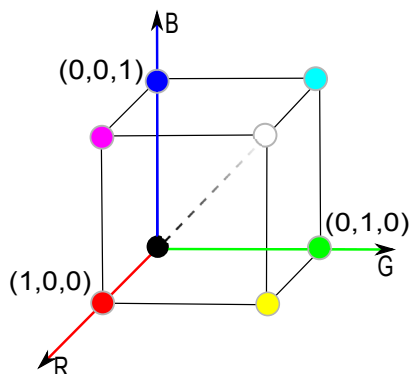
R – (Red) červená,
G – (Green) zelená,
B – (Blue) modrá.

Ostatné farby sa vyjadrujú aditívnym skladaním týchto farieb. Vyjadrujeme ich pomocou váhového súčtu jednotlivých zložiek. Môžeme to prirovnať k rozsvecovaniu farebných svetiel. Čím majú zložky jednotlivých farieb väčšiu hodnotu, tým je výsledná farba svetlejšia. Skladaním primárnych farieb tvoríme takzvané sekundárne farby. V tomto modeli sú to modrozelená (cyan), fialová (magenta) a žltá (yellow).

Model RGB je reprezentovaný jednotkovou kockou umiestnenou v začiatku súradnicovej sústavy Euklidovského priestoru, E^3 , v jeho prvom oktante (obr. 20.1). Jednotlivé osi súradnicového systému reprezentujú množstvo príslušnej farebnej zložky vo výslednej farbe. Intenzita základných farieb sa v tomto modeli pohybuje v intervale $(0, 1)$. Z toho vyplýva, že množina základných farieb vo vrchoch kocky predstavuje 8 farieb, ktoré tvoria primárne, sekundárne farby, biela a čierna. Vrchol $[0, 0, 0]$ (tj. stred súradnicového systému) zodpovedá čiernej farbe (black). Naproti tomu vrchol $[1, 1, 1]$ zodpovedá bielej farbe (white). Farby ležiace na diagonále medzi týmito vrcholmi zodpovedajú odtieňom šedej (gray).

Pri technickej implementácii je rozsah najčastejšie kódovaný pomocou 8 bitov (tj. 256 dielov). Intenzitu každého obrazového bodu môžeme teda opísať 24 bitmi, čo predstavuje presnosť resp. farebnú hĺbku „true color“. V RGB farebnom modeli môžeme vytvoriť $256^3 = 16,8$ milióna farieb. Pre ľudské oko by postačovalo aj delenie na 100 dielov namiesto 256. Pri praktickom nastavovaní sa preto používa percentuálne nastavenie jednotlivých zložiek.

Daný model je v porovnaní s inými modelmi najviac technicky orientovaný. Tento model sa používa pre farebné monitory. Všetky monitory a televízory tvoria farby tak, že na obrazovke na každý bod svetia tri lúče: červený, zelený a modrý. Avšak RGB model nie je veľmi efektívny pri



Obrázok 20.1: Model RGB reprezentovaný jednotkovou kockou. V osiach sa nachádzajú primárne farby R, G, B a v počiatočnom bode $[0, 0, 0]$ sa nachádza čierna farba.

práci s reálnymi obrazmi. Medzi zložkami v RGB modeli je veľká korelácia a pri akejkoľvek modifikácii obrazu treba spravidla modifikovať všetky tri zložky.

20.2.2 Model RGBA

V tomto prípade je k červenému, zelenému a modrému kanálu pridaný ďalší, tzv. **alfa-kanál**. Ten uchováva informáciu o **priehľadnosti** (*transparency*). Alfa-kanál má väčšinou rovnaký rozsah ako ostatné tri kanály a určuje, v akom pomere bude daná farba zmiešaná s farbou na pozadí obrazu

- Hodnota 0 znamená úplne priehľadný objekt,
- Hodnoty z intervalu $(0, 1)$ znamenajú čiastočne priesvitný objekt,
- Hodnota 1 znamená úplne nepriesvitný objekt.

20.2.3 Nelineárne priestory RGB

Aby sme mohli jednoznačne definovať farebný priestor, potrebujeme chromatické súradnice troch primárnych farieb a referenčnú bielu farbu. V praxi bolo vyvinutých množstvo RGB priestorov s rôznymi primárnymi farbami a referenčnou bielou, aj keď väčšinou ich hodnoty nie sú prístupné pre používateľa. Pri manipulácii s multimédiami používateľa väčšinou predpokladajú, že všetky obrazy sú reprezentované rovnakým RGB

priestorom, avšak ak nemáme k dispozícii chromatické súradnice primárnych farieb, akékoľvek manipulovanie s obrazom nie je presné.

Gama korekcia

Rôzne zariadenia majú odlišné fotosenzory. Aby sme dostali presnú reprodukciu farby na rôznych zariadeniach, pridáva sa tzv. **gama korekcia**. Gama korekcia slúži na kompenzáciu nelinearity televíznej obrazovky, kde napríklad trojnásobne silnejšiemu napätiu nezodpovedá $3\times$ jasnejší svit daného miesta obrazovky. Pôvodné lineárne signály sa označujú RGB, po gama korekcii R'G'B'. Pre rôzne zariadenia môže byť použitá rôzna hodnota gama.

Aj keď by bolo možné zvoliť tri primárne farby tak, aby pokryli čo najväčšiu plochu v CIE xy-diagrame, sú väčšinou volené tak, aby luminofory boli vyrobiteľné a energetické nároky neboli príliš vysoké. Zakreslením bodov reprezentujúcich jednotlivé primárne farby do CIE xy-chromatického diagramu dostaneme paletu farieb zobraziteľných na danom zariadení. Toto sa nazýva **gamut** zariadenia.

Každý RGB farebný priestor je definovaný pre špecifické zariadenie so svojimi RGB a gama hodnotami. Referenčná biela je farba referenčného bieleho podnetu. Ako referenčná biela sa väčšinou používa jeden z CIE štandardných iluminantov (D65, C, D50...). Bolo navrhnutých veľa rozličných množín kriviek trichromatických členiteľov, aby sme mohli správne vizualizovať farbu na displejoch, čo má za následok rôzne RGB štandardy, napr. NTSC-RGB, PAL-RGB a sRGB. Proces prispôsobenia hodnôt farieb pre rôzne zariadenia (napr. kameraRGB, monitorRGB, printerRGB) má na starosti systém správy farieb (CMS).

Farebný priestor sRGB

Farebný priestor **sRGB** má primárne farebné podnety definované v štandarde ITU-R BT.709 (tab. 20.1).

Na rozdiel od iných RGB priestorov, gama pre sRGB sa nedá vyjadriť jedným číslom. Gama korekcia pozostáva z lineárnej časti (gama 1,0) v tmavých hodnotách a nelineárnej časti na zvyšku intervalu. Hodnoty gama prechádzajú od 1,0 až po 2,4. Celková približná gama hodnota je 2,2.

Transformácia z CIE XYZ na sRGB je nasledovná

$$\begin{bmatrix} R_{sRGB} \\ G_{sRGB} \\ B_{sRGB} \end{bmatrix} = \begin{bmatrix} 3,2410 & -1,5374 & -0,4986 \\ -0,9692 & 1,8760 & 0,0416 \\ 0,0556 & -0,2040 & 1,0570 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}. \quad (20.10)$$

Tabuľka 20.1: Chromatické súradnice primárnych podnetov R, G, B priestoru sRGB dané vo farebnom priestore xyY.

Chromatické súradnice	Červená R	Zelená G	Modrá B	Biely bod D_{65}
x	0,6400	0,3000	0,1500	0,3127
y	0,3300	0,6000	0,0600	0,3290
Y	0,2126	0,7153	0,0721	1,0000

sRGB je odporúčaný farebný priestor pre Internet a od roku 2007 je najviac používaným RGB priestorom. Môže byť použitý pre editovanie a ukladanie obrazov určených na publikovanie na webe, avšak jeho farebná škála je limitovaná. Pokrýva približne 36 % viditeľných farieb. Obrazy určené na profesionálnu tlač môžu preto stratiť na kvalite. V tom prípade je lepšie použiť iný farebný priestor, napr. Adobe RGB, ktorý má inú farebnú škálu.

Farebný priestor Adobe RGB

Farebný priestor **Adobe RGB** vyvinula firma Adobe Systems v roku 1998 tak, aby obsahoval väčšinu farieb dosiahnuteľných v CMYK tlačiarňach pomocou reprezentácie hodnôt RGB na monitore. Tento priestor obsahuje približne 52 % viditeľných farieb, a oproti sRGB má rozsiahlejší gamut hlavne v oblasti fialovo – zelených farieb.

Farebný priestor Adobe Wide Gamut RGB

Spoločnosť Adobe System navrhla aj ďalší RGB priestor **Adobe Wide Gamut RGB**. Tento pokrýva až 77,6 % viditeľných farieb vďaka čisto spektrálnym primárnym podnetom, ktorých chromatické súradnice spolu s bielym bodom D_{50} sú v tabuľke 20.2.

20.2.4 Model CMY

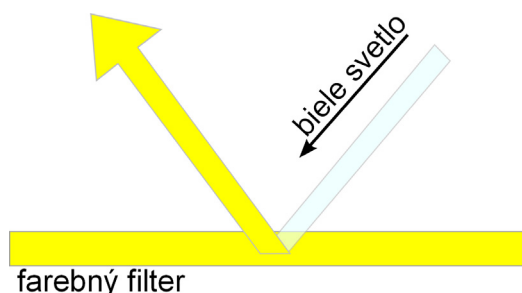
Základnými farbami modelu CMY sú

- C** – (Cyan) modrozelená, tyrkysová,
- M** – (Magenta) fialová, purpurová,
- Y** – (Yellow) žltá.

Tabuľka 20.2: Chromatické súradnice primárnych podnetov R, G, B priestoru Adobe Wide Gamut RGB.

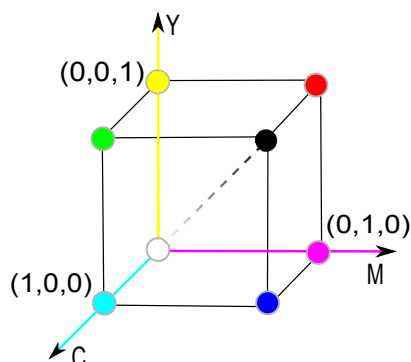
Chromatické súradnice	Červená R	Zelená G	Modrá B	Biely bod D_{50}
CIE x	0,7347	0,1152	0,1566	0,3457
CIE y	0,2653	0,8264	0,0177	0,3585
Vlnové dĺžky	700 nm	525 nm	450 nm	

Model CMY vychádza z maliarskeho miešania farieb, pri ktorom pridávaním hodnôt jednotlivých farebných zložiek dostávame výslednú farbu tmavších odtieňov, ide o tzv. subtraktívne miešanie farieb. Môžeme si to predstaviť ako nanášanie filtrov, kde napr. žltý pri dopade bieleho svetla všetky vlnové dĺžky pohltí, odrazí iba žltú (obrázok 20.2).



Obrázok 20.2: Princíp maliarskeho miešania farieb. Z bieleho svetla sa po dopade na žltý filter odrazia iba vlnové dĺžky zodpovedajúce žltej farbe, všetky zvyšné vlnové dĺžky pohltí filter.

Farebný model CMY sa najčastejšie reprezentuje ako jednotková koc-ka umiestnená v osiach C, M, Y (obr. 20.3). Intenzita základných farieb sa v tomto modeli pohybuje v intervale $\langle 0, 1 \rangle$. Vo vrcholoch kocky sa nachádzajú tzv. základné farby. Vrchol $[0, 0, 0]$ zodpovedá bielej farbe a vrchol $[1, 1, 1]$ čiernej farbe. Farby ležiace na diagonále medzi týmito vrcholmi zodpovedajú odtieňom šedej. Ich nárast je v opačnom smere ako v modeli RGB.



Obrázok 20.3: Model CMY reprezentovaný jednotkovou kockou. V osiach sa nachádzajú sekundárne farby C, M, Y a v počiatočnom bode $[0,0,0]$ sa nachádza biela farba.

Transformácie medzi modelom RGB a modelom CMY

Transformácia RGB na CMY je daná

$$\begin{aligned} \text{Cyan} &= 1 - \text{Red}, \\ \text{Magenta} &= 1 - \text{Green}, \\ \text{Yellow} &= 1 - \text{Blue}. \end{aligned} \quad (20.11)$$

Inverzná transformácia CMY na RGB je nasledovná

$$\begin{aligned} \text{Red} &= 1 - \text{Cyan}, \\ \text{Green} &= 1 - \text{Magenta}, \\ \text{Blue} &= 1 - \text{Yellow}. \end{aligned} \quad (20.12)$$

20.2.5 Model CMYK

Subtraktívne miešanie je vhodné na tlačiarenské účely, hlavne pri atramentových tlačiarnach, kde sa uplatňuje model CMYK (K – black or key), v ktorom je pridaná čierna farba. Pri tlači sa výsledná farba v skutočnosti nemieša, ale postupne sa nanášajú jednotlivé farby v podobe vzorov natočených pod rôznym uhlom. Takýto postup ale nie je dostatočne presný a pri tlačení čiernej farby dostaneme skôr tmavohnedú, resp. tmavošedú. Preto sa pridáva zvlášť čierna farba, ktorá je väčšinou v samostatnom toneri a tlačí sa samostatne. Ďalší dôvod je ekonomický, pri tvorbe čiernej sa mňajú všetky tri farebné tonery, čo je finančne náročné. Vytlačený do-

kument alebo obraz odráža svetlo, ktoré ho osvetľuje. Niektoré svetlo je pohltené, zvyšné odrazené.

Priestor CMYK je taktiež neintuitívny pre používateľov a takisto ako iné uvedené modely farieb nelineárny na vnímanie.

Farebné priestory RGB a CMYK sú závislé od zariadenia, na ktorom sa používajú. CMYK taktiež závisí od atramentu a od použitého papiera. S rozličnými zariadeniami dostaneme rôzne prenosové charakteristiky medzi špecifikovanými CMY(K) hodnotami a množstvom farby na papieri. Konverzie sú dané systémom správy farieb (CMS), ktoré využívajú profily farieb. Napriek tomu konverzie nemôžu byť úplne presné, pretože sú tu veľmi odlišné gamuty.

Transformácie medzi CMY a CMYK

Na prevod farieb medzi CMY a CMYK je často uvádzaná nasledovná veľmi jednoduchá transformácia

Transformácia CMY na CMYK je daná

$$\begin{aligned} K &= \min(C, M, Y), \\ C &= (C - K) / (1 - K), \\ M &= (M - K) / (1 - K), \\ Y &= (Y - K) / (1 - K). \end{aligned} \quad (20.13)$$

Transformácia CMYK na CMY je nasledovná

$$\begin{aligned} C &= \min(1, C(1 - K) + K), \\ M &= \min(1, M(1 - K) + K), \\ Y &= \min(1, Y(1 - K) + K). \end{aligned} \quad (20.14)$$

Táto nenáročná transformácia je skôr ilustratívna, ak sa použije na tlač, výsledná reprodukcia farieb bude veľmi zlá. Vhodnejšie je vypočítať konverziu z RGB do CMY(K) cez trichromatické zložky XYZ.

20.3 Modely farieb v televíznej a videotechnike

Modely farieb opísané v tejto časti sú asi najmenej známe zo všetkých uvedených v tejto kapitole. Používajú sa najmä tam, kde sa spracovanie

obrazu a počítačové videnie stretáva s televíznou alebo videotechnikou.

Po rozšírení čiernobieleho televízneho signálu na farebný bolo potrebné nájsť spôsob, ako tento signál komprimovať, pretože prenos farebného signálu v podobe RGB by vyžadoval trikrát väčšiu kapacitu čiernobieleho signálu. S využitím vlastností ľudského oka, ktoré je oveľa citlivejšie na zmeny v jase ako vo farbe bola oddelená jasová zložka od zložiek vyjadrujúcim farebu. Zároveň bolo treba zachovať spätnú kompatibilitu, aby sa mohli naďalej používať aj čiernobiely prijímače. Pridaný farebný signál nesmel viditeľne ovplyvňovať kvalitu čiernobieleho signálu a zároveň tento musel byť jednoducho vyfiltrovateľný. V rôznych častiach sveta vznikli viaceré systémy na kódovanie farebného signálu a každý riešil spätnú kompatibilitu na čiernobiely signál po svojom.

Spoločnou vlastnosťou všetkých modelov je ich oddelená jasová zložka (**luma**) od chromatickej. Keďže veľké percento jasu je tvorené zelenou primárnou farbou, chromatické zložky sú založené na zvyšných dvoch primárnych farbách červenej a modrej. Prvá chromatická zložka je vytvorená odpočítaním hodnoty lummy od nelineárnej modrej ($B'-Y'$) a druhá odpočítaním lummy z nelineárnej červenej ($R'-Y'$). Tieto zložky sa nazývajú **chroma**. Zložky modelov sú počítané z nelineárnych gama korigovaných hodnôt $R'G'B'$ a preto, vždy keď je použitá gama korekcia, používa sa apostrof väčšinou pri jasovej zložke, napríklad $Y'CbCr$.

V rôznych aplikáciách sa používajú rozdielne škálovacie faktory na ($B'-Y'$) a ($R'-Y'$). Škálovacie faktory $Y'PbPr$ sú optimalizované pre analógové video. V modeli $Y'CbCr$ je škálovanie vhodné pre digitálne video ako štúdiové video, JPEG a MPEG. Kodak Photo YCC používa škálovacie faktory optimalizované pre gamut farebného filmu. $Y'UV$ je škálovanie vhodné pre medzikrok vo formovaní kompozitných NTSC a PAL video signálov.

Vo všeobecnosti môžeme zapísať

$$\begin{aligned} Y &= k_r R + k_g G + k_b B, \\ Cb &= B - Y, \\ Cr &= R - Y, \end{aligned} \tag{20.15}$$

kde k_r , k_g a k_b sú vhodné koeficienty a $k_r + k_g + k_b = 1$.

Polárne súradnice

Chromatické komponenty ($B'-Y'$) a ($R'-Y'$) môžeme vyjadriť pomocou polárnych súradníc, aby sme lepšie reprezentovali perceptuálne atribúty

ako odtieň farby a saturáciu. Hodnoty môžeme vypočítať nasledovne:

$$\begin{aligned} H &= \tan^{-1} \left(\frac{B' - Y'}{R' - Y'} \right), \\ S &= \sqrt{(B' - Y')^2 + (R' - Y')^2}. \end{aligned} \quad (20.16)$$

Tieto modely sú vnemovo nehomogénne, takže nie je vhodné používať na meranie rozdielu medzi dvomi farbami euklidovskú vzdialenosť.

20.3.1 Model YIQ

Farebný model YIQ bol zavedený americkou komisiou NTSC (National Television Standard Committee) v roku 1953. Je tu oddelená jasová zložka od farebných. Y zložka (luma) zodpovedá vnímanému jasú a je to zložka, ktorú využívajú čiernobiele televízne prijímače, kým zložky I a Q kódujú chromatickú informáciu. Vďaka oddelenej jasovej zložke je možná manipulácia s jasom bez ovplyvnenia farebnej zložky, napríklad pri ekvalizácii histogramu.

Zložka I znamená *in-phase* (**činná zložka**) a nesie informáciu o tóne farby v rozmedzí medzi oranžovou a modrozelenou. Zložka Q znamená *quadrature* (**kvadrant**), a nesie informáciu o tóne farby medzi zelenou a purpurovou. Chromatické zložky spoločne kódujú odtieň farby aj saturáciu.

Tento model berie do úvahy fakt, že ľudské oko je výrazne citlivejšie na zmeny v jase ako v odtieni farby alebo saturácii, a preto sa používa väčšia šírka pásma pre Y ako pre I a Q. Väčšia šírka pásma pre zložku Y umožňuje aj lepšie vysporiadanie sa s vizuálnymi chybami spôsobenými kvantizáciou počas prenosu alebo spracovania. Zložka Y je vytvorená tak, že gama korekcia sa najskôr aplikuje na lineárny model RGB a potom sa vypočíta váhovaný súčet nelineárnych komponentov R'G'B'. Výsledná zložka vyjadruje jas, nie je však rovnaká ako CIE Y, napriek tomu, že použité označenie je rovnaké.

Presné škálovanie farebných osí bolo definované tak, aby zodpovedalo spektrám žiarenia luminoforov používaným v televíznych obrazovkách v rôznych častiach sveta. Gama korekcia používaná NTSC pre CRT je 2,2. Biely bod systému je štandardný iluminant C. Koeficienty, ktoré zodpovedajú NTSC červenému, zelenému a modrému luminoforu CRT obrazovky, sú štandardizované v ITU-R Rec. BT. 601. Luma je vypočítaná z nelineárnych komponentov R'G'B' nasledovne

$$Y'_{601} = 0,299R' + 0,587G' + 0,114B'. \quad (20.17)$$

Transformácie medzi RGB a YIQ

Transformácia z R'G'B' do Y'IQ je daná

$$\begin{bmatrix} Y' \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0,299 & 0,587 & 0,114 \\ 0,596 & -0,274 & 0,322 \\ 0,211 & -0,523 & 0,312 \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix}. \quad (20.18)$$

Inverzná transformácia z Y'IQ do R'G'B' je

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} 1,000 & 0,956 & 0,621 \\ 1,000 & -0,272 & -0,647 \\ 1,000 & -1,105 & 1,702 \end{bmatrix} \begin{bmatrix} Y' \\ I \\ Q \end{bmatrix}. \quad (20.19)$$

Polárne súradnice

Ak uvažujeme polárne súradnice, transformáciu pre odtieň farby a saturáciu môžeme vyjadriť nasledovne

$$\begin{aligned} H_{YIQ} &= \tan^{-1} \left(\frac{Q}{I} \right), \\ S_{YIQ} &= \sqrt{I^2 + Q^2}. \end{aligned} \quad (20.20)$$

20.3.2 Model YUV

Tento model sa používal na prenos televízneho signálu v európskej norme PAL a SECAM. Podobne ako pri modeli YIQ, aj v tomto modeli je oddelená jasová zložka od farebnej. Zvyšné dva signály obsahujú informáciu o veľkosti farebných zložiek obrazu. Symboly U a V tu nesmieme spájať s U a V CIE YUV (1960).

Zložka Y priestoru YUV je identická s Y v priestore YIQ. Zložka U sa niekedy označuje aj ako B-Y a V ako R-Y. Farebné zložky sa pohybujú v rozsahu -0,5 do +0,5 a jasová zložka má rozsah 0 až 1.

YUV je podobný priestoru YIQ, rozdiel je iba v 33-stupňovom otočení farebných osí

$$\begin{aligned} I &= -U \cdot \sin(33^\circ) + V \cdot \cos(33^\circ), \\ Q &= U \cdot \sin(33^\circ) + V \cdot \cos(33^\circ). \end{aligned} \quad (20.21)$$

Norma pre YUV bola špecifikovaná Európskou vysielacou úniou (European Broadcasting Union – EBU).

Transformácie medzi RGB a YUV

Biely bod systému je D65 a súradnice farieb YUV dostaneme transformáciou z RGB nasledovne

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0,299 & 0,587 & 0,114 \\ -0,147 & -0,289 & 0,436 \\ 0,615 & -0,515 & -0,100 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \quad (20.22)$$

Inverzná transformácia je potom daná

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1,000 & 0,000 & 1,140 \\ 1,000 & -0,395 & -0,581 \\ 1,000 & 2,032 & 0,000 \end{bmatrix} \begin{bmatrix} Y \\ U \\ V \end{bmatrix} \quad (20.23)$$

Y'UV nie je absolútny farebný priestor. Je to v podstate spôsob kódovania RGB informácie a aktuálna zobrazená farba závisí od množiny primárnych farebných podnetov RGB.

20.3.3 Model YPbPr

Y'PbPr je model luma-chroma pre prenos analógového videa cez sieť. Používajú sa rozdielne koeficienty pre SDTV a HDTV, ktoré upravujú štandardy ITU (International Telecommunication Union). Koeficienty pre štandardné video SDTV sú dané normou ITU-R BT.601-5, a pre HDTV ich upravuje štandard ITU-R BT.709-5.

Transformácie medzi RGB a YPbPr

Nasledujúci vzťah opisuje konverziu z RGB formátu do farebného formátu YPbPr. V prípade SDTV je použitá nasledovná matica

$$\begin{bmatrix} Y \\ Pb \\ Pr \end{bmatrix} = \begin{bmatrix} 0,299 & 0,587 & 0,114 \\ -0,169 & -0,331 & 0,500 \\ 0,500 & -0,419 & -0,081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (20.24)$$

Súčet koeficientov v prvom riadku matice je 1. V každom zo zvyšných dvoch riadkov je súčet koeficientov rovný 0. Hodnota 0,5 vyjadruje maximálny rozsah zložiek Pb a Pr pre modrú a červenú primárnu farbu. Inverzná transformácia pre SDTV je daná vzťahom

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1,000 & 0,000 & 1,402 \\ 1,000 & -0,334 & -0,714 \\ 1,000 & 1,772 & 0,000 \end{bmatrix} \begin{bmatrix} Y \\ Pb \\ Pr \end{bmatrix}. \quad (20.25)$$

20.3.4 Model YCbCr

Tento medzinárodný štandard pre digitálne video nezávisí od štandardu skenovania a primárnych parametrov systému, preto nemá chromatické súradnice, CIE XYZ matice ani úvahu o bielom bode, či CRT gama. Zaoberá sa len digitálnym vyjadrením R'G'B' signálov v Y'CbCr podobe. S týmto modelom sa môžeme stretnúť napríklad pri zápise rastrových obrazov do JPEG formátu.

Existuje viacero spôsobov vzorkovania kompletnej farebnej informácie obrazu. Uvedieme si tie najpoužívanejšie:

- 4:4:4** nekomprimovaný obraz, šírka pásma pre farebné zložky je rovnaká ako pre jasovú, teda pre každý pixel máme aj jasovú aj kompletnú chromatickú informáciu
- 4:2:2** farebné zložky sú skomprimované na polovicu, t. j. pre každý druhý pixel máme iba jasovú informáciu
- 4:1:1** horizontálne rozlíšenie farby je štvrtinové, čiže šírka pásma pre farbu je polovičná oproti šírke pásma pre zložku luma
- 4:2:0** horizontálne vzorkovanie je oproti 4:1:1 dvojnásobné, ale Cb, Cr farebné zložky sú vzorkované iba v každom druhom riadku, teda vertikálne rozlíšenie je oproti 4:1:1 polovičné. Takže celkové množstvo dát je rovnaké ako pri 4:1:1

Y'CbCr je definované v štandarde Rec. ITU-R BT. 601 pre štúdióvu kvalitu komponentného digitálneho videa. Signál pre jas je kódovaný 8 bitmi. Nie je však využitý celý rozsah 256 hodnôt. Y' s ofsetom 16 má rozsah 219, od hodnoty 16 pre čiernu po 235 pre bielu farbu. Chromatické zložky s ofsetom 128 majú rozsah ± 112 , od 16 po hodnotu 240 vrátane.

Transformácie medzi RGB formátom a YCbCr pre SDTV

Nasledovné rovnice opisujú konverziu farby z RGB formátu do YCbCr pre TV so štandardným rozlíšením (SDTV) (podľa ITU-R BT.601)

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 0,257 & 0,504 & 0,098 \\ -0,148 & -0,291 & 0,439 \\ 0,439 & -0,368 & -0,071 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \quad (20.26)$$

Inverzná konverzia je daná vzťahom

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1,000 & 0,000 & 1,140 \\ 1,000 & -0,395 & -0,581 \\ 1,000 & 2,032 & 0,000 \end{bmatrix} \begin{bmatrix} Y - 16 \\ Cb - 128 \\ Cr - 128 \end{bmatrix}. \quad (20.27)$$

Rozsah hodnôt pre jasovú a chromatické zložky rezervuje priestor, ktorý je potrebný pri pretečení napríklad v kombinácii so zariadením analógového videa. Pre počítačové aplikácie používajúce farebné formáty RGB a YCbCr je v mnohých prípadoch použitý plný 8-bitový rozsah bez rezervovaného priestoru. Farebný formát s plným rozsahom sa používa napríklad v JPEG obrazoch.

Transformácie medzi R'G'B' a Y'CbCr s využitím 8-bitového rozsahu

Konverzia R'G'B' farieb do Y'CbCr s využitím celého 8-bitového rozsahu je nasledovná

$$\begin{bmatrix} Y' \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 0,229 & 0,587 & 0,114 \\ -0,169 & -0,331 & 0,500 \\ 0,500 & -0,419 & -0,081 \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix}. \quad (20.28)$$

Inverzná transformácia farebného formátu YCbCr s využitím celého 8 bitového rozsahu naspäť do RGB je

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1,000 & 0,000 & 1,400 \\ 1,000 & -0,343 & -0,711 \\ 1,000 & 1,765 & 0,000 \end{bmatrix} \begin{bmatrix} Y \\ Cb - 128 \\ Cr - 128 \end{bmatrix}. \quad (20.29)$$

Y'CbCr pre HDTV

Pre HDTV sa používajú odlišné koeficienty ako pre SDTV, rozsah použitých hodnôt je však rovnaký ako pre SDTV s rezervovaným priestorom v záhlaví a päte. Parametre priestoru Y'CbCr pre HDTV sú definované v štandarde ITU-R BT.709.

20.3.5 Model Kodak Photo YCC

Tento farebný priestor vyvinuli vo firme Kodak na kódovanie obrazov systémom FotoCD. Je odvodený z gama korigovaných R'G'B' hodnôt s použitím bieleho bodu, primárnych podnetov a hodnoty gama definovaných v štandarde ITU-R BT.709. Je podobný priestoru Y'CrCb, rozdiel je len v gamute týchto priestorov. Gamut priestoru YCC bol prispôbený farebnému gamutu fotografického filmu, zatiaľ čo gamut Y'CrCb je orientovaný na luminofory NTSC systému.

Transformácia RGB dát na 8-bitový formát Y'C1C2

Pri kódovaní dát najprv aplikujeme gama korekciu na hodnoty RGB. Potom R'G'B' dáta transformujeme na zložky luma a dve chroma zložky

takto

$$\begin{aligned} Y' &= 0,299R' + 0,587G' + 0,114B', \\ C1 &= -0,299R' - 0,587G' + 0,886B', \\ C2 &= 0,701R' - 0,587G' - 0,114B'. \end{aligned} \quad (20.30)$$

Koeficienty pre C1 dostaneme, ak uvažujeme, že signál $(B - Y)$ dosahuje extrémne hodnoty v modrej t. j. $R = 0, G = 0, B = 1$ a teda $Y = 0,114$ a $B - Y = 0,886$, a v žltej t. j. $R = 1, G = 1, B = 0, Y = 0,886$ a $B - Y = -0,886$. Podobne dostaneme hodnoty pre C2. Extrémy signálu $(R - Y)$ sú pre červenú a modrozelenú a teda $\pm 0,701$.

Následne dáta s reálnymi hodnotami prevedieme na FotoYCC dáta s celočíselnými hodnotami, ktoré sú uložené ako 8-bitové čísla.

$$\begin{aligned} Y'_{8bit} &= \frac{255}{1,402} Y', \\ C1_{8bit} &= 156 + 111,4.C1, \\ C2_{8bit} &= 137 + 135,64.C2. \end{aligned} \quad (20.31)$$

Obidve chroma zložky sú komprimované faktorom 2 horizontálne aj vertikálne. Asymetrická škála medzi C1 a C2 je navrhnutá firmou KODAK tak, aby zachytila rozsiahly farebný gamut podobne ako fotografický film a vystihla typické rozloženie farieb v reálnych scénach.

Kodak YCC môže obsiahnuť viac informácií než bežné zobrazovacie jednotky. Prevod z YCC do RGB preto nie je jednoduchý opak prevodu RGB na YCC, ale závisí od cieľového zobrazovacieho systému.

Inverzná transformácia 8-bitového formátu Y'C1C2 na 24-bitové RGB

Prevod dát z formátu FotoYCC naspäť na 24-bitové RGB dostaneme tak, že najprv obnovíme zložky luma a chroma s reálnymi hodnotami nasledovne

$$\begin{aligned} Y' &= 1,3584.Y_{8bit}, \\ C1 &= 2,2179.(C1_{8bit} - 156), \\ C2 &= 1,8215.(C2_{8bit} - 137). \end{aligned} \quad (20.32)$$

A potom vypočítame nelineárne $R'G'B'$ zložky

$$\begin{aligned} R' &= Y + C2, \\ G' &= Y - 0,194C1 - 0,509C2, \\ B' &= Y + C1. \end{aligned} \tag{20.33}$$

Treba dbať na dve skutočnosti

a) výsledok RGB hodnôt je v rozsahu 0 až 346 (namiesto zvyčajne používaného rozsahu 0 až 255). Na konverziu pomocou nelineárnej funkcie na 8-bitové dáta je obvyčajne použitá vyhl'adávacia tabuľka.

b) ak sa luminofory displeja líšia od ITU-R BT.709 primárnych zložiek, potom bude potrebná ďalšia konverzia cez farebný priestor nezávislý od zariadenia, ako napríklad CIE XYZ.

20.4 Používateľsky orientované modely

Doposiaľ opísané modely sú hardvérovo orientované a pre bežného používateľa môže byť veľmi ťažké vyjadriť niektoré farby (napr. hnedú) pomocou jednotlivých zložiek napríklad v systéme RGB. Preto boli navrhnuté takzvané používateľsky orientované farebné modely, vhodnejšie pre intuitívne vnímanie farieb. Základnými použitými parametrami v nich sú

- farebný odtieň (hue),
- sýtosť farby (saturation),
- jasová zložka (value, lightness, brightness, luminance).

Tieto modely umožňujú používateľovi intuitívne špecifikovať farby a jednoducho si vybrať danú farbu. Vychádzajú z maliarskeho spôsobu vnímania farieb. Tieto modely sú však len jednoduchou lineárnou transformáciou z RGB priestoru, a tak majú aj podobné vlastnosti. Sú vhodné napríklad na výber farieb (color picker) alebo v softvéroch zameraných na editovanie obrazu. Všetky používateľsky orientované modely sú vne-movo nehomogénne.

Používateľské modely na rozdiel od hardvérovo orientovaných ako RGB alebo CMY, umožňujú meniť iba jeden parameter farby, kým zvyšné dva zostanú zachované. Túto dôležitú vlastnosť môžeme využiť napríklad pri ekvalizácii histogramov.

Predstaviteľmi tejto skupiny sú modely HSV (HSB), HSL (HLS), HSI a HCI (Hue Chroma Intensity). Model HSV (V – value) sa niekedy nazýva HSB (B – brightness). Model HSL používa ako tretiu zložku svetlost' (L – lightness), zatiaľ čo model HSI intenzitu. Všetky tieto modely sú podobné, výraznejšie sa líšia v jasovej zložke. V nasledujúcom texte uvádzame stručné vysvetlenie jednotlivých pojmov.

Intenzita svetla (*Intensity*) alebo aj svietivosť je celkové množstvo svetla prechádzajúce cez určitú oblasť. Meria sa v kandelách (cd).

Jas (*Luminance*) je fotometrická veličina, definovaná ako merná veličina svietivosti. Označuje sa L a vyjadruje sa v kandelách na meter štvorcový (cd/m²).

Jasnosť alebo aj vnímaný jas (*Brightness*) je absolútna vnímaná intenzita svetla v psychologickom zmysle.

Svetlosť (*Lightness*) je vnímaná intenzita svetla vzhľadom k referenčnej bielej farbe.

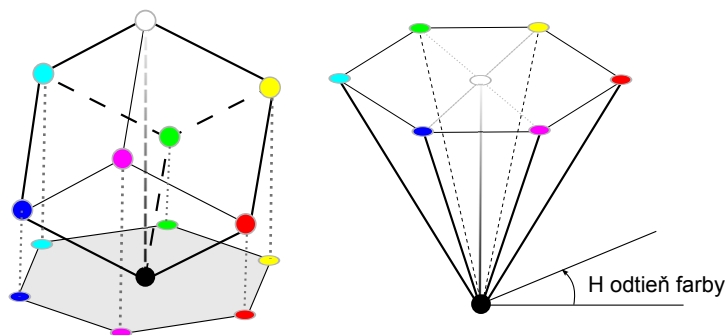
Intenzita a jas sú fyzikálne veličiny, zatiaľ čo vnímaný jas a svetlosť sú atribúty vizuálnej percepcie.

20.4.1 Model HSV

Hlavnými parametrami v modeli HSV sú odtieň farby (H – hue), sýtosť (S – saturation) a jasová hodnota (V – value). Odtieň farby určuje dominantnú spektrálnu farbu, sýtosť určuje prímes ostatných farieb a jasová hodnota množstvo bieleho (bezfarebného) svetla.

Model HSV pochádza zo 70. rokov. Formálne ho prvýkrát opísal Alvy Ray Smith v roku 1978 v augustovom čísle časopisu Computer Graphics. V rovnakom čísle opísali Joblove a Greenberg model HSL, ktorého dimenzie označili ako hue, relative chroma, a intensity.

Pre priestorové zobrazenie HSV modelu väčšinou používame tvar obráteného pravidelného šesťbokého ihlana (hexcone model), ktorého vrchol leží v začiatku súradnicovej sústavy, v bode [0,0,0] a predstavuje čiernu farbu. HSV ihlan vytvoríme deformáciou RGB kocky (obr. 20.4). Ak si predstavíme RGB kocku, potom hlavná diagonála vedúca z čierneho bodu do bieleho bodu definuje jasovú os. Keď pozeráme pozdĺž tejto diagonály na RGB kocku, vidíme šesťuholník. Je to vlastne premietnutie RGB kocky na rovinu kolmú na telesovú uhlopriečku. Ak pozeráme na menšiu kocku, vidíme menší šesťuholník. Každá rovina s konštantnou súradnicou V v priestore HSV vytvára šesťuholník.



Obrázok 20.4: Vľavo: projekcia RGB kocky pozdĺž hlavnej diagonály. Vpravo: HSV model.

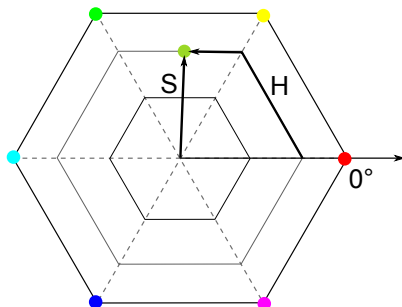
Saturácia (sýtosť farby) sa určuje ako vzdialenosť od stredovej osi jasu po danú farbu. Tón farby, súradnica H je však uhlová z intervalu $\langle 0^\circ, 360^\circ \rangle$. Jasová hodnota V je určená stredovou osou. Hodnoty jasu na nej klesajú od bielej farby v strede podstavy ihlanu až k čiernej vo vrchole ihlanu. V ostatných vrcholoch šesťuholníka na obvode podstavy ihlanu sú umiestnené základné farby (červená – R, žltá – Y, zelená – G, modrozelená – C, modrá – B a fialová – M), kde červená je pri 0° , a zvyšné sú rozdelené po 60° . Komplementárne farby sú od seba vzdialené 180° . Pri cyklickej zmene farby H sa mení vždy len jedna zo súradníc RGB. Veľkosť najväčšej zo súradníc RGB je zhodná s hodnotou V . Farby s plnou saturáciou sa nachádzajú na plášti ihlanu.

Vzhľadom na to, že odtieň farby sa meria ako uhol, je nespojitý v oblasti červených farieb v bode $0^\circ = 360^\circ$. Toto si treba uvedomiť pri riešení úloh počítačového videnia.

Nedostatky modelu HSV

Nedostatkom tohto modelu je jeho ihlanový tvar, ktorý spôsobuje, že bod s konštantnou hodnotou jasu V a konštantnou saturáciou S sa pri zmene hodnoty H pohybuje po šesťuholníku a preto zmena farebného odtieňa nie je plynulá (obr. 20.5). Tento nedostatok odstraňuje HSL model, kde je „šesťuholník nahradený kruhom“. Plynulý nie je ani prechod medzi čiernou a bielou farbou.

Farby s nízkou intenzitou môžu byť zobrazené ako plne saturované. Napríklad farba $[R, G, B] = [0, 0, 0]$ má saturáciu 1, aj keď pre všetky praktické účely je nerozlíšiteľná od čiernej, ktorá má saturáciu 0. A preto



Obrázok 20.5: Saturácia a odtieň farby určený v modeli HSV.

tmavé regióny obrazu majú často veľmi zašumené hodnoty saturácie. Ďalším problémom tohto modelu sú singularity. Pre achromatické farby je saturácia 0 a odtieň farby je nedefinovaný.

Prevod z RGB do HSV

Uvedieme vzťahy pre konverziu z RGB do HSV podľa Trávise ([Tra91]). Pri konverzii RGB na HSV (uvažujúc normalizované hodnoty RGB) najprv nájdeme maximálne a minimálne hodnoty z RGB trojice. Potom saturácia S je

$$S = \frac{\max - \min}{\max} \quad (20.34)$$

a hodnota (Value) V sa rovná maximu

$$V = \max. \quad (20.35)$$

Odtieň farby H sa potom vypočíta takto. Najprv vypočítame $R'G'B'$

$$\begin{aligned} R' &= \frac{\max - R}{\max - \min}, \\ G' &= \frac{\max - G}{\max - \min}, \\ B' &= \frac{\max - B}{\max - \min}. \end{aligned}$$

Ak saturácia $S = 0$, potom odtieň farby nie je definovaný (farba nemá žiadne zafarbenie, preto je achromatická).

$$\begin{array}{lll}
 \text{Ak } R = \max & \text{a } G = \min, & \text{tak } H = 5 + B' \\
 \text{Ak } R = \max & \text{a } G \neq \min, & \text{tak } H = 1 - G' \\
 \text{Ak } G = \max & \text{a } B = \min, & \text{tak } H = R' + 1 \\
 \text{Ak } G = \max & \text{a } B \neq \min, & \text{tak } H = 3 - B' \\
 \text{Ak } R = \max, & & \text{tak } H = 3 + G' \\
 & \text{Inak} & H = 5 - R'.
 \end{array}$$

Odtieň farby je potom konvertovaný na stupne vynásobením 60, čím dostaneme HSV s S medzi 0 a 1, a H medzi 0 až 360.

Spätná konverzia z HSV do RGB

Na spätnú konverziu z HSV na RGB najprv zoberieme odtieň farby H v rozsahu 0 až 360 a vydělíme 60

$$Hex = \frac{H}{60}.$$

Potom vypočítame hodnoty a, b, c, p, q

$$\begin{array}{ll}
 p &= [Hex], \\
 q &= Hex - p, \\
 a &= V(1 - S), \\
 b &= V(1 - S * q), \\
 c &= V(1 - S * (1 - q)).
 \end{array}$$

kde $[]$ je dolná celá časť čísla. Nakoniec vypočítame RGB takto:

$$\begin{array}{ll}
 \text{Ak } p = 0, & \text{tak } [R, G, B] = [V, c, a] \\
 \text{Ak } p = 1, & \text{tak } [R, G, B] = [b, V, a] \\
 \text{Ak } p = 2, & \text{tak } [R, G, B] = [a, V, c] \\
 \text{Ak } p = 3, & \text{tak } [R, G, B] = [a, b, V] \\
 \text{Ak } p = 4, & \text{tak } [R, G, B] = [c, a, V] \\
 \text{Ak } p = 5, & \text{tak } [R, G, B] = [V, a, b]
 \end{array}$$

Treba si pamätať, že HSV je iba mapovanie určitého RGB farebného priestoru. Je to viac-menej vyjadrenie originálneho RGB priestoru v inom

súradnicovom priestore. Saturované farby v HSV sú farby nachádzajúce sa na hrane zodpovedajúceho RGB trojuholníka v chromatickom diagrame. Keďže sa neuvádza, či je HSV odvodené od lineárneho alebo nelineárneho priestoru RGB a väčšinou nevieme ani presné hodnoty trichromatických zložiek RGB, hovoríme, že HSV je systém závislý od zariadenia, vo význame, že nie je absolútny kolorimetrický priestor, ale relatívny vzhľadom ku gamutu RGB, ktorý opisuje.

20.4.2 Model HSL

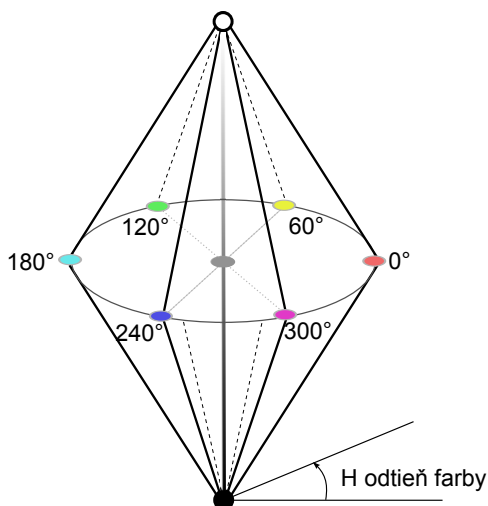
Podobný modelu HSV je aj model HSL, ktorý je na rozdiel od HSV reprezentovaný dvojitém kužeľom (bi-cone model) (obr. 20.4). Niekedy reprezentuje aj pomocou kužeľa alebo gule. Vo vrcholoch hore a dole sa nachádza biela a čierna farba. Stredovú os tvorí svetlosť (lightness). Svetlosť používame v súvislosti s farbami, napríklad svetlozelená, tmavozeleň farba.

V HSL dvojitém kuželi klesá počet farieb od stredu smerom hore k bielej farbe aj smerom dole k čiernej farbe. Maximum farieb v tomto modeli je pri hodnote svetlosti 0,5. V HSV ihlne počet farieb klesá s hodnotou parametra jas. Hore pri základni, kde je aj najväčší počet farieb, je hodnota parametra jas rovná 1, čiže najvyšší jas. Zatiaľ čo HSV má veľké množstvo farieb s maximálnou hodnotou jas, HSL má iba jednu – bielu. Dôsledkom tohto je, že dve farby s rovnakým odtieňom farby a maximálnou saturáciou budú v modeli HSV svetlejšie než v modeli HSL. Tvar modelu HSL zodpovedá viac skutočnosti, že najviac rôznych farieb vnímame pri priemernej svetlosti. Schopnosť rozlišovania farebných odtieňov skutočne klesá so stmavovaním aj so zosvetľovaním základnej čistej farby. V HSV 100% jas zodpovedá veľmi saturovaným farbám, čo ich robí neprírodnými.

Saturácia v modeli HSL sa pohybuje pri danej svetlosti od plne nasýtenej farby po sivú. V modeli HSV sa saturácia pohybuje hore na základni od plne nasýtených až po bielu farbu.

Nedostatky

Nevýhodou priestorov HSV a HSL je, že odtieň farby v týchto priestoroch nie je lineárne závislý od hodnoty farebného odtieňa v priestore RGB a ani jeho zmena nie je spojitá. Preto sa hrana v kocke RGB zobrazí ako oblúk v modeloch HSV a HSL. Toto je nevýhoda hlavne pri interpoláciách. Keďže obidva farebné modely sú odvodené od RGB modelu, majú



Obrázok 20.6: Model HSL. Po obvodu sa mení odtieň farby, na strednej osi sa mení svetlosť (lightness) a saturácia sa pohybuje od strednej osi po okraj.

aj jeho gamut. Transformácia medzi nimi prebieha v konštantnom čase bezstratovým algoritmom.

Pri prevode HSL na RGB musíme dávať pozor, aby sme netransformovali hodnoty mimo RGB kocky. Predpokladajme, že máme červenú farbu s $[R, G, B] = [1, 0, 0]$. Ak ju transformujeme do HSL, dostaneme $[H, S, L] = [0, 1, 1/2]$. Toto je maximálne satureovaná farba pri intenzite $1/2$ a je nemožné ďalej zvyšovať intenzitu na hodnotu 1 a zároveň zachovať rovnakú saturáciu.

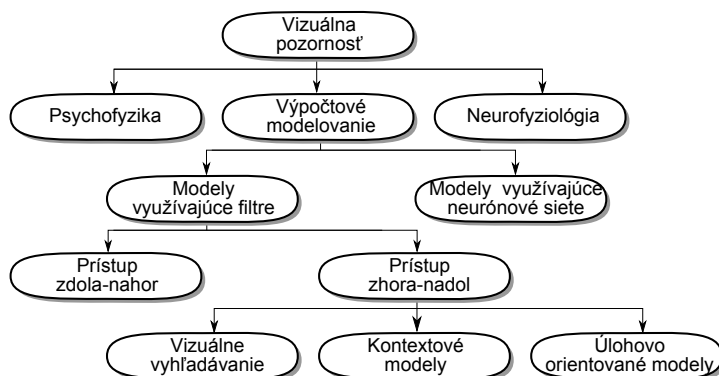
Vizuálne vnímanie a pozornosť

V tejto časti sa budeme venovať vizuálnemu vnímaniu a pozornosti. Do ľudského oka vstupuje každú sekundu veľké množstvo dát ($10^8 - 10^9$ bitov). Keďže ľudské vnímanie je obmedzené a ľudia nedokážu vnímať všetky tieto vstupné informácie naraz, je veľmi dôležité informácie triediť na základe ich dôležitosti. Tento proces prebieha v ľudskom vizuálnom systéme automaticky, avšak z hľadiska aplikácií je potrebné zistiť, na ktoré oblasti v scéne sa upriamuje pozornosť človeka a kde v scéne sa tieto oblasti nachádzajú. Takéto oblasti sú označované ako významné.

Ľudské vizuálne vnímanie závisí od množstva faktorov – vizuálnych vlastností pozorovanej scény, medzi ktoré patrí napr. farba, textúra, intenzita, orientácia. Preto treba zohľadňovať tieto faktory pri skúmaní scény a určovaní významných oblastí v nej. Avšak ľudia pri vnímaní scény využívajú aj vnemovú inteligenciu – vedomosť, ktorú získavajú skúsenosťami pri pozorovaní. Táto vedomosť veľkou mierou prispieva ku konečnému vizuálnemu vnemu.

Významné oblasti scény nesú pre pozorovateľa potenciálne dôležitú informáciu a poznatky o ich výskyte v obraze majú veľmi široké využitie v praxi, napr. pri kompresii a kódovaní obrazu, v rozpoznávaní, detekcii, kategorizácii objektov, pri segmentácii obrazu, odstraňovaní artefaktov, renderovaní, spracovaní multimédií, orezávaní obrazu [Gol08].

V nasledujúcich častiach si objasníme pojem vizuálnej pozornosti. Budeme sa bližšie venovať dvom najzákladnejším spôsobom určovania oblastí, ktoré priťahujú ľudskú pozornosť: sledovaniu pohybu očí a tvorbe máp významných oblastí.



Obrázok 21.1: Graf rozdelenia štúdií vizuálnej pozornosti.

21.1 Vizualna pozornost'

Vizuálna pozornosť veľmi vplýva na vnímanie, na pamäť (ľudia sú schopní si niečo zapamätať, len ak na to sústredia svoju pozornosť), na jazyk (čítanie zahŕňa pozornosť na text slovo po slove) aj na riešenie problémov (úspech v riešení problémov závisí od toho, ktorý problém zaujme pozornosť človeka) [Gol08].

Na obrázku 21.1 je zobrazený graf, ktorý znázorňuje rozdelenie štúdií vizuálnej pozornosti. Z neho je zrejmé, že vizuálna pozornosť je oblasť výskumu, ktorá je veľmi zaujímavá a zaoberajú sa ňou vedci z rôznych vedných odvetví [BI13].

Pri vnímaní scény sa rozlišujú dva hlavné prístupy

- zdola-nahor (bottom-up),
- zhora-nadol (top-down).

Pomocou týchto dvoch prístupov sa vizuálna pozornosť rozdeľuje na *endogénnu*, ktorá využíva prístup zhora-nadol a *exogénnu* založenú na prístupe zdola-nahor. Exogénna pozornosť je zameraná na vonkajšie podnety, upozorňuje na výrazné časti automaticky a má rýchly časový priebeh. Existuje široká škála exogénnych vizuálnych atribútov na zachytenie pozornosti, ako sú napríklad priestorové vnemy, náhle vizuálne zmeny jasu, farby atď.

V endogénnej časti je pozornosť určovaná subjektom. Je známa ako cieľovo riadená pozornosť. Endogénna pozornosť je riadená pozorovateľom, závislá na ňom a je pomalá z dlhodobého časového hľadiska.

Existuje viacero modelov vizuálneho vnímania, ktoré využívajú tieto prístupy oddelene. Avšak vo všeobecnosti takmer každý model vizuálneho vyhľadávania je založený na vzájomnej interakcii medzi týmito prístupmi. Oba hlavné prístupy sú navzájom nezávislé.

Prístup zdola-nahor

Postupnosť udalostí, ktoré sa začali stimuláciou receptorov, je zaradovaná medzi exogénne. Je riadená stimulmi, kde sa automaticky, bez vonkajších podnetov, pozornosť zameriava na výraznejšie časti obrazu. Túto pozornosť môže ovplyvniť veľké množstvo podnetov, ako sú napríklad zmeny vo farbe, zmena polohy, intenzity. Všetky tieto zmeny ľudské oko zachytáva a podvedome upriamuje pozornosť práve na ne.

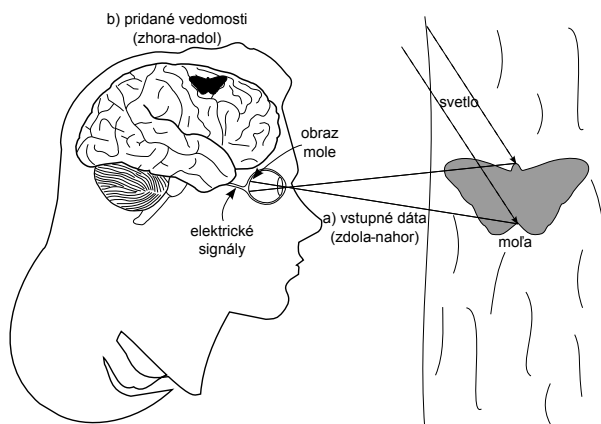
Množstvo psychologických výskumov ukázalo, že vizuálny systém je vysoko citlivý na vlastnosti v obraze, ako sú napríklad hrany, náhle zmeny vo farbe, intenzite, náhle pohyby.

Procesy zdola-nahor sú nezávislé na úlohe, ktorú pozorovateľ rieši. Snažia sa predvídať, ktoré časti pozorovanej scény môžu pritiahnúť viac pozornosti, a podľa toho vytvárajú mapy významných oblastí v obraze. Tieto procesy môžu byť využívané napríklad pri automatickej detekcii významných oblastí v scénach, v strojovom videní, pri inteligentnej kompresii obrazu atď. Pri analyzovaní scény pomocou prístupov zdola-nahor sú medzi významné objekty zaradené napríklad červené jablko na zelenom strome, horiaca sviečka v tmavej miestnosti, alebo pery a oči na tvári, ktoré sú považované za jej najvýraznejšie časti.

Pri pozorovaní scény, v ktorej sa nachádza veľké množstvo rovnakých objektov, ktoré by samostatne boli určené ako výrazné, sa tieto objekty stávajú nevýraznými. Keď pozorovateľ porovnáva dve fotografie jabloní s červenými jablkami, kde na jednej je iba jedno a na druhej je veľké množstvo jablák, v prvom prípade si jablko všimne, ale v druhom prípade sú jablká odfiltrované a pozorovateľ si všimne iné výrazné oblasti a objekty [CCM03].

Prístup zhora-nadol

Vizuálne vnímanie je oveľa komplikovanejšie, než len vnímanie energie na receptoroch, čo sa deje pri procesoch založených na prístupe zdola-nahor. Preto je pri procesoch založených na prístupe zhora-nadol pozorovanie scény čiastočne ovládané pozorovateľom. Toto ovládanie sa deje pridaním dodatočnej informácie ako napríklad tým, že človek je pri pozorovaní nových scén ovplyvnený vecami, ktoré už videl v minulosti. Toto



Obrázok 21.2: Pozorovanie mole je ovplyvnené kombináciou a) vstupných dát a b) predošlých vedomostí.

pozorovanie zahŕňa skúsenosti pozorovateľa, avšak môže byť veľmi pomalé.

Ako príklad prístupov zhora-nadol je možné použiť pozorovanie kôry stromu. Pri procesoch zdola-nahor sa určí farba kôry, rozlišujú sa hrany a textúra, ale až pri pridaní ďalšej informácie je možné určiť, že na kôre stromu sa nachádza mola (obr. 21.2).

Pri využívaní prístupov zhora-nadol môže byť ako prídavná informácia použitá úloha, ktorú má pozorovateľ vyriešiť. Avšak, ak má pozorovateľ napríklad nájsť určitý objekt v scéne, môže prehliadnuť iné výrazné oblasti v nej. Preto môže nastať situácia, kde niektoré objekty významné pri detekcii pomocou prístupov zdola-nahor nemusia byť považované za významné pri prístupe zhora-nadol.

Skúmaním vizuálnej pozornosti a procesov zhora-nadol sa zaoberal aj ruský psychológ Yarbus už v roku 1965. Jeho výskum spočíval v zaznamenávaní fixácií a skokov oka pozorovateľov, ktorí mali úlohu skúmať danú scénu a odpovedať na položené otázky. Pohyby očí sa výrazne menili v závislosti od daných otázok [CCM03].

21.2 Určovanie významných oblastí v obraze

Pri mnohých aplikáciách z oblasti grafiky, dizajnu a HCI je veľmi dôležité porozumenie tomu, kam človek zameriava svoju vizuálnu pozornosť. Tieto informácie sa využívajú napríklad pri automatickom orezávaní.

vaní obrazu, vytváraní náhl'adov, pri vyhľadávaní konkrétneho obrazu v databáze. Môžu byť použité taktiež pri kompresii videa, nefotorealistic-kom renderingu, adaptívnom zobrazovaní obrazu na malých displejoch.

Vo všeobecnosti je vzhľadom na uvedené aplikácie potrebné čo najlepšie, najpresnejšie a najrýchlejšie detegovať oblasti záujmu. Na určenie týchto oblastí sú k dispozícii dva hlavné prístupy

- sledovanie pohybu očí („eye tracking“),
- modely na detekciu významných oblastí.

Pre niektoré aplikácie sú používané dáta získané pomocou sledovania pohybu očí. Hlavný princíp spočíva v sledovaní a zaznamenávaní pohybov očí pozorovateľa. Pozorovateľ sleduje scénu a systém zaznamenáva pohyby jeho oka, čím sa zisťuje, kam sa zameriava jeho pozornosť. Avšak takýto spôsob zisťovania významných oblastí je finančne a časovo náročný, a preto nie je vždy možné ho používať. Z týchto dôvodov je veľmi potrebné detegovať významné oblasti aj iným spôsobom. Ako alternatíva sú preto navrhované a používané modely na detekciu pravdepodobnosti výskytu významných oblastí v obraze.

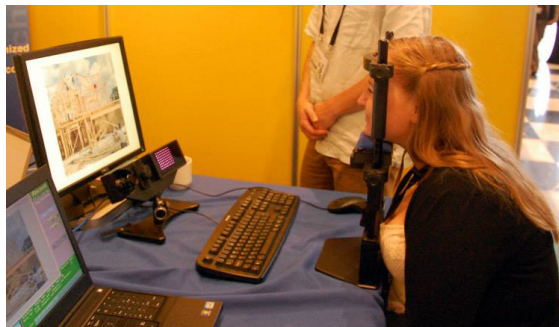
21.2.1 Sledovanie pohybu očí

Oči a ich pohyby sú veľmi dôležité pri vyjadrení potrieb, túžob a emocionálneho rozpoloženia človeka. Z toho dôvodu návrh robustných techník na detekciu očí a ich pohybu zohráva kľúčovú úlohu pri vývoji rôznych techník interakcie človeka s počítačom, vytváranie používateľských rozhraní, zisťovanie oblastí záujmu pozorovateľa.

Detekcia očí a sledovanie ich pohybu je atraktívnou oblasťou výskumu už mnoho rokov a v súčasnosti je využívaná v mnohých aplikáciách. Je to jedna z netradičných techník, ktorá sa využíva pri interakcii človeka s počítačom.

Prehľad techník na sledovanie pohybu očí

Geometrické a pohybové charakteristiky očí sú unikátne, čo umožňuje využitie ich detekcie pri rôznych aplikáciách. Hardvérový systém na sledovanie polohy a pohybu očí sa nazýva „eye tracker“. Vo všeobecnosti existujú dva hlavné prístupy na monitorovanie pohybov očí – meranie pozície očí v závislosti na hlave pozorovateľa a meranie orientácie očí v priestore. V rámci softverového riešenia „eye trackera“ existuje veľké množstvo rôznych prístupov [AF13]. V tejto časti si bližšie opíšeme dva z nich



Obrázok 21.3: Sledovanie pohybov očí s fixnou polohou hlavy.

- sledovanie pohybu očí pomocou senzorov,
- sledovanie pohybu očí pomocou techník počítačového videnia.

Sledovanie pohybu očí pomocou senzorov

Niektoré systémy analyzujú pohyby očí na základe elektrických potenciálov pomocou elektrookulogramu (EOG). Pre každé oko sú v jeho blízkosti umiestnené dva páry elektród, ktoré zaznamenávajú elektrický potenciál. Na základe týchto informácií je možné určiť polohu očí.

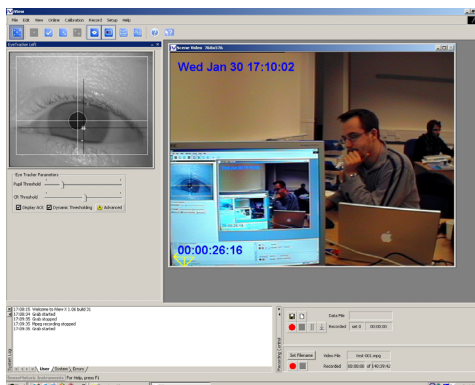
Sledovanie pohybu očí pomocou techník počítačového videnia

Veľké množstvo metód na sledovanie polohy a pohybu očí využíva techniky počítačového videnia s použitím kamery, ktorá sníma jedno, alebo obe oči a následne spracováva tieto informácie. Pozícia kamery môže byť rôzna. V niektorých prístupoch sa využíva kamera pripevnená k hlave pozorovateľa, čo je veľmi výhodné pri systémoch, kde používateľ pozoruje svoje okolie. V prípade sledovania obrazov na monitore sa bežne využíva fixné postavenie kamery a hlavy pozorovateľa (obrázok 21.3). Pomocou týchto techník sa riešia dve hlavné otázky

- určovanie pozície očí vzhľadom na pozorovanú scénu,
- sledovanie pohybu očí.

Určovanie pozície očí a sledovanie ich pohybu je stále veľmi náročná úloha, pri ktorej je potrebné riešiť viacero problémov ako napríklad stupeň otvorenia očí, variabilita vo veľkosti očí, pozícia hlavy a podobne.

Rôzne **techniky rozpoznávania obrazov a redukcie dimenzie príznakov** sa ukázali ako veľmi efektívne pri sledovaní pohybu očí. Redukcia



Obrázok 21.4: Ukážka systému na sledovanie pohybov očí pomocou techník počítačového videnia.

príznakov pomocou PCA efektívne pomohla znížiť počet hlavných komponentov obrazu oka pre lepšiu detekciu. Taktiež sa využíva neurónová sieť, ktorá je trénovaná počas kalibrácie kamery, kde sa používateľovi postupne zobrazuje niekoľko bodov, ktoré má pozorovať. Avšak využívanie klasifikácie čiastočne spomaľuje celý proces určovania pozície a pohybu očí.

Jedným z ďalších prístupov sú **systémy využívajúce kruhový tvar zreničky a detekciu hrán**. V [CK07] bol navrhnutý systém využívajúci detekciu tváre a Gaborove filtre. Potenciálna oblasť zodpovedajúca tvári je určená detekciou pokožky a následne sa tento predpoklad overí s geometrickým vzorom štruktúry tváre. Ďalej sú na oblasti s pravdepodobným umiestnením očí aplikované štyri Gaborove filtre ($0, \pi/4, \pi/2, 3\pi/4$), keďže zrenička nemá žiaden dominantný smer, a teda ich kombináciou môže byť jednoducho určená. V prípade [KBW08] je využívaný algoritmus detekcie najdlhšej čiary, ktorý je založený na predpoklade, že zrenička má kruhový tvar. Je nájdená najdlhšia vertikálna a horizontálna čiara, kde ich prienik určuje stred zreničky. Veľké množstvo systémov využíva tiež kruhovú Houghovu transformáciu a jej kombináciu s detekciou hrán.

Ďalší spôsob detekcie pohybu očí je **určovanie polohy zreničky**, kde existujú dva hlavné prístupy: metóda tmavej zreničky a svetlej zreničky. Metóda využívajúca tmavú zreničku je založená na kontraste zreničky s dúhovkou. Avšak v tomto prípade môže nastať problém pri používaní ľudí s veľmi tmavými očami, kde je kontrast príliš nízky. Metóda svetlej zreničky využíva odraz infračerveného svetla od sietnice, čo spôsobuje,

že zrenička sa zdá byť biela v obraze oka, čím sa rieši problém s kontrastom.

Niektoré zo systémov sú založené na **porovnávaní oka s jeho vzorom**. Pomocou tohto prístupu bol navrhnutý aj systém, ktorý umožňoval prirodzené pohyby hlavy pri sledovaní očí a minimalizoval potrebu kalibrácie systému.

Kombináciou rôznych techník sledovania pohybu očí sú vytvárané **hybridné systémy**. Jeden z takýchto systémov je založený na detekcii zreničky. Vychádza z jej intenzity, tvaru a veľkosti. Je využívané špeciálne infračervené osvetlenie, pri ktorom sa zrenička javí svetlejšia ako zvyšok snímanej tváre. Jej intenzita je používaná ako primárny príznak pre detekciu, avšak v snímanom obraze sa môžu objaviť iné rovnako svetlé objekty. Preto sa využívajú aj prídavné informácie o zreničke, ako jej veľkosť a tvar. V iných prípadoch sa využíva napríklad kombinácia systémov zachytávania tmavej a svetlej zreničky, odrazu od rohovky s metódou iteratívnej krivky na efektívnu detekciu zreničky a mnoho ďalších.

Využitie systémov na sledovanie pohybu očí

Oblasť výskumu určovania polohy a pohybu očí sa stala atraktívnou hlavne kvôli veľkému množstvu aplikácií, ktoré môžu využívať robustné metódy sledovania pohybu očí. Jednou z hlavných oblastí využitia sú alternatívne spôsoby pre komunikáciu človeka s počítačom. Považujú sa za jednoduchšie ako napríklad ovládanie hlasom, alebo EEG/ECG signálmi. Taktiež dosahujú vyššiu presnosť. Ako príklad môžeme uviesť systémy, ktoré sa využívajú pre **kontrolu prístupu a podpory technológií** pomocou snímania očí. Jeden z takýchto systémov je určený na písanie pomocou pohybov očí, ktorého výstupom je text konvertovaný na reč. Rovnako je možné pomocou očí kontrolovať rôzne systémy vrátane elektronických vozíkov, alebo televízneho ovládača.

Detekcia očí môže byť taktiež využívaná pri **biometrickom rozpoznávaní používateľov**. V tomto prípade sa systém používa na lokalizáciu očí a následne dúhovky, čo je prvý a veľmi dôležitý krok pri rozpoznaní používateľa na základe dúhovky.

Ďalšou veľmi zaujímavou aplikáciou týchto systémov je **využitie v optometrii**. Použitie týchto systémov môže dopomôcť k odhadu zorného uhla jednotlivca, najmä určiť jeho slepý uhol.

E-learningové systémy sú v súčasnosti veľmi rozšírené. Avšak aj napriek tomu, že používatelia sú väčšinou zvyknutí na prácu s počítačom, učenie s použitím takýchto systémov môže byť dosť odlišné. Pri interakcii používateľa s vyučujúcim je veľmi dôležitá emočná časť tohto vzťahu,

ktorá však v takomto prípade chýba. Preto je možné využívať snímanie očí a ich pohybu na sledovanie používateľových aktivít, správanie a emočné rozpoloženie používateľa. Na základe týchto informácií môže byť určený čas na vykonávanie danej úlohy. Na základe interpretácie pohybov očí môže byť rozpoznávané, či má používateľ ťažkosti s porozumením úlohy, je unavený alebo v strese.

V automobilovom priemysle sa sledovanie pohybu očí využíva na **rozvoj monitorovania a asistencie pri systémoch využívaných v autách**. Sledovanie pohybu očí je napríklad využívané na varovanie vodiča v prípade, že začne byť veľmi unavený alebo zaspí počas jazdy. Únava vodiča môže byť detekovaná analýzou prahu žmurkania, stavu očí (otvorené / zatvorené) a rovnako aj určovaním času, kedy pohľad vodiča zostáva v rovnakom smere.

21.2.2 Modely na detekciu významných oblastí

Model vizuálnej pozornosti je opis pozorovaného a predpovedaného správania ľudskej vizuálnej pozornosti. Modely na detekciu významných oblastí môžu využívať prirodzený jazyk, blokové diagramy, matematické prístupy, algoritmy alebo výpočty. Ich úlohou je detegovať oblasti, ktoré s najväčšou pravdepodobnosťou priťahujú pozornosť pozorovateľa.

Modelovaniu systémov na simuláciu vizuálnej pozornosti sa vedci aktívne venujú už vyše 25 rokov. Bolo vytvorené veľké množstvo modelov, ktoré sa úspešne používajú v počítačovom videní, robotike, kognitívnych systémoch. Tieto modely sa zameriavajú na zachytenie významných oblastí v obraze a ich ďalšie použitie [BI13], [Gol08].

V súčasnosti existuje veľké množstvo modelov na detekciu významných oblastí inšpirovaných biologickými procesmi, ktoré sú založené na prístupe zdola-nahor (napr. [IKN98], [HRC05]). Tieto modely sú však iba základným opisom správania sa ľudského vizuálneho systému. Detegujú významné oblasti použitím nízkoúrovňových príznakov, ako sú farba, intenzita, orientácia, textúra, pohyb. Taktiež používajú stratégiu víťaza, nemožnosť návratu, detekciu na rôznych úrovniach a podobne. Najväčším obmedzením týchto modelov je neprítomnosť ďalšej informácie, ktorá by v kombinácii s nízkoúrovňovými príznakmi pomohla presnejšie určiť významné oblasti v obraze. Preto je pri navrhovaní nových a presnejších modelov veľmi dôležitý kognitívny proces a využitie sémantickej informácie.

Prvá otázka, ktorá sa naskytne pri uvažovaní nad využitím kognitívneho procesu, je existencia miesta v ľudskom mozgu, kde sa lokalizuje mapa významných oblastí. Avšak podľa doterajších výskumov také mie-

sto neexistuje. Toto vysvetľuje fakt, že významnosť oblastí v pozorovanom obraze neurčujú iba nízkoúrovňové príznaky, ale taktiež naše vedomosti, spomienky, očakávania.

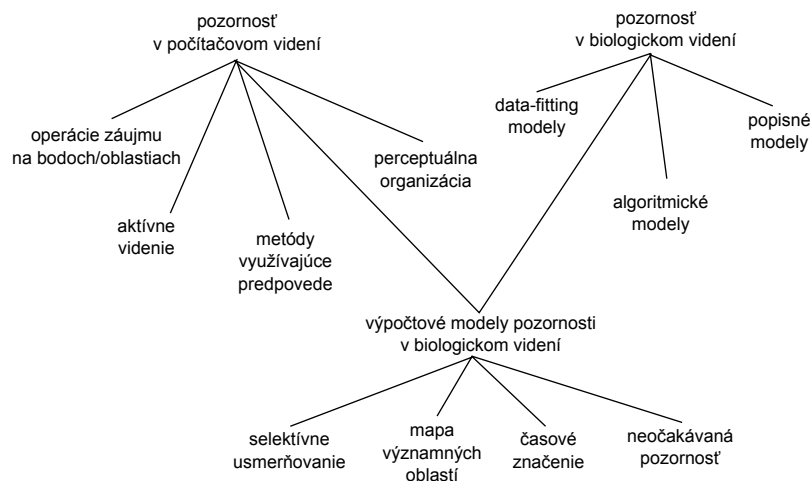
Vizuálne a kognitívne procesy sú veľmi úzko prepojené a v súčasnosti nie je možné určiť, ktoré z nich v danom čase pozorovania určujú rozloženie vizuálnej pozornosti. Pri určovaní presnosti modelov založených na procesoch zdola-nahor sa preto využíva sledovanie očí. Avšak aj pri sledovaní očí nemôžeme zabúdať na procesy, ktoré ovplyvňujú vnímanie zo sémantického hľadiska. Napríklad pri voľnom pozorovaní scény sú výrazné iné oblasti ako pri pozorovaní scény s určitou úlohou. Výraznosť týchto oblastí je ovplyvnená aj správaním, myšlienkami a spomienkami pozorovateľa. Veľmi dobrým prístupom je charakterizácia významnosti oblastí pomocou analýzy dát získaných sledovaním očí. Tento bol využitý v práci [Jud+09] a dáva veľmi dobré výsledky.

Rozdelenie modelov na detekciu významných oblastí

V roku 1998 bol vytvorený jeden z prvých modelov založený na simulácii biologického vnímania scény využívajúci procesy zdola-nahor, detekciu nízkoúrovňových príznakov [IKN98]. Odvtedy vzrastal záujem o tému detekcie významných oblastí. Niektoré ďalšie modely už využívajú kombináciu nízkoúrovňových a vysokoúrovňových príznakov (napr. detekcia pokožky, tváre, textu). Tieto kombinácie dokážu pomôcť pri riešení veľmi špecifických úloh ako napríklad detekcia chodcov. V súčasnosti sa výskum nachádza iba na začiatku tvorby týchto výpočtových modelov, ktoré by popri nízkoúrovňových príznakoch využívali aj pridanú informáciu, čím by sa zlepšila a spresnila detekcia významných oblastí obrazu.

Existujú dva hlavné prístupy vo výskume vizuálnej pozornosti. Prvý z prístupov sa zameriava na využitie metód pozornosti v počítačovom videní. Táto časť demonštruje dôležitosť redukcie informácií pre ďalšie spracovanie. Hlavnou motiváciou pre druhý prístup, ktorý skúma pozornosť v biologickom videní, je vysvetlenie procesov pozornosti pre charakteristiky biologického, špeciálne ľudského videnia (obr. 21.5). Výpočtové modely patria do prieniku týchto dvoch prístupov skúmania pozornosti. Tieto modely zahŕňajú formálny popis vizuálnej pozornosti. Hlavnou motiváciou pre ich vytvorenie je redukcia množstva informácií, ktoré je potrebné spracovať [Tso11].

Modely na detekciu významných oblastí v obraze môžu byť kategorizované na základe veľkého množstva rôznych príznakov a prístupov. V [BI13] bolo preskúmaných viac ako 65 modelov na detekciu význam-



Obrázok 21.5: Diagram znázorňujúci výskum vizuálnej pozornosti.

ných oblastí. Následne bol navrhnutý prístup kategorizácie týchto modelov, ktorý na základe 13 rôznych faktorov rozdeľuje modely do 8 základných skupín.

Faktory rozdelenia modelov do skupín sú nasledovné:

1. Procesy zdola-nahor
2. Procesy zhora-nadol
3. Informácia o priestore a čase
4. Typ úlohy (voľné pozorovanie scény, vizuálne vyhľadávanie, interaktívna úloha)
5. Priestorovo alebo objektovo založené
6. Príznamy (intenzita, farba, textúra, orientácia, pohyb, tváre, detekcia horizontu, optický tok atď.)
7. Typ modelu
8. Statické vizuálne stimuly (napr. fotografie)
9. Dynamické vizuálne stimuly (napr. videá, hry)

10. Syntetické stimuly (hry, virtuálne prostredie, maľby)
11. Prírodné stimuly (fotografie, videá prírodných scén)
12. Spôsob hodnotenia modelu (Kullbackova-Leiberova divergencia, použitie ROI, ...)
13. Použitá databáza

Faktory 1 až 8 priamo opisujú vlastnosti modelov a faktory 9 až 13 s nimi nie sú priamo spojené.

Na základe predchádzajúcich faktorov môžeme modely rozdeliť do ôsmich kategórií, avšak niektoré modely môžu patriť aj viacerým kategóriám.

Prvou z kategórií sú **kognitívne modely**. Skoro všetky modely sú priamo, alebo nepriamo spojené s kognitívnymi procesmi. Jeden z najznámejších modelov patriaci do tejto kategórie [IKN98] je založený na správaní a neurónovej architektúre vizuálneho systému primátov. Tento model sa stal základom pre mnoho ďalších modelov (napr. [Dha03]) a stal sa štandardom pre porovnávanie nových modelov. Existuje veľa jeho modifikácií ako napríklad pridanie detekcie kontrastu v pohybe, slúžiaco na detekciu významných oblastí vo videu.

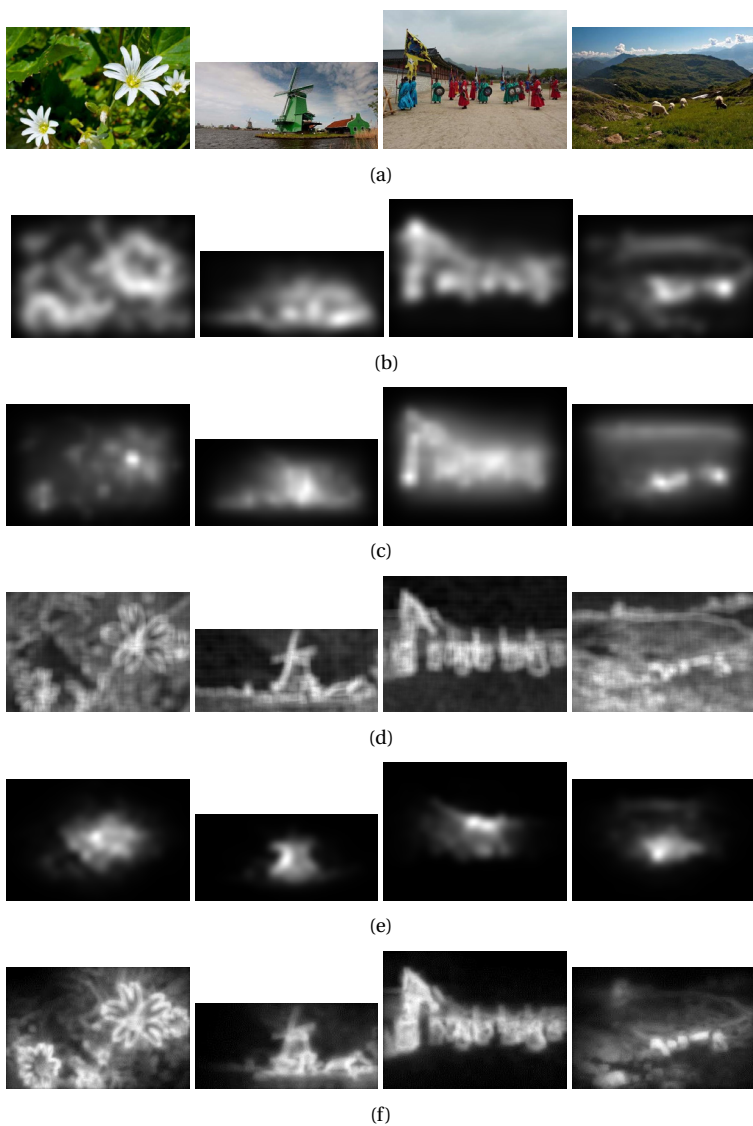
Bayesovské modely (napr. [Oli+03]) využívajú kombináciu prístupu zdola-nahor s predchádzajúcimi znalosťami, napríklad kontext scény. Pretože majú možnosť učiť sa z dát, môžu napríklad využiť štatistiku prírodných scén alebo iných príznakov, ktoré priťahujú pozornosť pozorovateľa.

Informačné modely sú založené na predpoklade, že určená poloha významných oblastí slúži na maximalizáciu informácie získanej z prostredia. Tieto modely (napr. [BT05]) vyberajú tie oblasti v obraze, ktoré nesú najviac informácií a ostatné vyradia.

Modely patriace do kategórie **rozhodovacích modelov** (napr. [GV04]) boli veľmi úspešné v aplikáciách počítačového videnia, ako je klasifikácia, keďže dosahovali vysokú presnosť v predpovedi fixácií pohľadu používateľa.

Grafové modely (napr. [AL10]) využívajú techniky ako skryté Markovove modely, dynamické bayesovské siete a podmienené náhodné polia na určenie vzťahov medzi náhodnými premennými. Na základe týchto vzťahov sa tvoria mapy významných oblastí.

Medzi **modelmi využívajúcimi spektrálnu analýzu** a modelmi ostatných kategórií je veľký rozdiel. Kým iné modely spracovávajú obraz



Obrázok 21.6: Modely na detekciu významných oblastí: (a) Originálny obraz, Mapa významných oblastí detekovaná pomocou (b) [IKN98], (c) [HKP07], (d) [Zha+08], (e) [EE13], (f) [GZT10].

v priestorovej doméne, tento typ modelov (napr. [Ach+09]) deteguje významné oblasti v obraze vo frekvenčnej doméne.

V modelovaní vizuálnej pozornosti sa využívajú aj prístupy strojového učenia s využitím učenia modelov na základe informácií získaných použitím sledovania pohybov očí alebo významných oblastí označených používateľmi. V tomto prípade je dôležité podotknúť, že modely patriace do kategórie **modelov s využitím klasifikácie príznakov** (napr. [Kie+09]), nemusia byť založené len na prístupe zdola-nahor, keďže využívajú príznaky, ktoré sú význačné pre prístupy zhora-nadol (napr. detekcia tváří, textu).

Modely, ktoré nepatria do žiadnej z predchádzajúcich kategórií tvoria skupinu **iných modelov**. Napríklad v [RC02] bolo predstavené meranie významnosti oblastí použitím viacnásobných príznakov založených na koncepte teórie hier. Ďalší spôsob detekcie významných oblastí bol postavený na základe porovnávania s predlohou [Rao+02]. Schéma detekcie významných oblastí založená na hranách v šedotónovom obraze bola navrhnutá v [Ros09].

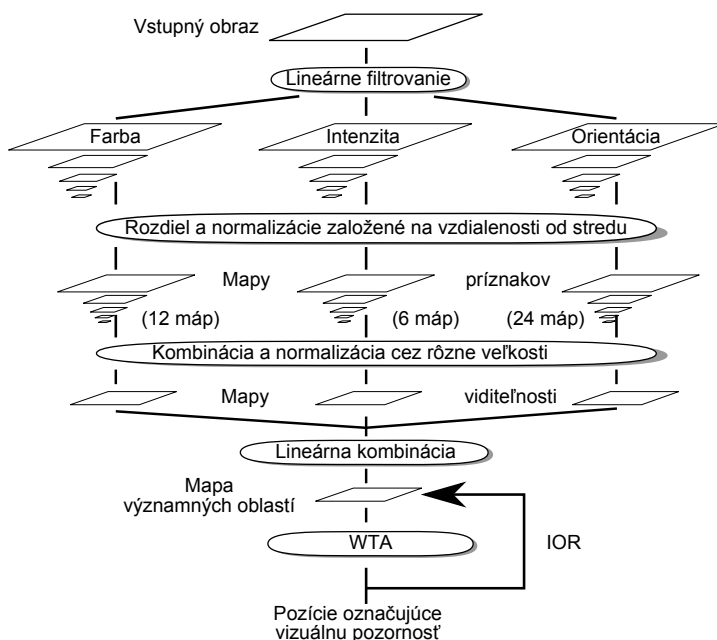
Na obrázku 21.6 sú znázornené mapy významných oblastí. Každý zo zobrazených modelov patrí do inej kategórie. Výsledné mapy významných oblastí (saliency maps) sú veľmi rozdielne v závislosti od prístupu a v súčasnosti nie je možné presne určiť najlepší prístup na určovanie významných oblastí v obraze.

21.2.3 Model pre rýchlu analýzu scény

V tejto časti si opíšeme jeden zo základných modelov [IKN98]. Bol navrhnutý v roku 1998 a slúži ako základ pre veľký počet modelov, rovnako ako aj referenčný model pre porovnanie úspešnosti detekcie. Je založený na správaní a štruktúre neurónov vizuálneho systému primátov. Kombinuje viacúrovňové príznaky do jednej mapy významných oblastí. Používa dynamickú neurónovú sieť, ktorá vyberá oblasti záujmu v poradí od najvýraznejších po menej výrazné. Tento systém rieši zložitý problém porozumenia scény výpočtovo efektívnym spôsobom.

Na vstupe sú použité farebné obrazy v rozlíšení 640×480 pixelov, z ktorých je pomocou Gaussových pyramíd vytvorených 9 obrazov rôznych veľkostí určených na ďalšie spracovanie. Tieto sú zmenšenou kópiou pôvodného obrazu v pomere 1 : 1 (škála 0) až 1 : 256 (škála 8).

Každý príznak je vypočítaný pomocou operácií zameraných na stred (center-surround), keďže vizuálne neuróny sú najcitlivejšie v malej oblasti (stred), zatiaľ čo na podnety zo širšej oblasti sú neuróny menej citlivé. Zameranie na stred je v tomto modeli implementované pomocou



Obrázok 21.7: Základná architektúra modelu pre rýchlu analýzu scény.

rozdielu medzi jemnými a hrubými mierkami. V strede je reprezentácia pixela s veľkosťou $c \in \{2, 3, 4\}$, v širšom okolí je pixel reprezentovaný veľkosťami $s = c + \delta$, kde $\delta \in \{3, 4\}$. Rozdiel medzi dvoma mapami „ $\ominus$ “ je získaný pomocou interpolácie na jemnejšiu veľkosť a následným odčítaním. Na obrázku 21.7, je znázornená schéma modelu [IKN98].

Určovanie príznačov

Zo vstupného obrazu vo formáte RGB (r – červená, g – zelená, b – modrá) je získaná mapa pre intenzitu $I = (r + g + b)/3$, ktorá slúži ako vstup na vytvorenie Gaussovskej pyramídy $I(\delta)$, kde $\delta \in \{0, \dots, 8\}$. Kanály r, g, b sú normalizované. Následne sú vytvorené farebné kanály: $R = r - (g + b)/2$ pre červenú, $G = g - (r + b)/2$ pre zelenú, $B = b - (r + g)/2$ pre modrú a $Y = (r + g)/2 - |r - g|/2 - b$ pre žltú (záporné hodnoty sú nastavené na 0). Z týchto kanálov sú vytvorené štyri Gaussovské pyramídy $R(\delta)$, $G(\delta)$, $B(\delta)$ a $Y(\delta)$.

Neuróny cicavcov sú citlivé na náhlu zmenu intenzity ako napríklad tmavý stred obklopený svetlým okolím, alebo svetlý stred obklopený tma-

vým okolím. Oba spomenuté typy citlivosti na zmenu intenzity model prepája, čím vznikne 6 máp $I(c, s)$, kde $c \in \{2, 3, 4\}$ a $s = c + \delta$, kde $\delta \in \{3, 4\}$

$$I(c, s) = |I(c) \ominus I(s)|. \quad (21.1)$$

Druhou množinou sú mapy pre farebné kanály, ktoré sú v mozgovej kôre reprezentované pomocou systému „dvojíc oponentov“. V strede oblasti vnímania sú neuróny stimulované jednou farbou (napr. červená) a potláčané druhou farbou (napr. zelená), avšak v širšej oblasti je to naopak. Takto navzájom pôsobiace dvojice farieb sú červená/zelená, zelená/červená, modrá/žltá a žltá/modrá. Toto viedlo k vytvoreniu $RG(c, s)$ mapy pre červenú a zelenú farbu a $BY(c, s)$ mapy pre modrú a žltú farbu

$$RG(c, s) = |(R(c) - G(c)) \ominus (G(s) - R(s))|, \quad (21.2)$$

$$BY(c, s) = |(B(c) - Y(c)) \ominus (Y(s) - B(s))|. \quad (21.3)$$

Posledná množina máp pre orientáciu je získaná z I pomocou orientovaných Gaborových pyramíd $O(\delta, \theta)$, kde $\delta \in \{0, \dots, 8\}$ reprezentuje veľkosť a $\theta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ orientáciu. Mapy pre orientáciu predstavujú lokálne zmeny medzi stredom a okolím

$$O(c, s, \theta) = |O(c, \theta) \ominus O(s, \theta)|. \quad (21.4)$$

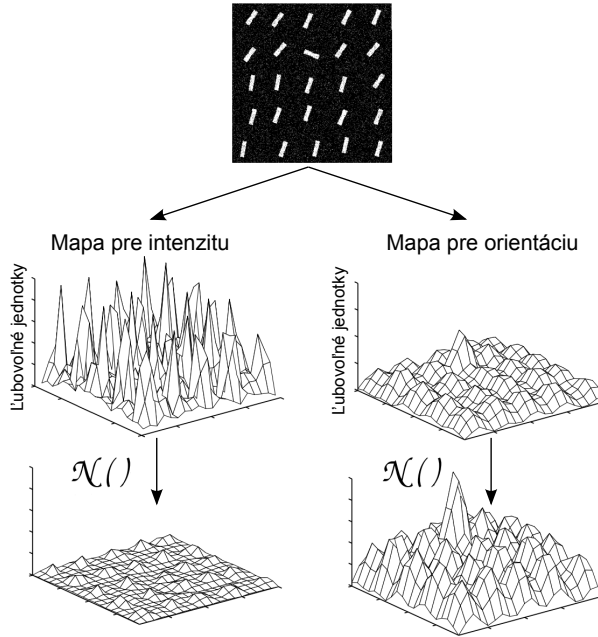
Pomocou týchto vzťahov je vytvorených 42 máp príznakov: šesť pre intenzitu, dvanásť pre farbu a dvadsaťštyri pre orientáciu.

Tvorba mapy významných oblastí

Pri kombinácii všetkých máp nastáva situácia, ktorá spôsobuje, že oblasti, ktoré sú veľmi výrazné iba v niektorých z kombinovaných máp, môžu zaniknúť. Preto treba uplatniť normalizačný operátor $\mathcal{N}()$, ktorého aplikácia pozostáva z troch krokov

1. normalizovanie hodnôt mapy na pevný rozsah v rozmedzí $\{0, \dots, M\}$ s cieľom odstrániť rozdiely medzi mapami,
2. nájdenie pozície lokálneho maxima M a vypočítanie priemernej hodnoty $\overline{m}$ všetkých týchto maxím,
3. globálne vynásobenie mapy $(M - \overline{m})^2$.

V prípade, že je v mape veľký rozdiel, najvýraznejšia oblasť vystúpi do popredia. Ak je rozdiel malý, mapa neobsahuje nič výrazné a je potlačená.

Obrázok 21.8: Normalizačný operátor $\mathcal{N}()$.

Mapy sú následne kombinované do troch „máp viditeľnosti“ pre intenzitu $\bar{I}$, farbu $\bar{C}$ a orientáciu $\bar{O}$ s veľkosťou $\delta = 4$. Tieto mapy sú získané pomocou sčítania „ $\oplus$ “, ktoré sa skladá z redukcie každej mapy na rovnakú veľkosť a následnom sčítaní pixel po pixeli.

Mapy pre intenzitu $\bar{I}$ a farbu $\bar{C}$ sú získané nasledovne

$$\bar{I} = \oplus_{c=2}^4 \oplus_{s=c+3}^{c+4} \mathcal{N}(I(c, s)), \quad (21.5)$$

$$\bar{C} = \oplus_{c=2}^4 \oplus_{s=c+3}^{c+4} [\mathcal{N}(RG(c, s)) + \mathcal{N}(BY(c, s))]. \quad (21.6)$$

Pre orientáciu sú najprv vytvorené štyri mapy kombináciou šiestich máp s danou orientáciou θ , ktoré sú následne sčítané

$$\bar{O} = \sum_{\theta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}} \mathcal{N} \left(\oplus_{c=2}^4 \oplus_{s=c+3}^{c+4} \mathcal{N}(O(c, s, \theta)) \right). \quad (21.7)$$

Tieto tri mapy sú normalizované a sčítané do výslednej mapy významných oblastí

$$S = \frac{1}{3} \left(\mathcal{N}(\bar{I}) + \mathcal{N}(\bar{C}) + \mathcal{N}(\bar{O}) \right). \quad (21.8)$$

Maximum výslednej mapy v danom čase ukazuje najvýznamnejšie miesto, na ktoré by mala byť zameraná pozornosť.

Ľudská vizuálna psychika funguje na princípe skúmania scény so zákazom vracat' sa na už navštívené oblasti. Tento fakt je využitý v danom modeli, kde je najprv spustená súbežná extrakcia príznakov, výstupom sú tri mapy kontrastov pre farbu $\bar{C}$, intenzitu $\bar{I}$ a orientáciu $\bar{O}$. Tieto mapy sú kombinované, čím vznikne mapa významných oblastí. Následne sa táto mapa prechádza a pomocou kroku IOR sa postupne určujú najvýznamnejšie oblasti.

V neurónovej implementácii, bol vytvorený model mapy významných oblastí (saliency map, SM) ako 2D vrstvy presakujúcich neurónov s vlastnosťou „integrate-and-fire“ a s veľkosťou štyri. Tieto neuróny majú kapacitu, ktorá zahŕňa určitú hodnotu a prah. Keď je daný prah dosiahnutý, vytvorí sa „hrot“ a kapacitný náboj je nastavený na nulu. Potom je SM daná ako 2D víťaz berie všetko (winner take all, WTA) neurónovej siete, v ktorej nájdená výrazná oblasť ostáva a ostatné sú potlačené.

Neuróny dostávajú vstupy z mapy S a sú všetky nezávislé. Potenciál neurónov v SM patriacich najvýraznejšej oblasti vzrastá najrýchlejšie. Každý SM neurón excituje jeho príslušný WTA neurón. Všetky WTA neuróny sa tiež vyvíjajú nezávisle od seba pokiaľ jeden („víťaz“) nedosiahne prahovú hodnotu. Toto vyvolá tri súbežné mechanizmy

1. oblasť pozornosti (focus of attention, FOA) je posunutá na miesto víťazného neurónu,
2. globálne sa spustí WTA a inicializuje na predvolenú hodnotu všetky neuróny,
3. v SM sa dočasne aktivuje lokálne potlačenie v oblasti s veľkosťou a umiestnením novej FOA. Takto je možná dynamická zmena FOA prostredníctvom povolenia nasledujúcej výraznej oblasti stať sa „víťazom“. Taktiež sa nevracame na už spracované FOA.

Tento model je výpočtovo nenáročný a má jednoduchú architektúru. Dokáže dobre detegovať oblasti vizuálnej pozornosti v scéne. Rýchlo určuje napríklad dopravné značky rôznych tvarov, farieb a textúru aj napriek tomu, že nebol navrhnutý na túto úlohu. Z výpočtového hľadiska je hlavnou výhodou tohto modelu paralelná implementácia nielen vo výpočtovo náročnej fáze extrakcie príznakov, ale aj v systéme zaostrovania na výrazné oblasti. Využitie neurónových sietí sa ukázalo ako účinné pri reprodukcii niektorých výkonov vizuálneho systému primátov. Jeho účinnosť však závisí od implementovaných príznakov.

21.3 Využitie detekcie významných oblastí v obraze

Detekcia významných oblastí v obraze a sledovanie pohybu očí majú široké využitie. Sledovanie pohybu očí sa využíva aj pri zisťovaní chorôb pacientov, ktoré sa môžu prejavovať napríklad pomalšími pohybmi očí. Detekcia významných oblastí v obraze má široké využitie v počítačovom videní, segmentácii obrazu, detekcii objektov, dynamickom osvetľovaní, interakcii medzi človekom a počítačom, kompresii obrazu, pri tvorbe náhľadov, retargetingu. Rovnako sa využíva pri navrhovaní reklamných plátov, dizajne webových stránok. V tejto časti si podrobne opíšeme niektoré spôsoby využitia detekcie významných oblastí v obraze.

21.3.1 Hľadanie najlepších pohľadov na objekty

V prípade virtuálnych múzeí a tvorbe ciest po týchto múzeách je veľmi užitočné nájsť najlepšie pohľady na zobrazované exponáty, pomocou ktorých môžeme čo najlepšie navrhnuť virtuálnu prechádzku daným múzeom. Jeden zo spôsobov hľadania najlepších pohľadov na objekty je založený na detekcii významných oblastí v obraze [KVC13]. Hlavná myšlienka spočíva v predpoklade, že čím viac detailov môžeme vidieť na danom objekte, tým je pohľad naň lepší. V prípade, že máme viacero možných pohľadov na daný objekt detegujeme významné oblasti na každom z nich. Následne použijeme segmentáciu obrazu pomocou povodí (watershed). Čím viac segmentov získame, tým je daný pohľad na objekt detailnejší.

21.3.2 Kompresia obrazu

Existujú dva hlavné spôsoby kompresie dát: stratová a bezstratová. Stratová kompresia sa používa v prípade informácií ako napríklad zvuk a obraz. Avšak pri veľmi veľkej kompresii strácame potrebnú informáciu. Tento problém rieši kompresia obrazu, ktorá využíva oblasti záujmu (region of interest ROI). Jej princíp spočíva v určení ROI a kódovaní týchto oblastí s väčšou kvalitou v porovnaní s ostatnými oblasťami v obraze. Existuje viacero spôsobov na určovanie týchto oblastí a jedným z nich je práve využitie detekcie významných oblastí v obraze. Jedným z najrozšírenejších spôsobov kompresie dát s použitím ROI je štandard JPEG 2000.

21.4 Príklad využitia: Hodnotenie kvality videa

Video je jedným z najrozšírenejších spôsobov zachytávania vizuálnej informácie. Pre väčšinu populácie slúži na relax, zábavu, avšak veľmi dôležitú úlohu zohráva pri komunikácii sluchovo hendikepovaných osôb. Keďže je vizuálna informácia rozhodujúca pre sluchovo hendikepovaných ľudí, umožňuje im komunikovať osobne používaním posunkovej reči, niektoré časti hovoriacej osoby sú dôležitejšie, ako iné. Preto je informácia o týchto vizuálne relevantných častiach veľmi dôležitá. S použitím informácie o výskyte týchto oblastí môžu byť navrhnuté objektívne metódy hodnotenia kvality videa, ktoré môžu pomôcť k skvalitneniu komunikácie nielen sluchovo hendikepovaných osôb.

Existuje veľké množstvo objektívnych metrík na hodnotenie kvality videa. Medzi takéto metriky patria napríklad SSIM (Structural Similarity), VQM (Video Quality Metrics), PSNR (Peak signal-to-noise Ratio). Všetky zo spomínaných metód však pracujú s obrazom ako s celkom. Existuje niekoľko modifikácií ako napríklad metrika 3SSIM, ktorá obraz rozdelí na základe textúry na 3 úrovne, kde každej z úrovní pridelí určitú váhu. Avšak ani toto rozdelenie nie je dostatočné.

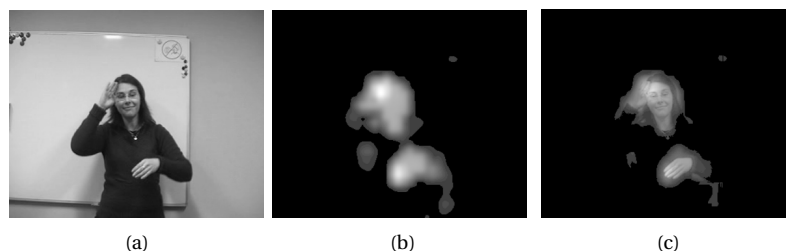
Zrozumiteľnosť posunkovej reči je veľmi dôležitá pre dorozumievanie sa sluchovo hendikepovaných osôb. Preto bola navrhnutá technika hodnotenia kvality videa, ktorá berie do úvahy informáciu o významných oblastiach v obraze. Pomocou tejto techniky je každému pixelu obrazu pridelená váha s hodnotou od 0 do 255. Toto zvýšenie počtu váh na pixel dáva veľmi dobré výsledky v porovnaní s inými technikami [KPT12].

21.4.1 Detekcia významných oblastí

Pre pozorovateľa sú veľmi dôležité zmeny v nízkoúrovňových príznakoch, ako sú farba, intenzita, textúra. Avšak zo sémantického hľadiska sa pozorovatelia zameriavajú na osoby, tváre, objekty. Pri komunikácii sluchovo hendikepovaných sú z tohto hľadiska najdôležitejšie ruky a tvár, ktoré je možné spoločne určiť pomocou detekcie pokožky.

V navrhnutom prístupe bol na detekciu významných oblastí použitý model, ktorý je založený na kombinácii príznakov farby, intenzity, textúry a detekcie pokožky. Základný model pozostáva z nasledovných krokov

1. Tvorba máp príznakov pre farbu, intenzitu, textúru a detekciu pokožky,
2. Tvorba mapy potlačenia,



Obrázok 21.9: Ukážka detekcie významných oblastí na účely hodnotenia kvality videa: (a) vstupná snímka, (b) mapa významných oblastí, (c) kombinácia Y kanála vstupnej snímky z modelu YUV s mapou významných oblastí.

3. Tvorba výslednej mapy významných oblastí.

V prvom kroku sa vytvoria mapy pre príznaky farby, intenzity, textúry a pre detekciu pokožky. Následne sa z týchto máp vytvorí mapa potlačenia a mapa kombinácií. Mapa kombinácií vzniká súčtom všetkých máp príznakov. Táto mapa obsahuje významné oblasti, avšak obsahuje aj nesprávne určené významné oblasti. Preto je vytvorená mapa potlačenia, ktorá potlačí tieto nesprávne významné oblasti. Kombináciou mapy potlačenia a mapy kombinácií vzniká výsledná mapa významných oblastí (obr. 21.9(b)).

21.4.2 Hodnotenie kvality videa

Hlavná myšlienka tohto prístupu spočíva v použití máp detegovaných významných oblastí na určenie váh pre hodnotenie kvality videa. Pomocou tejto mapy môže byť každá snímka videa rozdelená na 256 rôznych úrovní podľa významnosti.

Prístup bol testovaný na videách zachytávajúcích osobu komunikujúcu slovenskou znakovou rečou, ktoré mali rôzne úrovne kvality. Výsledky získané pomocou objektívnych metrík hodnotenia kvality videa sa pomocou Pearsonovho korelačného koeficientu porovnávali s hodnotením sluchovo hendikepovanými používateľmi, ktorých úlohou bolo pre každé video určiť mieru zrozumiteľnosti. Základné metriky, ako SSIM, VQM a PSNR boli modifikované použitím máp významných oblastí. Tieto nové metriky sú označené ako S-SSIM, S-VQM a S-PSNR. V tabuľke 21.1 sú znázornené výsledné hodnoty Pearsonovho korelačného koeficientu pre porovnávané techniky. Čím je jeho hodnota bližšia 1, tým je daná

metrika úspešnejšia. Z výsledkov zobrazených v tabuľke je zrejmé, že použitie máp detekcie významných oblastí na určenie váh pre každý pixel v obraze značne zvyšuje úspešnosť daných metrík.

Objektívna metrika	Pearsonov korelačný koeficient	Poradie
S-VQM	0,98447	1
S-SSIM	0,97869	2
VQM	0,97624	3
SSIM	0,964678	4
3SSIM	0,93052	5
S-PSNR	0,91398	6
PSNR	0,91	7

Tabuľka 21.1: Porovnanie objektívnych metrík na ohodnocovanie kvality videa pomocou Pearsonovho korelačného koeficientu



Zoznam použitej literatúry

- [Ach+09] R. Achanta a kol. „Frequency-tuned Salient Region Detection“. *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR 2009)*. 2009, str. 1597–1604.
- [AF13] A. Al-Rahayfeh a M. Faezipour. „Eye Tracking and Head Movement Detection: A State-of-Art Survey“. *IEEE Journal of Translational Engineering in Health and Medicine* 1 (2013).
- [AL10] T. Avraham a M. Lindenbaum. „Esaliency (Extended Saliency): Meaningful Attention Using Stochastic Image Modeling“. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32.4 (2010), str. 693–708.
- [BI13] A. Borji a L. Itti. „State-of-the-Art in Visual Attention Modeling“. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35.1 (2013), str. 185–207.
- [BT05] N. D. B. Bruce a J. K. Tsotsos. „Saliency Based on Information Maximization.“ *NIPS*. 2005.
- [BVM08] L. Busin, N. Vandenbroucke a L. Macaire. „Color spaces and image segmentation“. *Advances in Imaging and Electron Physics* 151 (2008), str. 65–168.
- [CCM03] A. Chalmers, K. Cater a D. Mafloli. „Visual Attention Models for Producing High Fidelity Graphics Efficiently“. *Proceedings of the 19th Spring Conference on Computer Graphics*. SCCG '03. Budmerice, Slovakia: ACM, 2003, str. 39–45.

- [CIE04] CIE. *15: Technical Report: Colorimetry, 3rd edition*. Tech. spr. International Commission on Illumination, 2004.
- [CIE13] CIE. *Slovenský národný komitét medzinárodnej komisie pre osvetľovanie*. 2013.
- [CK07] Y.-W. Chen a K. Kubo. „A Robust Eye Detection and Tracking Technique Using Gabor Filters“. *IIH-MSP*. IEEE, 2007, str. 109–112.
- [Dém13] A. Démuth. *Teórie percepcie*. Filozofická fakulta Trnavskej univerzity v Trnave, 2013.
- [Dha03] L. I. N. Dhavale. „Realistic avatar eye and head animation using a neurobiological model of visual attention“. *Proc. SPIE*. SPIE Press, 2003, str. 64–78.
- [Dob05] M. Dobeš. *Základy neuropsychológie*. Spoločenskovedný ústav SAV, Košice, 2005.
- [Ebn07] M. Ebner. *Color Constancy*. Wiley & Sons, Inc., 2007.
- [EE13] E. Erdem a A. Erdem. „Visual saliency estimation by nonlinearly integrating features using region covariances“. *Journal of Vision* 13.4 (2013), str. 1–20.
- [Fai05] M. D. Fairchild. *Color Appearance Models*. Wiley-IS&T, Chichester, UK, 2005.
- [FBH97] H. S. Fairman, M. H. Brill a H. Hemmendinger. „How the CIE 1931 color-matching functions were derived from Wright-Guild data“. *Color Research & Application* 22.1 (1997), str. 11–23.
- [FP02] D. A. Forsyth a J. Ponce. *Computer Vision: A Modern Approach*. Prentice Hall Professional Technical Reference, 2002.
- [FR98] A. Ford a A. Roberts. *Colour Space Conversions*. 1998.
- [Gev+12] T. Gevers a kol. *Color in Computer Vision: Fundamentals and Applications*. Wiley & Sons, Inc., 2012.
- [Gev01] T. Gevers. „Color-Based Retrieval“. *Advances in Computer Vision and Pattern Recognition*. Ed. M. S. Lew. Springer, 2001, str. 11–49.
- [Gol08] B. E. Goldstein. *Cognitive Psychology: Connecting Mind, Research and Everyday Experience*. Belmont: Thomson Wadsworth, 2008.

- [Gra53] H. Grassmann. „Theory of Compound Colors“. *Annalen der Physik und Chemie* 19 (1853), str. 53–60. Translation published in *Philosophical Magazine*, 4, pp. 254–264 (1854). Reprinted in *Selected Papers in Colorimetry – Fundamentals*, D. L. MacAdam, Editor, SPIE Milestone Series, Volume MS77, (1993).
- [GV04] D. Gao a N. Vasconcelos. „Discriminant saliency for visual recognition from cluttered scenes“. *In Proc. NIPS*. 2004, str. 481–488.
- [GW08] R. C. Gonzalez a R. E. Woods. *Digital Image Processing*. Treťie vyd. Prentice Hall, 2008.
- [GZT10] S. Goferman, L. Zelnik-manor a A. Tal. „Context-aware saliency detection“. *in IEEE Conf. on Computer Vision and Pattern Recognition*. 2010.
- [HKP07] J. Harel, C. Koch a P. Perona. „Graph-based visual saliency“. *ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS 19*. MIT Press, 2007, str. 545–552.
- [Hor89] P. Horňák. *Svetelná technika*. ALFA, 1989.
- [HRC05] Y. Hu, D. Rajan a L.-T. Chia. „Adaptive local context suppression of multiple cues for salient visual attention detection.“ *ICME*. IEEE, 2005, str. 346–349.
- [IKN98] L. Itti, C. Koch a E. Niebur. „A Model of Saliency-Based Visual Attention for Rapid Scene Analysis“. *IEEE Trans. Pattern Anal. Mach. Intell.* 20.11 (nov. 1998), str. 1254–1259.
- [Jud+09] T. Judd a kol. „Learning to Predict Where Humans Look“. *IEEE International Conference on Computer Vision (ICCV)*. 2009.
- [KBW08] T. Kocejko, A. Bujnowski a J. Wtorek. „Eye mouse for disabled“. *Human System Interactions, 2008 Conference on*. Máj 2008, str. 199–202.
- [Kie+09] W. Kienzle a kol. „Center-surround patterns emerge as optimal predictors for human saccade targets“. *en. Journal of Vision* 9.5:7 (máj 2009), str. 1–15.
- [KPT12] J. KucEROVA, J. Polec a D. Tarcsiova. „Video Quality Assessment using Visual Attention Approach for Sign Language“. 6.5 (2012), str. 181–186.

- [KVC13] J. Kucerova, I. Varhanikova a Z. Cernekova. „Best View Methods Suitability for Different Types of Objects“. *Proceedings of the 28th Spring Conference on Computer Graphics. SCCG '12*. Budmerice, Slovakia: ACM, 2013, str. 55–61.
- [Mus10] V. Muszková. „Fyziologické a Psychologické Princípy Zrakovej Percepcie“. Diplomová práca. Masarykova univerzita Brno, Lékařská fakulta, 2010.
- [Oic07] B. Oicherman. „Effects of colorimetric additivity failure and of observer metamerism on cross-media colour matching“. Dizertačná práca. The University of Leeds UK, Department of Colour Science, 2007.
- [Oli+03] A. Oliva a kol. *Top-Down Control of Visual Attention in Object Detection*. 2003.
- [Poy97] C. Poynton. *Frequently Asked Questions about Color*. 1997.
- [PP10] M. Petrou a C. Petrou. *Image Processing: The Fundamentals*. Druhé vyd. Wiley & Sons, Inc., 2010.
- [Pra07] W. K. Pratt. *Digital Image Processing*. Štvrté vyd. Wiley & Sons, Inc., 2007.
- [Rao+02] R. P. Rao a kol. „Eye movements in iconic visual search“. *Vision Research* 42 (2002), str. 1447–1463.
- [RC02] O. Ramström a H. I. Christensen. „Visual Attention Using Game Theory.“ *Biologically Motivated Computer Vision*. Ed. H. H. Bülthoff a kol. Zv. 2525. Lecture Notes in Computer Science. Springer, 2002, str. 462–471.
- [Reb10] Rebrová, K. and Takáč, M. „Vnímanie a pomenovávanie farieb a farebných kategórií“. *Umelá inteligencia a kognitívna veda II*. Ed. V. Kvasnička a kol. Bratislava, 2010, str. 411–436.
- [RF95] E. Ružický a A. Ferko. *Počítačová Grafika a Spracovanie Obrazu*. Sapientia, 1995.
- [Ros09] P. L. Rosin. „A simple method for detecting salient regions.“ *Pattern Recognition* 42.11 (2009), str. 2363–2371.
- [Rus11] J. C. Russ. *The Image Processing Handbook*. 6. vyd. CRC Press, 2011.
- [SB59] W. S. Stiles a J. M. Burch. „N.P.L. colour-matching investigation: final report (1958)“. *Optica Acta: International Journal of Optics* 6.1 (1959), str. 1–26.

- [SH11] R. Shapley a M. J. Hawken. „Color in the Cortex: single- and double-opponent cells“. *Vision Research* 51.7 (2011), str. 701–717.
- [SHB07] M. Sonka, V. Hlavac a R. Boyle. *Image Processing, Analysis, and Machine Vision*. Tretie vyd. Cengage Learning, 2007.
- [Ska93] V. Skala. *Světlo, barvy a barevné systémy v počítačové grafice*. Advances in Computer Vision and Pattern Recognition. Academia Praha, 1993.
- [ŠMS11] A. Štrba, V. Mesároš a D. Senderáková. *SVETLO – vlny, lúče, fotóny*. Enigma publishing s.r.o., Nitra, 2011.
- [SS01] L. G. Shapiro a G. C. Stockman. *Computer Vision*. Prentice Hall, 2001.
- [Tra91] D. Travis. *Effective Color Displays. Theory and Practice*. Academic Press, 1991.
- [Tso11] J. K. Tsotsos. *A Computational Perspective on Visual Attention*. MIT Press, 2011, str. I–XVI, 1–308.
- [Umb10] S. E. Umbaugh. *Digital Image Processing and Analysis: Human and Computer Vision Applications with CVIPtools*. Druhé vyd. CRC Press, 2010.
- [VV11] M. Vik a M. Viková. „Odchylky od platnosti Grassmanových zákonů – problém současné kolorimetrie?“. *Světlo* 1 (2011), str. 52–54.
- [WS00] G. Wyszecki a W. S. Stiles. *Color Science, Concepts and Methods, Quantitative Data and Formulae*. Druhé vyd. Wiley & Sons, Inc., 2000.
- [Žár+04] J. Žára a kol. *Moderní počítačová grafika*. Druhé vyd. Computer Press, Brno, 2004.
- [Zha+08] L. Zhang a kol. „SUN: A Bayesian framework for saliency using natural statistics.“ *J Vis* 8.7 (2008), str. 32.1–20.

ČASŤ V

DODATKY

Vybrané témy z pravdepodobnosti

A.1 Pravdepodobnosť

Pravdepodobnosť môžeme definovať rôznymi spôsobmi.

Klasická (Laplaceova) definícia. $P(A) = \frac{N_A}{N}$. Pravdepodobnosť je daná relatívnou početnosťou výsledku A v N pokusoch.

Limitná definícia. $P(A) = \lim_{N \rightarrow \infty} \frac{N_A}{N}$. Pravdepodobnosť intuitívne chápeme ako relatívnu početnosť v nekonečne veľkom počte pokusov.

Axiomatická (Kolmogorovova) definícia. Pole javov $\mathcal{A}$ je σ -algebra na Ω , t. j. systém podmnožín priestoru Ω obsahujúci Ω a uzavretý voči doplnku a zjednoteniu

$$\Omega \in \mathcal{A}, \quad (\text{A.1})$$

$$A \in \mathcal{A} \Rightarrow A^C \in \mathcal{A}, \quad (\text{A.2})$$

$$(\forall n \in \mathbb{N} : A_n \in \mathcal{A}) \Rightarrow \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}. \quad (\text{A.3})$$

Pravdepodobnosť je funkcia $P : \mathcal{A} \rightarrow \langle 0, 1 \rangle$, ktorá spĺňa nasledovné vlastnosti

$$P(\Omega) = 1, \quad (\text{A.4})$$

$$P\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} P(A_n), \quad (\text{A.5})$$

ak javy A_n sú po dvoch disjunktné.

Usporiadaná trojica $(\Omega, \mathcal{A}, P)$ tvorí tzv. pravdepodobnostný priestor. Funkcia P môže predstavovať mieru na σ -algebre $\mathcal{A}$. Ostatné definície sú špeciálnymi prípadmi axiomatickej definície, v praxi sa však pri výpočte pravdepodobnosti často využívajú.

Množina všetkých možných výstupov náhodnej premennej sa nazýva výberový priestor. Jednotlivé (rovnako pravdepodobné) prvky sa nazývajú elementárne javy. Každú podmnožinu výberového priestoru nazveme (náhodný) jav. Jav môže byť diskretný (napr. hod kockou alebo mincou, počet detí v rodine...) alebo spojitý (napr. rozmer nejakého výrobku, výška hráčov basketbalového mužstva...).

Nech A je jav reprezentujúci hod kockou. Potom výberový priestor je $1, 2, \dots, 6$ a elementárne javy sú počty bodov na stranách kocky. Javom môže byť „padne párne číslo“.

Rozkladom výberového priestoru nazývame súbor vzájomne disjunktívnych neprázdnych javov, ktorých zjednotením je celý výberový priestor.

Opačný jav k javu A , t. j. A^C , sa nazýva aj jav doplnkový alebo komplementárny. Zrejme A a A^C je najjednoduchší rozklad. Pre každý jav platí $P(A^C) = 1 - P(A)$.

Prienik $A \cap B$ nastáva práve vtedy, keď nastanú oba javy A a B . Pravdepodobnosť $P(A \cap B)$ sa nazýva vzájomná alebo združená pravdepodobnosť a označuje sa aj $P(A, B)$.

Zjednotenie $A \cup B$ nastáva práve vtedy, keď nastane aspoň jeden z javov A a B .

Vlastnosti pravdepodobnosti

Nech $A, B \in \mathcal{A}$, potom

$$P(\emptyset) = 0 \quad (\text{A.6})$$

$$P(A^C) = 1 - P(A) \quad (\text{A.7})$$

$$A \subseteq B \Rightarrow P(A) \leq P(B) \quad (\text{A.8})$$

$$P(B \setminus A) = P(B) - P(A \cap B) \quad (\text{A.9})$$

$$A \subset B \Rightarrow P(B \setminus A) = P(B) - P(A) \quad (\text{A.10})$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (\text{A.11})$$

$$A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B) \quad (\text{A.12})$$

$$P(A \cup B) = 1 - P((A \cup B)^C) = 1 - P(A^C \cap B^C) \quad (\text{A.13})$$

$$P(A \cap B) = 1 - P((A \cap B)^C) = 1 - P(A^C \cup B^C) \quad (\text{A.14})$$

A.1.1 Nezavislosť javov, podmienená pravdepodobnosť

Podmienená pravdepodobnosť $P(A|B)$ je definovaná vzťahom

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad (\text{A.15})$$

kde $P(B) \neq 0$. Podmienená pravdepodobnosť hovorí o pravdepodobnosti, že nastane jav A za podmienky, že nastal jav B . Z tejto definície vyplýva $P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$.

Javy A a B nazveme nezávislé, ak platí $P(A \cap B) = P(A) \cdot P(B)$. Pre nezávislé javy preto platí $P(A|B) = P(A)$. To znamená, že to, či nastane jav B neovplyvní pravdepodobnosť, že nastane jav A . A platí to aj naopak $P(B|A) = P(B)$.

Príklad 1.1: Aká je pravdepodobnosť, že pri hode mincou padne hlava dvakrát za sebou?

Pri férovej minci je pravdepodobnosť, že padne hlava 0,5. Výsledok druhého hodu (B) je nezávislý od výsledku prvého hodu (A). Preto výsledná pravdepodobnosť $P(A \cap B) = P(A) \cdot P(B) = 0,5 \cdot 0,5 = 0,25$. ■

Príklad 1.2: Z krabice, v ktorej sa nachádza 10 červených a 5 modrých ponožiek, vytiahneme naslepo postupne dve ponožky. Aká je pravdepodobnosť, že sme ako druhú vytiahli modrú ponožku?

V tomto prípade sa nejedná o nezávislé javy, lebo vytiahnutím prvej ponožky sa zmení počet kusov v krabici a tým aj pravdepodobnosť vytiahnutia druhej ponožky. Označme si A jav, keď bola vytiahnutá ako prvá červená a B , keď modrá ponožka. Jav C je vytiahnutie modrej ponožky v druhom ťahu. Potom platí $P(A) = 10/15$ a $P(B) = 5/15$. Podmienené pravdepodobnosti sú nasledovné $P(C|A) = P(C \cap A)/P(A)$ a $P(C|B) = P(C \cap B)/P(B)$. Po vytiahnutí prvej ponožky zostane v krabici už len 14 kusov. Keď sme vytiahli červenú ponožku, zostalo 9 červených a 5 modrých ponožiek, a preto $P(C|A) = 5/14$. Keď sme vytiahli modrú, zostalo 10 červených a 4 modré ponožky, preto $P(C|B) = 4/14$. Výsledná pravdepodobnosť, že v druhom ťahu vytiahneme modrú ponožku je $P(C) = P((C \cap A) \cup (C \cap B))$, pretože buď vytiahneme červenú a modrú ($C \cap A$) alebo ($\cup$) dve modré ($C \cap B$). Tieto javy sú nezlučiteľné (nemôžu nastať oba naraz), a preto $P(C) = P(C \cap A) + P(C \cap B)$. Z definície podmienenej pravdepodobnosti potom vyplýva

$$P(C) = P(C|A) \cdot P(A) + P(C|B) \cdot P(B) = \frac{5}{14} \cdot \frac{10}{15} + \frac{4}{14} \cdot \frac{5}{15} = \frac{7}{21} = \frac{1}{3}. \quad (\text{A.16})$$

Teda pravdepodobnosť, že v druhom ťahu vytiahneme modrú ponožku, je $1/3$. ■

Nech $\{A_i\}_{i=1}^M$ je rozkladom priestoru Ω (teda $\sum_{i=1}^M P(A_i) = 1$). Potom pravdepodobnosť javu B je daná

$$P(B) = \sum_{i=1}^M P(B|A_i)P(A_i). \quad (\text{A.17})$$

Je zrejmé, že ľubovoľný jav $B \in \Omega$ je zjednotením javov $(B \cap A_1), \dots, (B \cap A_M)$, čiže $B = \bigcup_{i=1}^M (B \cap A_i)$. Tieto javy sú disjunktné, preto $P(B) = \sum_{i=1}^M P(B \cap A_i)$. Z definície podmienenej pravdepodobnosti (A.15) potom vyplýva (A.17).

V reálnych aplikáciách nás často zaujíma, ktorý z javov $\{A_i\}_{i=1}^M$ viedol k javu B . Hľadáme teda $P(A_k|B)$. Z (A.15) a (A.17) vieme, že

$$P(A_k|B) = \frac{P(B|A_k) \cdot P(A_k)}{P(B)} = \quad (\text{A.18})$$

$$= \frac{P(B|A_k) \cdot P(A_k)}{\sum_{i=1}^M P(B|A_i)P(A_i)}. \quad (\text{A.19})$$

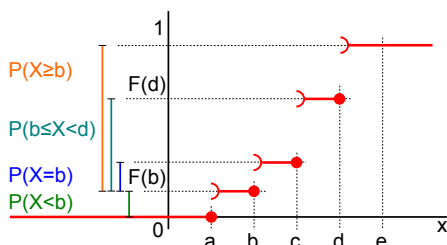
Vzťah (A.19) sa nazýva **Bayesov vzorec** a udáva, ako zistíme aposteriórnu pravdepodobnosť javu A_k za predpokladu, že nastal jav B ($P(A_k|B)$), ak pre javy $\{A_i\}_{i=1}^M$ poznáme apriórnu pravdepodobnosť javov ($P(A_i)$) a podmienené pravdepodobnosti $P(B|A_i)$.

Príklad 1.3: Na konci výrobnéj linky je tester, ktorý s pravdepodobnosťou 0,999 vyradí chybný výrobok. S pravdepodobnosťou 0,01 však vyradí aj kvalitný výrobok. Kazivosť výrobkov je 0,1 (jeden z desiatich výrobkov je chybný). Aká je pravdepodobnosť, že vyradený výrobok je skutočne chybný?

Priestor výrobkov rozložíme na chybné $P(A_{ch}) = 0,1$ a kvalitné $P(A_{kv}) = 0,9$. Vyradenie výrobku je jav B , pričom podmienené pravdepodobnosti sú $P(B|A_{ch}) = 0,999$ a $P(B|A_{kv}) = 0,01$. Pomocou Bayesovho vzorca hľadáme pravdepodobnosť

$$\begin{aligned} P(A_{ch}|B) &= \frac{P(B|A_{ch}) \cdot P(A_{ch})}{P(B|A_{ch})P(A_{ch}) + P(B|A_{kv})P(A_{kv})} = \\ &= \frac{0,999 \cdot 0,1}{0,999 \cdot 0,1 + 0,01 \cdot 0,9}. \end{aligned} \quad (\text{A.20})$$

Pravdepodobnosť, že vyradený výrobok je skutočne chybný, je približne 0,9174. ■



Obrázok A.1: Distribučná funkcia diskkrétnej premennej.

A.2 Náhodná premenná

Pojmom náhodná premenná (veličina) označujeme funkciu, ktorej hodnota je určená výsledkom náhodného pokusu. Prirad'uje číselnú hodnotu každému elementárnemu javu.

Pre náhodnú premennú $X : \Omega \rightarrow \mathbb{R}$ platí, že množina $\{\omega \in \Omega | X(\omega) < x\}$ je náhodný jav.

Hodnota funkcie sa nedá pred uskutočnením pokusu určiť, ale dokážeme ju opísať rozdelením pravdepodobnosti. Rozdelenie pravdepodobnosti náhodnej premennej definovanej na pravdepodobnostnom priestore $(\Omega, \mathcal{A}, P)$ opisujeme pomocou distribučnej funkcie $F : \mathbb{R} \rightarrow \langle 0, 1 \rangle$ definovanej ako

$$F(x) = P(X < x). \quad (\text{A.21})$$

$P(X < x)$ znamená pravdepodobnosť, s akou náhodná premenná X nadobúda hodnotu menšiu ako x . Distribučná funkcia je neklesajúca a zľava spojitá a platí $\lim_{x \rightarrow -\infty} F(x) = 0$ a $\lim_{x \rightarrow \infty} F(x) = 1$.

Náhodná premenná môže byť diskrétna alebo spojitá. Distribučná funkcia pre diskrétnu premennú je „schodovitá“ funkcia ako na obr. A.1. Príklad distribučnej funkcie spojitaj premennej je na obr. A.2.

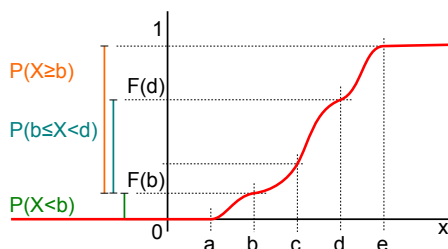
Vzťahy medzi pravdepodobnosťou a distribučnou funkciou sú nasledovné (ako vidno na obrázkoch A.1a A.2)

$$\forall a < b; a, b \in \mathbb{R}$$

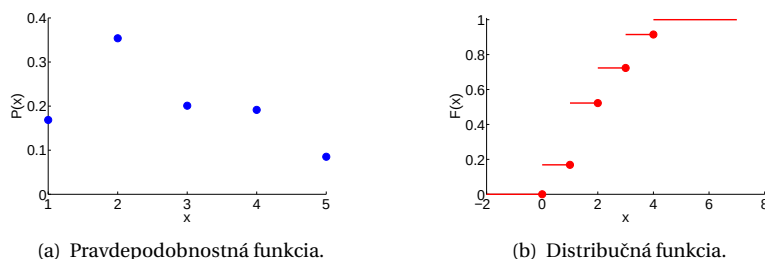
$$P(X \geq a) = 1 - F(a) = \int_a^\infty f(x) dx \quad (\text{A.22})$$

$$P(a \leq X < b) = F(b) - F(a) = \int_a^b f(x) dx \quad (\text{A.23})$$

$$P(X = a) = \begin{cases} \lim_{x \rightarrow a^+} F(x) - F(a), & \text{diskrétna premenná} \\ 0, & \text{spojitá premenná.} \end{cases} \quad (\text{A.24})$$



Obrázok A.2: Distribučná funkcia spojitkej premennej.



Obrázok A.3: Príklad pravdepodobnostnej a zodpovedajúcej distribučnej funkcie diskkrétnej premennej.

Distribučnú funkciu diskkrétnej premennej určíme ako

$$F(x) = \sum_{x_i < x} P(x_i) \quad (\text{A.25})$$

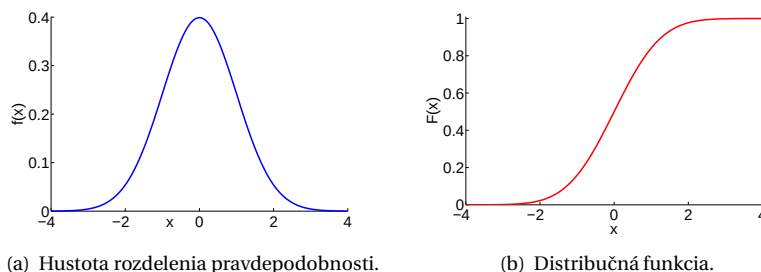
a pre spojitú premennú

$$F(x) = \int_{-\infty}^x f(t) dt, \quad (\text{A.26})$$

kde $f(t)$ je hustota rozdelenia pravdepodobnosti (ekvivalent pravdepodobnostnej funkcie $P(x)$). Hustota pravdepodobnosti je nezáporná funkcia a platí $\int_{-\infty}^{\infty} f(x) dx = 1$, $\lim_{x \rightarrow -\infty} f(x) = 0$ a $\lim_{x \rightarrow \infty} f(x) = 0$. Tieto vzťahy opisujú obrázky A.3 a A.4.

A.3 Číselné charakteristiky náhodnej premennej

Charakterizujú vlastnosti náhodnej premennej a umožňujú porovnávať premenné. V ďalšom texte opíšeme len charakteristiky spojitých premenných. Pre diskkrétne premenné platia obdobné vzťahy.



Obrázok A.4: Vzťah hustoty rozdelenia pravdepodobnosti a zodpovedajúcej distribučnej funkcie spojitej premennej.

A.3.1 Momenty

Začiatkový moment r -tého rádu je definovaný ako

$$\mu_r = E(X^r) = \int_{-\infty}^{\infty} x^r f(x) dx. \quad (\text{A.27})$$

Okrem začiatkových momentov sa používajú aj centrálné momenty, kde od hodnôt pôvodných dát odčítame ich priemer

$$\mu'_r = E((X - E(X))^r) = \int_{-\infty}^{\infty} (x - E(X))^r f(x) dx. \quad (\text{A.28})$$

Vyžívajú sa najmä centrálné momenty druhého, tretieho a štvrtého rádu.

Stredná hodnota

Vyjadruje hodnotu, „okolo“ ktorej sa sústred'ujú všetky hodnoty náhodnej premennej. Nazýva sa aj priemer (expected value, mean), značí sa $E(X)$ alebo μ

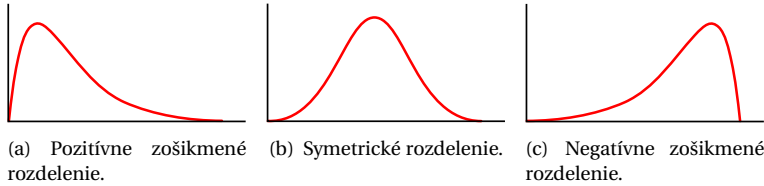
$$\mu = \int_{-\infty}^{\infty} x f(x) dx. \quad (\text{A.29})$$

Vidíme, že začiatkový moment prvého rádu je vlastne priemer.

Rozptyl

Rozptyl (tiež variancia, disperzia, stredná kvadratická odchýlka) je centrálny moment druhého rádu a vyjadruje variabilitu hodnôt dát okolo strednej hodnoty. Označuje sa σ^2 alebo $V(X)$, prípadne $\text{Var}(X)$

$$\text{Var}(X) = \int_{-\infty}^{\infty} (x - E(X))^2 f(x) dx. \quad (\text{A.30})$$



Obrázok A.5: Šikmosť.

Pri praktických výpočtoch sa používa vzťah

$$\text{Var}(X) = E(X^2) - (E(X))^2. \quad (\text{A.31})$$

Smerodajná (alebo štandardná) odchýlka označovaná σ alebo $\text{Std}(X)$ je daná druhou odmocninou rozptylu

$$\text{Std}(X) = \sqrt{\text{Var}(X)}. \quad (\text{A.32})$$

Šikmosť

Podiel tretieho centrálneho momentu a tretej mocniny smerodajnej odchýlky vyjadruje mieru symetrie rozdelenia pravdepodobnosti vzhľadom na normované (štandardné) normálne rozdelenie $N(0, 1)$ (pozri kapitolu A.6)

$$\alpha_3 = \frac{\mu'_3}{\sigma^3}. \quad (\text{A.33})$$

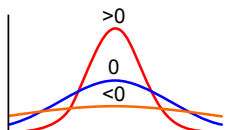
Podľa hodnoty, akú nadobúda α_3 hovoríme o symetrickom ($\alpha_3 = 0$), negatívne ($\alpha_3 < 0$) alebo pozitívne ($\alpha_3 > 0$) zošikmenom rozdelení (obr. A.5).

Špicatosť

Podiel štvrtého centrálneho momentu a štvrtej mocniny smerodajnej odchýlky vyjadruje relatívnu plochosť rozdelenia pravdepodobnosti vzhľadom na normované (štandardné) normálne rozdelenie $N(0, 1)$

$$\alpha_4 = \frac{\mu'_4}{\sigma^4}. \quad (\text{A.34})$$

Pre $N(0, 1)$ je $\alpha_4 = 3$, preto sa používa normovaná hodnota $\alpha_4 - 3$ a podľa tejto hodnoty hovoríme o normálnej (0), menšej (< 0) alebo väčšej (> 0) špicatosti rozdelenia (obr. A.6).



Obrázok A.6: Normálna (0), menšia (< 0) a väčšia (> 0) špicatosť rozdelenia.

A.3.2 Kvantily

Kvantil Q_p je hodnota, pre ktorú platí, že pravdepodobnosť, že náhodná veličina X má hodnotu menšiu ako Q_p , je $100p$ %

$$P(X < Q_p) = p. \quad (\text{A.35})$$

Kvantily tvoria vlastne inverznú funkciu k distribučnej funkcii. Pre spojité rozdelenie je kvantil Q_p číslo, pre ktoré platí

$$F(Q_p) = p. \quad (\text{A.36})$$

Hodnoty kvantilov jednotlivých rozdelení nájdeme v štatistických tabuľkách. Hodnoty charakterizujú polohu dát. Ich rozdiel (rozpätie) používame na charakteristiku variability.

Niektoré dôležité kvantily majú vlastné mená.

Medián

Medián je kvantil deliaci súbor na dve polovice. Je to teda $Q_{0,5}$. Znamená to, že polovica hodnôt dát je menšia a polovica väčšia ako medián.

Kvartil

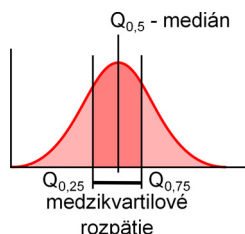
Kvartily delia súbor na štvrtiny. Sú to $Q_{0,25}$ (dolný kvartil), $Q_{0,5}$ a $Q_{0,75}$ (horný kvartil). Rozdiel horného a dolného kvartilu určuje medzikvartilové rozpätie $Q_{0,75} - Q_{0,25}$, ktoré je znázornené na obrázku A.7.

Decil

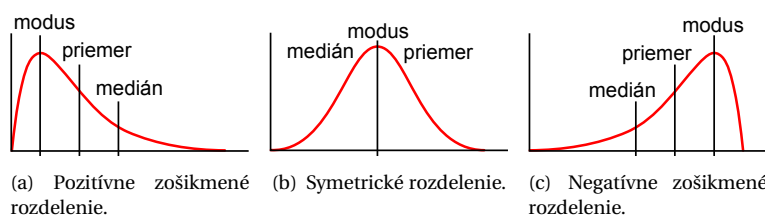
Decily delia súbor na desatiny. Sú to $Q_{0,1}, Q_{0,2}, \dots, Q_{0,9}$. Medzidecilové rozpätie je rozdiel $Q_{0,9} - Q_{0,1}$.

Percentil

Percentily delia súbor na stotiny. Sú to $Q_{0,01}, Q_{0,02}, \dots, Q_{0,99}$. Medzipercentilové rozpätie je rozdiel $Q_{0,99} - Q_{0,01}$.



Obrázok A.7: Medzikvartilové rozpätie.



Obrázok A.8: Modus, medián a priemer pre rôzne zošikmené rozdelenia pravdepodobnosti.

A.3.3 Modus

Modus je najčastejšie sa opakujúca hodnota v súbore. Vzt'ah medzi priemerom, modusom a mediánom je znázornený na obrázku A.8.

A.4 Náhodné vektory

Náhodnú veličinu chápeme ako výsledok náhodného pokusu. Takýto výsledok však nemusí byť opísaný jedným číslom, ale usporiadanou n -ticou čísiel. Príkladom je kontrola produkcie v potravinárskom priemysle, kde napríklad pri mäsových výrobkoch sa pri meraní zisťuje množstvo soli, obsah vody, obsah tuku, obsah dusitanov. Čiže každé meranie je reprezentované usporiadanou štvoricou.

Podobne ako rozdelenie pravdepodobnosti náhodnej premennej definujeme združené rozdelenie pravdepodobnosti náhodného vektora. Pre N -rozmerný vektor $(X_1, \dots, X_N)$ platí

$$\begin{aligned} F(x_1, \dots, x_N) &= P((X_1 < x_1) \& \dots \& (X_N < x_N)) = \\ &= P(X_1 < x_1, \dots, X_N < x_N). \end{aligned} \quad (\text{A.37})$$

Združenú distribučnú funkciu diskrétného náhodného vektora určíme ako

$$F(x_1, \dots, x_N) = \sum_{X_1 < x_1} \dots \sum_{X_N < x_N} P(x_1, \dots, x_N) \quad (\text{A.38})$$

a pre spojité premenné

$$F(x_1, \dots, x_N) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_N} f(t_1, \dots, t_N) dt_1 \dots dt_N, \quad (\text{A.39})$$

kde $f(t_1, \dots, t_N)$ je hustota rozdelenia pravdepodobnosti (ekvivalent pravdepodobnostnej funkcie $P(x_1, \dots, x_N)$).

A.4.1 Marginálne rozdelenie pravdepodobnosti

Pri náhodných vektoroch (v oblasti počítačového videnia a klasifikácie sú to vektory príznakov) potrebujeme často poznať aj marginálne rozdelenie jednotlivých zložiek vektora. Marginálne rozdelenie zložky i vypočítame pomocou združenej hustoty pravdepodobnosti nasledovne

$$\begin{aligned} F_i(x_i) &= P(X_1 < \infty, \dots, X_i < x_i, \dots, X_N < \infty) = \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{x_i} \dots \int_{-\infty}^{\infty} f(t_1, \dots, t_N) dt_1 \dots dt_N. \end{aligned} \quad (\text{A.40})$$

Zložky náhodného vektora sú vzájomne nezávislé, práve vtedy keď

$$F(x_1, \dots, x_N) = F_1(x_1) \dots F_N(x_N). \quad (\text{A.41})$$

A.5 Číselné charakteristiky náhodného vektora

Aj náhodné vektory opisujeme číselnými charakteristikami. Rodel'ujeme ich na skalárne, vektorové a maticové charakteristiky.

A.5.1 Združené momenty

Obecný združený moment $(r_1, \dots, r_N)$ -tého rádu je definovaný ako

$$\begin{aligned} \mu_{r_1, \dots, r_N} &= E(X_1^{r_1} X_2^{r_2} \dots X_N^{r_N}) = \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} x_1^{r_1} \dots x_N^{r_N} f(x_1, \dots, x_N) dx_1 \dots dx_N. \end{aligned} \quad (\text{A.42})$$

Keď od hodnôt dát odčítame priemer, môžeme vypočítať centrálny združený moment $(r_1, \dots, r_N)$ -tého rádu definovaný ako

$$\begin{aligned}\mu'_{r_1, \dots, r_N} &= E((X_1 - E(X_1))^{r_1} (X_2 - E(X_2))^{r_2} \dots (X_N - E(X_N))^{r_N}) = \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (x_1 - E(X_1))^{r_1} \dots (x_N - E(X_N))^{r_N} f(x_1, \dots, x_N) dx_1 \dots dx_N.\end{aligned}\quad (\text{A.43})$$

Príkladom použitia momentov v 2D úlohách počítačového videnia sú škálovo invariantné momenty (i, j) -tého rádu

$$\eta_{i,j} = \frac{\mu_{i,j}}{\mu_{0,0}^{\left(1 + \frac{i+j}{2}\right)}} \quad (\text{A.44})$$

a Huove momenty, ktoré sú škálovo, translačne aj rotačne invariantné

$$\begin{aligned}I_1 &= \eta_{2,0} + \eta_{0,2} \\ I_2 &= (\eta_{2,0} - \eta_{0,2})^2 + 4\eta_{1,1}^2 \\ &\dots\end{aligned}\quad (\text{A.45})$$

Zvyšné Huove momenty sú uvedené v [Hu62].

A.5.2 Marginálne charakteristiky

Marginálne charakteristiky opisujú jednotlivé zložky náhodného vektora. Opíšeme si strednú hodnotu a rozptyl, momenty vyšších rádov získame obdobne.

Stredná hodnota

Stredná hodnota $E(\mathbf{X})$ náhodného vektora $\mathbf{X}$ je vektor stredných hodnôt jednotlivých zložiek

$$E(\mathbf{X}) = (E(X_1), \dots, E(X_N)). \quad (\text{A.46})$$

Stredná hodnota $E(\mathbf{X})$ je vlastne vektorom začiatočných združených momentov $(r_1, \dots, r_N)$ -tého rádu, kde jeden z indexov r_i je rovný 1 a ostatné sú nulové.

Rozptyl

Podobne, rozptyl $\text{Var}(\mathbf{X})$ náhodného vektora $\mathbf{X}$ je vektor rozptylov jednotlivých zložiek

$$\text{Var}(\mathbf{X}) = (\text{Var}(X_1) \dots \text{Var}(X_N)). \quad (\text{A.47})$$

Rozptyl $\text{Var}(\mathbf{X})$ je vlastne vektorom centrálnych združených momentov $(r_1, \dots, r_N)$ -tého rádu, kde jeden z indexov r_i je rovný 2 a ostatné sú nulové.

A.5.3 Kovariancia

Vzťah medzi dvomi náhodnými veličinami X_i, X_j určuje kovariancia

$$\text{Cov}(X_i, X_j) = E((X_i - E(X_i))(X_j - E(X_j))), \quad (\text{A.48})$$

a teda centrálny moment $(r_1, \dots, r_N)$ -tého rádu, kde indexy $r_i = r_j = 1$ a ostatné sú nulové.

Kladná kovariancia hovorí, že hodnoty náhodných premenných sa pohybujú rovnakým smerom (zvýšením jednej sa zvýši aj druhá).

Z vlastností kovariancie vyplýva aj vzťah, ktorý sa používa na praktický výpočet jej hodnoty

$$\text{Cov}(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j). \quad (\text{A.49})$$

Ďalej platí

$$\text{Cov}(X_i, X_i) = E(X_i^2) - (E(X_i))^2 = \text{Var}(X_i) \quad (\text{A.50})$$

a

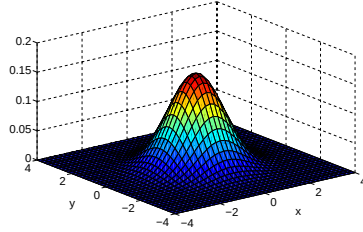
$$\text{Cov}(X_i, X_j) = \text{Cov}(X_j, X_i). \quad (\text{A.51})$$

Na opis vzťahov jednotlivých zložiek náhodného vektora $\mathbf{X}$ sa používa kovariančná matica Σ , kde

$$\Sigma_{ij} = \text{Cov}(X_i, X_j) = E((X_i - \mu_i)(X_j - \mu_j)), \quad (\text{A.52})$$

a teda

$$\Sigma = \begin{bmatrix} \text{Var}(X_1) & \text{Cov}(X_1, X_2) & \cdots & \text{Cov}(X_1, X_N) \\ \text{Cov}(X_2, X_1) & \text{Var}(X_2) & \cdots & \text{Cov}(X_2, X_N) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(X_N, X_1) & \text{Cov}(X_N, X_2) & \cdots & \text{Var}(X_N) \end{bmatrix}. \quad (\text{A.53})$$



Obrázok A.9: Dvojrozmerné normálne rozdelenie so nulovým priemerom a jednotkovou kovariančnou maticou.

A.6 Normálne rozdelenie

V oblasti spracovania obrazu, počítačového videnia a rozpoznávania objektov pracujeme väčšinou s predpokladom, že dáta majú normálne rozdelenie.

Hovoríme, že náhodná premenná má normálne (Gaussove) rozdelenie s parametrami μ a σ ($X \sim N(\mu, \sigma)$), keď pre jej funkciu hustoty platí

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \quad (\text{A.54})$$

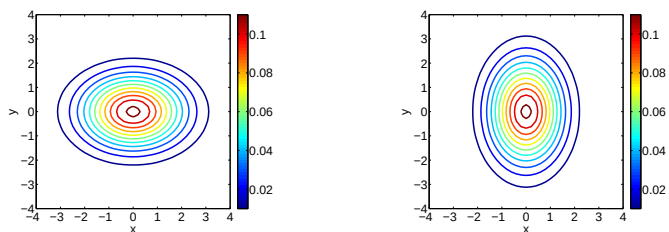
A.6.1 Viacrozmerné normálne rozdelenie

Zápis $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ znamená, že N -rozmerný vektor $\mathbf{X} = (X_1, \dots, X_N)$ má normálne (Gaussove) rozdelenie s parametrami $\boldsymbol{\mu}$ a $\boldsymbol{\Sigma}$. Toto platí, ak združená funkcia hustoty vektora $\mathbf{X}$ je

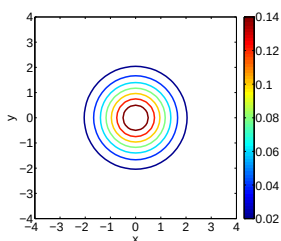
$$f(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^N |\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right), \quad (\text{A.55})$$

kde $\boldsymbol{\mu}$ je vektor stredných hodnôt a $\boldsymbol{\Sigma}$ je kovariančná matica vektora $\mathbf{X}$.

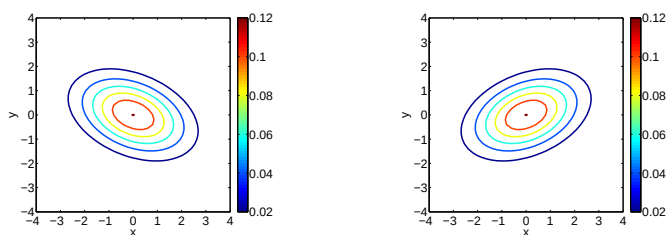
Pre dvojrozmerné normálne rozdelenie má funkcia hustoty pravdepodobnosti kopcovitý tvar (obr. A.9), ktorý závisí na kovariančnej matici, ako je znázornené na obrázku A.10.



(a) $\text{Var}(X) = 2, \text{Var}(Y) = 1, \text{Cov}(X, Y) = 0$ (b) $\text{Var}(X) = 1, \text{Var}(Y) = 2, \text{Cov}(X, Y) = 0$



(c) $\text{Var}(X) = \text{Var}(Y) = 1, \text{Cov}(X, Y) = 0$



(d) $\text{Var}(X) = 2, \text{Var}(Y) = 1, \text{Cov}(X, Y) = -0,5$ (e) $\text{Var}(X) = 2, \text{Var}(Y) = 1, \text{Cov}(X, Y) = 0,5$

Obrázok A.10: Vrstevnice dvojrozmerného normálneho rozdelenia so stredom $(0, 0)$ a rôznymi kovariančnými maticami.



Vybrané témy z maticového počtu

B.1 Singulárny rozklad matice

Singulárne číslo a singulárne vektory matice $\mathbf{A}$ sú nezáporné číslo σ a nenulové vektory $\mathbf{u}$ a $\mathbf{v}$ také, že

$$\mathbf{A}\mathbf{v} = \sigma\mathbf{u}, \quad (\text{B.1})$$

$$\mathbf{A}^H\mathbf{u} = \sigma\mathbf{v}, \quad (\text{B.2})$$

kde $\mathbf{A}^H$ je Hermitovský (komplexne združená) transponovaná matica. Pre reálne matice, s ktorými budeme pracovať, platí $\mathbf{A}^H = \mathbf{A}^T$. Z toho vyplýva, že $\mathbf{v}$ je vlastný vektor matice $\mathbf{A}^T\mathbf{A}$ a $\mathbf{u}$ vlastný vektor matice $\mathbf{A}\mathbf{A}^T$ prislúchajúce vlastnému číslu $\lambda = \sigma^2$. Dôkaz uvidíme neskôr v kapitole B.1.2.

V maticovom zápise:

$$\mathbf{A}\mathbf{V} = \mathbf{U}\mathbf{S}, \quad (\text{B.3})$$

$$\mathbf{A}^T\mathbf{U} = \mathbf{V}\mathbf{S}^T, \quad (\text{B.4})$$

kde $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_N]$, $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_D]$ a $\mathbf{S} = \text{diag}(\sigma_1, \dots, \sigma_{\text{rank}(\mathbf{A})})$.

Zápis

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T \quad (\text{B.5})$$

sa nazýva singulárny rozklad matice $\mathbf{A}$. Rozmerovo

$$[\mathbf{A}]_{N \times D} = [\mathbf{U}]_{N \times N} [\mathbf{S}]_{N \times D} [\mathbf{V}^T]_{D \times D}. \quad (\text{B.6})$$

B.1.1 Úsporná forma SVD

Ak $r = \text{rank}(\mathbf{A}) \leq \min(N, D)$ a matice $\mathbf{U}$, $\mathbf{S}$ a $\mathbf{V}$ sa dajú zapísať ako blokové nasledovne

$$\mathbf{U} = \left[\begin{array}{c|c} \mathbf{U}_1 & \mathbf{U}_2 \\ \hline \mathbf{U}_3 & \mathbf{U}_4 \end{array} \right], \quad (\text{B.7})$$

$$\mathbf{S} = \left[\begin{array}{c|c} \text{diag}(\sigma_1, \dots, \sigma_r) & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array} \right] = \left[\begin{array}{c|c} \mathbf{S}_1 & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array} \right], \quad (\text{B.8})$$

$$\mathbf{V}^T = \left[\begin{array}{c|c} \mathbf{V}_1 & \mathbf{V}_3 \\ \hline \mathbf{V}_2 & \mathbf{V}_4 \end{array} \right], \quad (\text{B.9})$$

kde $\mathbf{0}$ je nulová matica, potom

$$\mathbf{A} = \left[\begin{array}{c|c} \mathbf{U}_1 & \mathbf{U}_2 \\ \hline \mathbf{U}_3 & \mathbf{U}_4 \end{array} \right] \left[\begin{array}{c|c} \mathbf{S}_1 & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array} \right] \left[\begin{array}{c|c} \mathbf{V}_1 & \mathbf{V}_3 \\ \hline \mathbf{V}_2 & \mathbf{V}_4 \end{array} \right] = \quad (\text{B.10})$$

$$= \left[\begin{array}{c} \mathbf{U}_1 \\ \hline \mathbf{U}_3 \end{array} \right]_{(N \times r)} [\mathbf{S}_1]_{(r \times r)} [\mathbf{V}_1 | \mathbf{V}_3]_{(r \times D)}. \quad (\text{B.11})$$

Takémuto vyjadreniu sa hovorí úsporná forma SVD, pretože rátame len r stĺpcov matice $\mathbf{U}$ a r riadkov matice $\mathbf{V}^T$.

B.1.2 Vlastnosti SVD

Ak $\mathbf{A}$ je reálna štvorcová symetrická matica, potom

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{U}^T, \quad (\text{B.12})$$

kde $\mathbf{S}$ má na diagonále vlastné čísla matice $\mathbf{A}$ a $\mathbf{U}$ sú ortonormálne vlastné vektory matice $\mathbf{A}$.

Dôkaz. $\mathbf{S}$ má na diagonále vlastné čísla matice $\mathbf{A}$, $\mathbf{U}$ sú ortonormálne vlastné vektory matice $\mathbf{A}$, potom

$$\mathbf{A}\mathbf{U} = \mathbf{U}\mathbf{S}, \quad (\text{B.13})$$

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{U}^{-1}. \quad (\text{B.14})$$

Pre ortonormálne matice platí $\mathbf{U}^T = \mathbf{U}^{-1}$ a teda $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{U}^T$. $\square$

Ak $\mathbf{A}$ je reálna matica, tak

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T, \quad (\text{B.15})$$

kde $\mathbf{S}$ má na diagonále odmocninu vlastných čísel matice $\mathbf{A}\mathbf{A}^T$, $\mathbf{U}$ sú vlastné vektory matice $\mathbf{A}\mathbf{A}^T$ a $\mathbf{V}$ sú vlastné vektory matice $\mathbf{A}^T\mathbf{A}$.

Dôkaz. Pre singulárny rozklad matice platí

$$\mathbf{A}\mathbf{V} = \mathbf{U}\mathbf{S}, \quad (\text{B.16})$$

$$\mathbf{A}^T \mathbf{U} = \mathbf{V}\mathbf{S}^T, \quad (\text{B.17})$$

a teda

$$\mathbf{A}^T \mathbf{A}\mathbf{V} = \mathbf{A}^T \mathbf{U}\mathbf{S} = \mathbf{V}\mathbf{S}^T \mathbf{S}, \quad (\text{B.18})$$

$$\mathbf{A}\mathbf{A}^T \mathbf{U} = \mathbf{A}\mathbf{V}\mathbf{S}^T = \mathbf{U}\mathbf{S}\mathbf{S}^T. \quad (\text{B.19})$$

Z toho vyplýva

$$\mathbf{A}^T \mathbf{A} = \mathbf{V}\mathbf{S}^T \mathbf{S}\mathbf{V}^T, \quad (\text{B.20})$$

$$\mathbf{A}\mathbf{A}^T = \mathbf{U}\mathbf{S}\mathbf{S}^T \mathbf{U}^T, \quad (\text{B.21})$$

a teda $\mathbf{U}$ sú vlastné vektory matice $\mathbf{A}\mathbf{A}^T$ a $\mathbf{V}$ sú vlastné vektory matice $\mathbf{A}^T \mathbf{A}$.

Označme σ_i vlastné čísla matice $\mathbf{A}\mathbf{A}^T$. Pretože platí $\text{eig}(\mathbf{A}\mathbf{B}) = \text{eig}(\mathbf{B}\mathbf{A})$, σ_i sú zároveň vlastné čísla matice $\mathbf{A}^T \mathbf{A}$. Potom $\mathbf{S}^T \mathbf{S} = \mathbf{S}\mathbf{S}^T = \text{diag}(\{\sigma_i\})$.

Nech $\mathbf{S} = \text{diag}(\{\lambda_i\})$. Z toho vyplýva, že $\mathbf{S}^T \mathbf{S} = \mathbf{S}\mathbf{S}^T = \text{diag}(\{\lambda_i^2\})$, a teda $\lambda_i = \sqrt{\sigma_i}$, a teda $\mathbf{S}$ má na diagonále odmocniny vlastných čísel matice $\mathbf{A}\mathbf{A}^T$. $\square$



Vybrané kapitoly z lineárnej algebry

Pre hlbší popis uvedených pojmov odkážeme čitateľa na podrobnejšiu literatúru teórie maticového počtu ([Kra06]).

C.1 Vektorový priestor (Lineárny priestor)

Reálny vektorový priestor Reálny vektorový priestor V je neprázdna množina prvkov nazývaných vektory, v ktorej je definovaná operácia sčítanie, ktorá každej dvojici vektorov $u \in V, v \in V$ priradí vektor $u + v \in V$ a operácia násobenia vektorov reálnym číslom, ktorá každej dvojici $\alpha \in R, u \in V$, priradí vektor $\alpha u \in V$; pričom tieto operácie majú tieto vlastnosti:

1. $u + v = v + u$ pre všetky $u, v \in V$.
2. $(u + v) + w = u + (v + w)$ pre všetky $u, v, w \in V$.
3. Existuje vektor $o \in V$; zvaný nulový vektor, tak, že $v + o = v$ pre všetky $v \in V$.
4. Ku každému vektoru $v \in V$ existuje jednoznačne určený vektor $-v$ tak, že $v + (-v) = o$.
5. $(\alpha\beta)v = \alpha(\beta v)$ pre všetky $\alpha, \beta \in R$ a pre všetky $v \in V$.
6. $1 \cdot v = v$ pre všetky $v \in V$.

$$7. (\alpha + \beta)v = \alpha v + \beta v \text{ pre všetky } \alpha, \beta \in R \text{ a pre všetky } v \in V.$$

$$8. \alpha(u + v) = \alpha u + \alpha v \text{ pre všetky } u, v \in V.$$

Komplexný vektorový priestor Analogicky sa definuje komplexný vektorový priestor. Reálny a komplexný vektorový priestor budeme nazývať spoločným názvom vektorový priestor.

Číslo potom bude v prípade reálneho vektorového priestoru znamenať reálne číslo, v prípade komplexného vektorového priestoru číslo bude znamenať komplexné číslo. Čísla budeme tiež nazývať skaláry.

C.2 Lineárna kombinácia vektorov

Operácie sčítanie vektorov a násobenie vektorov číslami umožňujú zaviesť pojem lineárna kombinácia vektorov.

Ak sú $\alpha_1, \dots, \alpha_n$ ľubovoľné čísla a $v_1, \dots, v_n$ ľubovoľné vektory, tak vektor

$$\alpha_1 v_1 + \dots + \alpha_n v_n \quad (C.1)$$

nazývame

lineárnou kombináciou vektorov $v_1, \dots, v_n$ s koeficientami $\alpha_1, \dots, \alpha_n$.

C.3 Lineárna transformácia

Pojmom lineárne zobrazenie (lineárna transformácia) sa v matematike označuje také zobrazenie medzi vektorovými priestormi V_1 a V_2 , ktoré zachováva vektorové operácie sčítanie a násobenie skalárom.

C.4 Báza vektorového priestoru

Bázou vektorového priestoru V nazývame taký systém Φ vektorov vo V , pre ktoré sú splnené nasledovné podmienky:

1. Φ je lineárne nezávislý
2. Φ generuje celý priestor V (úplná báza)

Nech $a_1 \dots a_N \in V$. Potom $a_{\{1\}} \dots a_{\{N\}}$ tvoria bázu V práve vtedy, keď každý vektor $z \in V$ je možné vyjadriť ako lineárnu kombináciu vektorov $a_1 \dots a_N$ a to jednoznačne.

Ak vektory $a_1 \dots a_N$ tvoria bázu priestoru V , a pre $s \in V$ platí, že $s = s_1 a_1 + \dots + s_N a_N$. Potom **koeficienty $s_1 \dots s_N$ sa nazývajú súradnice vektora s vzhľadom k danej báze.**





Zoznam použitej literatúry

- [HT13] M. Hamala a M. Trnovska. *Nelineárne programovanie, teória a algoritmy*. Druhé vyd. EPOS, 2013.
- [Hu62] M. K. Hu. „Visual pattern recognition by moment invariants“. *Information Theory, IRE Transactions on* 8.2 (1962), str. 179–187.
- [Kor03] J. Korbaš. *Lineárna algebra a geometria I*. Vydavateľstvo UK, Bratislava, 2003.
- [Kra06] E. Krajník. *Základy maticového počtu*. ČVUT, Praha, 2006, str. 165.
- [Lit11] M. Litschmannová. *Vybrané kapitoly z pravdepodobnosti*. Skriptá. VŠB TU Ostrava, Fakulta elektrotechniky a informatiky, 2011.
- [Nav07] M. Navara. *Pravdepodobnosť a matematická statistika*. Skriptá. FEL ČVUT, Praha, 2007.
- [PJ11] A. Pázman a K. Janková. *Pravdepodobnosť a štatistika*. Vydavateľstvo UK, Bratislava, 2011.
- [Zla11] P. Zlatoš. *Lineárna algebra a geometria*. Skriptá. FMFI UK, Bratislava, 2011.

Počítačové videnie. Detekcia a rozpoznávanie objektov

Autorský kolektív:

RNDr. Elena Šikudová, PhD.

RNDr. Zuzana Černeková, PhD.

Ing. Vanda Benešová, CSc.

RNDr. Zuzana Haladová

RNDr. Júlia Kučerová

Recenzenti:

Prof. Ivan Bajla, ÚMER SAV, Bratislava

Prof. Jaroslav Polec, ÚT FEI STU, Bratislava

Doc. Jarmila Pavlovičová, ÚT FEI STU, Bratislava

Obálka: Spyridon Giochalias

Vydavateľstvo: Wikina, Livornská 445, 109 00 Praha 10

1. vydanie

Náklad: 100

Počet strán: 391

Nepredajné

Vydané s podporou grantu KEGA 068UK-4/2011

ISBN: 978-80-87925-06-5