

VYSOKÁ ŠKOLA BÁŇSKÁ – TECHNICKÁ UNIVERZITA OSTRAVA

Fakulta elektrotechniky a informatiky

**DIGITÁLNÍ ZPRACOVÁNÍ
A ANALÝZA OBRAZŮ**

Eduard Sojka

Ostrava

© Eduard Sojka, 2000
ISBN 80-7078-746-5

Obsah

Předmluva	4
1 Prostor obrazových signálů	5
1.1 Prostor signálů	5
1.1.1 Linearita prostoru signálů	5
1.1.2 Skalární součin nad prostorem signálů	6
1.1.3 Reprezentace signálu pomocí báze funkcí	8
1.2 Operace nad prostorem signálů	9
1.2.1 Lineární operace	9
1.2.2 Operace invariantní vůči posuvu	9
1.2.3 Diracův impulz	9
1.2.4 Lineární sumace bodových zdrojů, konvoluce	10
1.2.5 Diskrétní konvoluce	11
2 Transformace obrazových signálů	13
2.1 Báze δ	13
2.2 Transformace signálu	13
2.3 Fourierova transformace	14
2.3.1 Spojitá Fourierova transformace	14
2.3.2 Vlastnosti spojité Fourierovy transformace	17
2.3.3 Diskrétní Fourierova transformace	19
2.3.4 Vlastnosti diskrétní Fourierovy transformace	20
2.3.5 Rychlá Fourierova transformace	22
2.4 Kosinová transformace	24
2.5 Waveletové transformace	25
2.5.1 Spojitá waveletová transformace	25
2.5.2 Waveletová transformace s dyadickou bází	27
3 Stochastický přístup k popisu obrazových signálů	29
3.1 Náhodné pole	29
3.2 Momenty náhodného pole	29
3.3 Homogenní a ergodické náhodné pole	30
3.4 Výkonová spektrální hustota	31
3.5 Lineární operace nad náhodným polem	31
3.6 Statistická analýza obrazových transformací	32
4 Digitalizace obrazových signálů	34
4.1 Vzorkování deterministicky popsaných obrazů	34
4.2 Zpřesněné modely vzorkování	36
4.3 Aliasing	37
4.4 Praktické metody rekonstrukce obrazu	39
4.5 Kvantování obrazových signálů	41
5 Filtrace obrazových signálů	43
5.1 Nerekurzivní filtrace	43
5.1.1 Realizace nerekurzivního filtru v prostorové a frekvenční doméně	43
5.1.2 Filtry s nulovou fází	44
5.1.3 Návrh filtru s využitím výřezové funkce	44
5.1.4 Návrh optimálního filtru	45

5.2	Rekurzivní filtrace	46
5.2.1	Realizace rekurzivního filtru přímým výpočtem	46
5.2.2	Realizace rekurzivního filtru iteračním výpočtem	47
5.2.3	Návrh rekurzivního filtru	48
5.3	Obecný model degradace a rekonstrukce obrazu	48
5.4	Inverzní filtrace	49
5.5	Wienerův filtr ve frekvenční doméně	49
5.6	Wienerův filtr v prostorové doméně	50
6	Bodové, algebraické a geometrické operace s obrazy	52
6.1	Bodové operace	52
6.1.1	Vliv bodových operací na histogram jasu	53
6.1.2	Vyrovnění histogramu	54
6.2	Algebraické operace s obrazy	54
6.2.1	Vliv algebraických operací na histogram jasu	55
6.2.2	Využití průměrování k redukci šumu	55
6.3	Geometrické transformace obrazu	56
7	Komprese obrazu	60
7.1	Redukce objemu barvosné informace	60
7.2	Komprese JPEG	60
7.3	Komprese MPEG	64
7.4	MPEG komprese zvuku	68
7.5	Fraktální komprese	71
8	Segmentace obrazu	74
8.1	Detekce hran	74
8.1.1	Gradientní metody hledání hran	74
8.1.2	Detekce hran hledáním průchodu druhé derivace nulou	76
8.1.3	Parametrické modely hrany	78
8.1.4	Cannyho detektor hran	79
8.1.5	Stanovení hran na základě textury	82
8.1.6	Hledání hran v barevných obrazech	83
8.2	Spojování hran	84
8.2.1	Heuristické sledování hrany	84
8.2.2	Proložení přímky a křivky	85
8.2.3	Houghova transformace	86
8.3	Detekce oblastí	87
8.3.1	Prahování	87
8.3.2	Metoda spojování oblastí	89
8.3.3	Metoda dělení oblastí	89
8.4	Detekce rohů	90
8.4.1	Kitchen-Rosenfeldův detektor rohů	90
8.4.2	Harrisův detektor rohů	90
8.5	Zpracování binárních obrazů	91
8.5.1	Matematická morfologie	91
8.5.2	Ztenčování	92
8.5.3	Kostra	92
8.5.4	Indexování	92
9	Měření objektů pro příznakové rozpoznání	93
9.1	Momenty	93
9.2	Pravoúhlost a podlouhlost	94
9.3	Kruhovitost	94

9.4	Energie hranice	95
9.5	Průměrná vzdálenost pixelu od hranice	95
9.6	Popis tvaru objektu pomocí průběhu křivosti jeho hranice	95
9.7	Eulerovo číslo	96
9.8	Atributy odvozené z histogramu jasu	96
9.9	Atributy odvozené z frekvenčního spektra jasu	96
9.10	Algoritmus výpočtu plochy, obvodu a Eulerova čísla	97
9.11	Algoritmus určení hranice	98
9.12	Hodnocení vhodnosti a redukce zvolené množiny příznaků	98
9.13	Karhunen-Loéveho transformace	99
10	Příznakové metody analýzy obrazu	101
10.1	Klasifikátory	101
10.1.1	Klasifikace pomocí diskriminančních funkcí	102
10.1.2	Klasifikace pomocí minimální vzdálenosti	103
10.1.3	Určení diskriminačních funkcí minimalizací rizika	104
10.2	Neuronové sítě	106
10.2.1	Třívrstvá (vícevrstvá) síť s učením „back propagation“	107
10.2.2	Neuronová síť s kompeticí	108
11	Zpětná stereoprojekce	110
11.1	Model kamery	110
11.2	Případ dvou kamer s rovnoběžnými optickými osami	112
11.3	Absolutní kalibrace a rekonstrukce	112
11.4	Relativní kalibrace a rekonstrukce	114
11.5	Epipolára	117
11.6	Automatizované hledání korespondence	117
12	Analýza obrazů proměnných v čase	119
12.1	Diskrétní Kalmanův filtr	119
12.2	Aplikace Kalmanova filtru při sledování pohybujících se objektů	120
12.3	Alfa, beta filtr pro sledování pohybu bodu	123
12.4	Sledování objektů v obrazech získaných pohybující se kamerou	124
12.5	Optický tok	125
Dodatek A:	Základní vztahy počtu pravděpodobnosti	127
A.1	Náhodná proměnná	127
A.2	Podmíněná pravděpodobnost	128
A.3	Sdružená pravděpodobnost	128
Dodatek B:	Řešení předeterminovaných systémů lineárních rovnic	129
Dodatek C:	Hledání mimima s podmínkou	130
Dodatek D:	Hledání mimima gradientní metodou	131
	Literatura odkazovaná z textu	132
	Další doporučená literatura	132

Předmluva

Digitální zpracování a analýza obrazů jsou moderní a rychle se rozvíjející disciplíny. Jejich výsledky jsou dnes již běžně používány při úpravách a kompresi obrazů a videosekvencí, při rozpoznávání textů, v kontrolních a střežících systémech, v průmyslových robotech, ve fotogrametrii i jinde. Dostatečný výkon a technologická úroveň současných výpočetních systémů umožnila pronikání digitálního zpracování obrazu dokonce i do spotřební elektroniky. Příkladem je digitální fotografie, digitální video a digitální televize.

Na fakultě elektrotechniky a informatiky VŠB-TU Ostrava je ve studijním oboru „informatika“ problematika digitálního zpracování a analýzy obrazu přednášena v předmětech „Digitální zpracování obrazu“ a „Analýza obrazu a počítačové vidění“. Předkládaný učební text je určen pro potřeby obou těchto předmětů, vychází z několikaleté zkušenosti s výukou těchto předmětů a pokrývá značnou část látky, která je v těchto předmětech přednášena. Poněkud teoretická povaha textu je kompenzována ve cvičeních, která jsou naopak věnována implementačním a jiným praktickým pohledům na probíranou tematiku, jakož i ukázkám v praxi používaných programových systémů. Jsem si vědom toho, že i při velké péči, kterou jsem vypracování textu věnoval, se v něm mohou vyskytnout nejasnosti, překlepy nebo i chyby. Čtenářům budu proto vděčný za upozornění na veškerá případná nedopatření i za připomínky a komentáře.

Ostrava, září 1999

Autor

1 Prostor obrazových signálů

V této kapitole podáváme přehled základních teoretických nástrojů, které jsou potřebné pro kvalifikované zvládnutí problematiky digitálního zpracování obrazu. Protože každý optický obraz je signálem, budeme v této kapitole poněkud obecnější a místo o obrazu nebo obrazovém signálu budeme často obecně hovořit pouze o signálu. Závěry, které zde uvedeme, platí pro jakékoli signály, a tedy i pro signály obrazové. Obrazové signály jsou zpravidla dvojrozměrné. S ohledem na tuto skutečnost budeme vztahy, které budou později prakticky využívány, ihned formulovat také v jejich dvojrozměrných variantách.

1.1 Prostor signálů

Předpokládáme, že signál je definován nad jistou podmnožinou (označme ji Ω) m -rozměrného euklidovského prostoru E^m . Touto podmnožinou může být nějaká ohraničená nebo neohraničená souvislá oblast v E^m , množina jednotlivých bodů z E^m , ale také celý prostor E^m . Jako matematickou reprezentaci signálu zvolíme funkci $f: \Omega \rightarrow \gamma$, kde γ je obor hodnot signálu. Množina γ může být např. množinou celých nebo reálných čísel (obrazy ve stupních šedi), množinou komplexních čísel (obrazy vznikající některými transformacemi), množinou vektorů o třech složkách (v barevných obrazech bývá často barva každého bodu popsána trojicí R, G, B , jejíž jednotlivé členy udávají intenzitu červené, zelené a modré barevné složky) nebo podmnožinou některé z uvedených množin. O signálu f řekneme, že je m -rozměrným signálem. Množinu funkcí realizujících zobrazení $\Omega \rightarrow \gamma$ nazveme m -rozměrným signálovým prostorem \mathcal{S} . Je tedy $\mathcal{S} = \{f | f: \Omega \rightarrow \gamma\}$. Úvahy provedené v této kapitole se budou týkat zejména signálů reálných a komplexních. Doplnující informace, které se týkají signálů reprezentovaných celými čísly a vektory, uvedeme na vhodném místě později.

1.1.1 Linearita prostoru signálů

Při zpracování signálů dosti často předpokládáme, že pro signály platí princip superpozice. Matematickým vyjádřením této představy je, že signálový prostor předpokládáme lineární. Zopakujme, že v lineárním prostoru je definována operace sčítání prvků a násobení prvku konstantou (konstantou může být reálné nebo komplexní číslo). Připomeňme dále, že v lineárním prostoru existuje nulový prvek a ke každému prvku existuje prvek opačný.

Uvažujme nejprve prostor jednorozměrných diskretních signálů (označme jej \mathcal{S}), kde definičním oborem Ω je množina M bodů. Každý bod je popsán svým indexem m . Bez újmy na obecnosti lze předpokládat, že index nabývá hodnot $0, 1, \dots, M-1$. V prostoru \mathcal{S} uvažujme množinu $\{\varphi_k(m) | k=0, 1, \dots, N-1\}$ obsahující N diskretních funkcí definovaných nad Ω . Funkce této množiny nazveme lineárně závislými, jestliže lze najít koeficienty a_k (z nichž alespoň jeden je různý od nuly) tak, že platí vztah (1.1). V opačném případě funkce nazveme lineárně nezávislými.

$$\sum_{k=0}^{N-1} a_k \varphi_k(m) = 0. \quad (1.1)$$

Množina $\{\varphi_k(m)\}$ je úplná v daném prostoru signálů, jestliže je možné každou funkci z tohoto prostoru vyjádřit jako lineární kombinaci funkcí z $\{\varphi_k(m)\}$. Lineárně nezávislý systém funkcí $\{\varphi_k(m)\}$, který je úplný, nazveme bází prostoru \mathcal{S} . Každou funkci $f(m)$ prostoru \mathcal{S} lze vyjádřit jako lineární kombinaci bázevých funkcí

$$f(m) = \sum_{k=0}^{M-1} F_k \varphi_k(m) = \sum_{k=0}^{M-1} F(k) \varphi_k(m), \quad (1.2)$$

kde F_k jsou koeficienty, které obecně mohou být komplexní. Protože na \mathcal{S} můžeme pohlížet jako na prostor M -rozměrných vektorů, je zřejmé, že bázevých funkcí je právě M . Pro zvolený systém bázevých funkcí lze proto každý signál $f(m)$ prostoru \mathcal{S} reprezentovat uspořádanou M -ticí $(F_0, F_1, \dots, F_{M-1})$

koeficientů. Na tuto M -tici můžeme ovšem také pohlížet jako na diskretní funkci $F(k)$, $k = 0, 1, \dots, M-1$, čehož jsme již ve vztahu (1.2) využili.

Analogicky lze postupovat také pro jednorozměrné spojité signály. V tomto případě je definičním oborem signálů nespočetná množina $\Omega \subseteq \mathbb{E}^1$, a proto lze očekávat, že i báze obecně bude nespočetnou množinou funkcí. Takovou množinu můžeme formálně zapsat jako jedinou funkci $\varphi(u, x)$ dvou argumentů, kde $x \in \Omega$ a $u \in U$, $U \subseteq \mathbb{R}$ (\mathbb{R} značí množinu reálných čísel). Tuto funkci budeme nazývat bázevým jádrem. Libovolný spojitý signál $f(x)$ lze pak vyjádřit jako lineární kombinaci pomocí vztahu

$$f(x) = \int_U F(u) \varphi(u, x) du. \quad (1.3)$$

S ohledem na zvolené bázevé jádro $\varphi(u, x)$ lze tedy každý prvek $f(x)$ prostoru spojitých signálů reprezentovat pomocí funkce $F(u)$. Uvedené závěry snadno zobecníme také na signály dvojrozměrné. Předpokládáme, že v případě dvojrozměrného diskretního signálu je $\Omega = \{(m, n) \mid m=0, 1, \dots, M-1; n=0, 1, \dots, N-1\}$. Celkový počet bázevých funkcí je proto $M \times N$. V případě dvojrozměrného spojitého signálu je podle očekávání poloha každého bodu popsána dvojicí souřadnic x, y a dále je $U \subseteq \mathbb{R} \times \mathbb{R}$. Pro dvojrozměrný diskretní a dvojrozměrný spojitý signál tak postupně máme

$$f(m, n) = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} F(k, l) \varphi_{k, l}(m, n), \quad (1.4)$$

$$f(x, y) = \iint_U F(u, v) \varphi(u, v, x, y) du dv. \quad (1.5)$$

Jak je zřejmé ze vztahu (1.4), nahlížíme ve shodě s poznámkou ke vztahu (1.2) v diskretním případě na koeficienty v lineární kombinaci jako na dvojrozměrnou funkci $F(k, l)$, $k \in \{0, 1, \dots, M-1\}$, $l \in \{0, 1, \dots, N-1\}$.

1.1.2 Skalární součin nad prostorem signálů

Řešení úloh se usnadní, jestliže nad prostorem signálů zavedeme skalární součin. Necht' \mathcal{S} je lineární prostor signálů. Skalární součin $\langle \varphi, \psi \rangle$ přiřazuje každým dvěma prvkům φ, ψ tohoto prostoru číslo (obecně se jedná o číslo komplexní) těchto vlastností:

$$\langle \varphi, \psi \rangle = \langle \psi, \varphi \rangle^*, \quad (1.6)$$

$$\langle a\varphi_1 + b\varphi_2, \psi \rangle = a\langle \varphi_1, \psi \rangle + b\langle \varphi_2, \psi \rangle, \quad (1.7)$$

$$\langle \varphi, \varphi \rangle \geq 0, \quad \langle \varphi, \varphi \rangle = 0 \Leftrightarrow \varphi = 0, \quad (1.8)$$

kde a, b jsou čísla (obecně opět komplexní) a $*$ označuje komplexně sdružené číslo. Prostor, v němž je zaveden skalární součin, se nazývá unitárním prostorem. Pomocí skalárního součinu lze zavést normu $\|\varphi\|$ prvku φ jako odmocninu ze skalárního součinu

$$\|\varphi\| = \sqrt{\langle \varphi, \varphi \rangle}. \quad (1.9)$$

Prvky φ, ψ unitárního prostoru nazveme ortogonálními, jestliže platí

$$\langle \varphi, \psi \rangle = 0. \quad (1.10)$$

Necht' $\{\varphi_k\}$ je systém prvků unitárního prostoru. Tento systém nazveme ortogonálním, jestliže platí

$$\langle \varphi_k, \varphi_l \rangle = \begin{cases} \neq 0, & k = l \\ = 0, & k \neq l \end{cases}. \quad (1.11)$$

Systém $\{\varphi_k\}$ nazveme ortonormálním, jestliže navíc platí

$$\langle \varphi_k, \varphi_l \rangle = \delta(k-l), \quad (1.12)$$

kde $\delta(i)$ je tzv. Kroneckerova funkce, která je definována nad množinou celých čísel vztahem

$$\delta(i) = \begin{cases} 1, & i = 0 \\ 0, & i = \pm 1, \pm 2, \dots \end{cases}. \quad (1.13)$$

Uvažujme nyní unitární prostor \mathcal{S} , ve kterém je systém $\{\varphi_k\}$ bázi. Jestliže je systém $\{\varphi_k\}$ ortogonální popř. ortonormální, pak hovoříme o ortogonální popř. ortonormální bázi. K prověření úplnosti systému funkcí $\{\varphi_k\}$ v \mathcal{S} lze použít větu: Systém funkcí $\{\varphi_k\}$ je úplný v \mathcal{S} právě tehdy, jestliže v \mathcal{S} neexistuje prvek (s výjimkou nulového), který by byl ortogonální ke všem prvkům systému $\{\varphi_k\}$ (větu uvádíme bez důkazu).

Zvolme nyní v prostoru \mathcal{S} diskretních signálů konkrétní skalární součin a ukažme důsledky této volby. Definičním oborem funkcí prostoru \mathcal{S} nechť je nejprve množina $\Omega = \{0, 1, \dots, M-1\}$ a $\varphi(m)$, $\psi(m)$ nechť jsou funkce z \mathcal{S} . Skalární součin v prostoru \mathcal{S} zavedme takto

$$\langle \varphi(m), \psi(m) \rangle = \sum_{m=0}^{M-1} \varphi(m) \psi^*(m), \quad (1.14)$$

kde $*$ označuje komplexně sdruženou funkci. Na čtenáři ponecháváme, aby ukázal, že takto definovaný skalární součin má vlastnosti (1.6) až (1.8). Nechť $\{\varphi_k(m)\}$ je báze v prostoru \mathcal{S} . Podmínku (1.12) ortonormality lze nyní rozepsat ve tvaru

$$\sum_{m=0}^{M-1} \varphi_k(m) \varphi_l^*(m) = \delta(k-l). \quad (1.15)$$

Lze ukázat, že uvedenému vztahu je ekvivalentní vztah

$$\sum_{k=0}^{M-1} \varphi_k(m) \varphi_k^*(n) = \delta(m-n). \quad (1.16)$$

Analogicky lze postupovat také ve spojitém případě. Nechť \mathcal{S} je prostor spojitých signálů. V \mathcal{S} uvažujme bázevé jádro $\varphi(u, x)$, $x \in \Omega$, $u \in U$. Podmínka ortonormality báze má nyní tvar

$$\langle \varphi(u, x), \varphi(v, x) \rangle = \delta(u-v), \quad (1.17)$$

kde $\delta(u)$ je v tomto případě tzv. Diracův impuls. (I v dalším textu budeme podle potřeby symbolu δ používat k označení buď Kroneckerovy funkce nebo Diracova impulsu. O kterou funkci jde, bude vždy zřejmé z kontextu). Diracův impuls nabývá nenulové hodnoty pouze pro $u=0$, kde je jeho funkční hodnota nekonečno. Podrobněji o této funkci pojednáme v odstavci 1.2.3. Nechť $\varphi(x)$, $\psi(x)$ jsou funkce z \mathcal{S} . V prostoru \mathcal{S} zavedeme skalární součin předpisem

$$\langle \varphi(x), \psi(x) \rangle = \int_{\Omega} \varphi(x) \psi^*(x) dx. \quad (1.18)$$

Ověření, že takto zavedený skalární součin má vlastnosti (1.6) až (1.8), opět ponecháváme čtenáři jako cvičení. Podmínku ortonormality báze můžeme nyní zapsat jako

$$\int_{\Omega} \varphi(u, x) \varphi^*(v, x) dx = \delta(u-v). \quad (1.19)$$

Opět lze ukázat, že uvedenému vztahu je ekvivalentní vztah

$$\int_U \varphi(u, x) \varphi^*(u, y) du = \delta(x-y). \quad (1.20)$$

Zobecnění dosud uvedeného na dvojrozměrné signály je přímočaré. V případě dvojrozměrných signálů máme $\Omega \subseteq E^2$. Vztahům (1.14) až (1.20) postupně odpovídají následující vztahy (vztahy (1.21), (1.24) podávají návod na výpočet skalárního součinu, vztahy (1.22), (1.23), (1.25), (1.26) vyjadřují podmínku ortonormality):

$$\langle \varphi(m, n), \psi(m, n) \rangle = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \varphi(m, n) \psi^*(m, n), \quad (1.21)$$

$$\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \varphi_{k_1, l_1}(m, n) \varphi_{k_2, l_2}^*(m, n) = \delta(k_1 - k_2) \delta(l_1 - l_2) = \delta(k_1 - k_2, l_1 - l_2), \quad (1.22)$$

$$\sum_{k=0}^{M-1} \sum_{l=0}^{N-1} \varphi_{k,l}(m_1, n_1) \varphi_{k,l}^*(m_2, n_2) = \delta(m_1 - m_2, n_1 - n_2), \quad (1.23)$$

$$\langle \varphi(x, y), \psi(x, y) \rangle = \iint_{\Omega} \varphi(x, y) \psi^*(x, y) dx dy, \quad (1.24)$$

$$\iint_{\Omega} \varphi(u_1, v_1, x, y) \varphi^*(u_2, v_2, x, y) dx dy = \delta(u_1 - u_2, v_1 - v_2), \quad (1.25)$$

$$\iint_U \varphi(u, v, x_1, y_1) \varphi^*(u, v, x_2, y_2) du dv = \delta(x_1 - x_2, y_1 - y_2). \quad (1.26)$$

1.1.3 Reprezentace signálu pomocí báзовých funkcí

Nechť \mathcal{S} je unitární prostor diskrétních signálů definovaných nad množinou M bodů, necht' $f(m)$ je signál z tohoto prostoru a $\{\varphi_k(m) \mid k = 0, 1, \dots, M-1\}$ necht' je ortonormální báze v \mathcal{S} . Ze vztahu (1.2) víme, že $f(m)$ je možné vyjádřit jako lineární kombinaci

$$f(m) = \sum_{l=0}^{M-1} F_l \varphi_l(m). \quad (1.27)$$

Provedeme-li skalární součin obou stran rovnice (1.27) s funkcí $\varphi_k(m)$ báze, pak získáme

$$\langle f(m), \varphi_k(m) \rangle = \sum_{l=0}^{M-1} F_l \langle \varphi_l(m), \varphi_k(m) \rangle. \quad (1.28)$$

S ohledem na ortonormalitu máme $F_k = \langle f(m), \varphi_k(m) \rangle$. (1.29)

Rozepíšeme-li skalární součin podle vztahu (1.14) a zapíšeme-li M -tici F_0, F_1, \dots, F_{M-1} formálně jako funkci, pak lze vztah (1.29) přepsat do tvaru

$$F(k) = \sum_{m=0}^{M-1} f(m) \varphi_k^*(m). \quad (1.30)$$

Vztah (1.30) podává návod, jak pro zadaný signál $f(m)$ a zadanou bázi $\{\varphi_k(m)\}$ prostoru \mathcal{S} určit M -tici koeficientů $(F_0, F_1, \dots, F_{M-1})$, které signál reprezentují.

Analogicky je opět možné postupovat také pro spojité signály. Necht' \mathcal{S} je unitární prostor jednorozměrných spojitých signálů, $\varphi(u, x)$ necht' je ortonormální bázevé jádro a $f(x)$ necht' je signál v \mathcal{S} . Ze vztahu (1.3) víme, že f je možné vyjádřit jako lineární kombinaci

$$f(x) = \int_U F(v) \varphi(v, x) dv. \quad (1.31)$$

Provedeme-li skalární součin obou stran rovnice (1.31) s $\varphi(u, x)$, pak po úpravě získáme

$$F(u) = \langle f(x), \varphi(u, x) \rangle. \quad (1.32)$$

Rozepíšeme-li skalární součin podle (1.18), pak lze vztah (1.32) přepsat na

$$F(u) = \int_{\Omega} f(x) \varphi^*(u, x) dx. \quad (1.33)$$

Vztah (1.33) opět podává návod, jak pro zadaný signál $f(x)$ a pro zadané ortonormální bázevé jádro $\varphi(u, x)$ určit funkci $F(u)$, která signál reprezentuje. Jak je již obvyklé, získáme podobně snadno také předpisy pro funkce $F(k, l)$, $F(u, v)$ použité ve vztazích (1.4), (1.5):

$$F(k, l) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n) \varphi_{k,l}^*(m, n), \quad (1.34)$$

$$F(u, v) = \iint_{\Omega} f(x, y) \varphi^*(u, v, x, y) dx dy. \quad (1.35)$$

1.2 Operace nad prostorem signálů

Nechť \mathcal{S} je prostor signálů. Matematickým modelem úpravy signálu je unární nebo *n*ární operátor \mathcal{O} realizující zobrazení $\mathcal{S} \rightarrow \mathcal{S}$, popřípadě $\mathcal{S} \times \mathcal{S} \times \dots \times \mathcal{S} \rightarrow \mathcal{S}$. Jestliže, s ohledem na naši aplikaci, uvažujeme prostor dvojrozměrných signálů a jestliže $f(x,y), f_1(x,y), \dots, f_n(x,y) \in \mathcal{S}$, jsou prvky tohoto prostoru, pak můžeme úpravu signálu popsat vztahy

$$g(x,y) = \mathcal{O}\{f(x,y)\}, \quad (1.36)$$

$$g(x,y) = \mathcal{O}\{f_1(x,y), f_2(x,y), \dots, f_n(x,y)\}, \quad (1.37)$$

kde $g(x,y)$ je opět prvkem uvažovaného signálového prostoru. Velmi často předpokládáme, že operátor \mathcal{O} má nějaké speciální vlastnosti. Nejčastěji se jedná o linearitu a invarianci vůči posuvu.

1.2.1 Lineární operace

Většina výsledků v teorii zpracování signálů je odvozena za předpokladu, že úprava signálu je popsána operátorem, který je lineární. Vysvětlíme tento pojem: Nechť $f(x,y), g(x,y) \in \mathcal{S}$ jsou signály; a, b nechť jsou komplexní nebo reálná čísla. Operátor \mathcal{O} nazveme lineárním operátorem, jestliže platí

$$\mathcal{O}\{af(x,y)+bg(x,y)\} = a\mathcal{O}\{f(x,y)\} + b\mathcal{O}\{g(x,y)\}. \quad (1.38)$$

Uvedený vztah můžeme samozřejmě zobecnit pro libovolný počet sčítanců.

1.2.2 Operace invariantní vůči posuvu

Někdy požadujeme, aby měl operátor kromě linearitu ještě některé další vlastnosti. Velmi často to bývá invariance operátoru vůči posuvu. Nechť $f(x,y)$ je signál a pro signál $g(x,y)$ nechť platí $g(x,y) = \mathcal{O}\{f(x,y)\}$. Operátor \mathcal{O} nazveme invariantním vůči posuvu, jestliže pro libovolný signál $f(x,y)$ a pro každou dvojici (a,b) vyplývá ze vztahu $g(x,y) = \mathcal{O}\{f(x,y)\}$, že platí také

$$g(x-a, y-b) = \mathcal{O}\{f(x-a, y-b)\}. \quad (1.39)$$

Méně formálně: Jestliže operátor invariantní vůči posuvu aplikujeme na dva vzájemně posunuté, ale jinak shodné signály, pak výsledkem operace jsou signály, které jsou také vzájemně posunuté, ale jinak shodné. Vzájemný posuv výsledných signálů je při tom stejný jako posuv signálů vstupních.

1.2.3 Diracův impulz

Mnoho pozoruhodných výsledků bylo ve zpracování signálů odvozeno za použití impulzové funkce - tzv. Diracova impulzu. Věnujme proto tento odstavec uvedené funkci a jejím vlastnostem. Pro další úvahy je užitečné nejprve zavést funkci $\text{rect}(x,y)$, která je definována předpisem

$$\text{rect}(x,y) = \begin{cases} 1, & \left(|x| \leq \frac{1}{2}\right) \cap \left(|y| \leq \frac{1}{2}\right) \\ 0, & \text{jinak} \end{cases}. \quad (1.40)$$

Dále zavedme funkci

$$\delta_n(x,y) = n^2 \text{rect}(nx, ny), \quad n = 1, 2, 3, \dots, \quad (1.41)$$

tj.
$$\delta_n(x,y) = \begin{cases} n^2, & \left(|x| \leq \frac{1}{2n}\right) \cap \left(|y| \leq \frac{1}{2n}\right) \\ 0, & \text{jinak} \end{cases}.$$

Je zřejmé, že platí
$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta_n(x,y) dx dy = 1. \quad (1.42)$$

Uvažujme případ, kdy $n \rightarrow \infty$. Funkce $\delta_{\infty}(x,y)$ se nazývá Diracův impulz. Pro Diracův impulz budeme v dalším textu používat jednoduššího zápisu $\delta(x,y)$. Je zřejmé, že pro všechna $(x,y) \neq (0,0)$ je hodnota Diracova impulzu $\delta(x,y) = 0$, v bodě $(0,0)$ je hodnota nekonečno. Dále je zřejmé, že platí

$$\delta(-x, -y) = \delta(x,y), \quad (1.43)$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(x, y) dx dy = 1. \quad (1.44)$$

Pro odvození jedné z nejvýznačnějších vlastností Diracova impulsu mějme funkci $f(x, y)$ a uvažujme integrál

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta_n(x, y) dx dy = n^2 \int_{-1/2n}^{1/2n} \int_{-1/2n}^{1/2n} f(x, y) dx dy. \quad (1.45)$$

Z výrazu na pravé straně rovnice (1.45) vidíme, že integrál na levé straně rovnice vyjadřuje střední hodnotu funkce $f(x, y)$ v oblasti nad čtvercem o stranách $1/n$ se středem v počátku souřadné soustavy. Pro $n \rightarrow \infty$ pak proto obdržíme střední hodnotu nad nekonečně malým čtvercem se středem v počátku, což je funkční hodnota $f(0, 0)$. Je tedy

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x, y) dx dy = f(0, 0). \quad (1.46)$$

Posuneme-li impuls do bodu (a, b) , pak analogicky máme

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x - a, y - b) dx dy = f(a, b). \quad (1.47)$$

Poznamenejme, že vlastnost popsaná vztahem (1.47) má zásadní význam a často se jí používá. Konečně na závěr tohoto odstavce uvedeme ještě jednu vlastnost, kterou později upotřebíme:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp[-j2\pi(ux + vy)] du dv = \delta(x, y), \quad \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp[-j2\pi(ux + vy)] dx dy = \delta(u, v), \quad (1.48)$$

kde j je komplexní jednotka. Doporučujeme čtenáři, aby se alespoň neformálně pokusil uvedenou vlastnost zdůvodnit. Návod: Uvažte, že pro $x=y=0$ se v prvním případě jedná o integraci funkce nabývající hodnoty 1 nad nekonečnou oblastí. Pro $x \neq 0, y \neq 0$ se podle Eulerova vztahu $e^{j\varphi} = \cos\varphi + j\sin\varphi$ jedná o integraci funkcí typu $\cos\varphi\cos\psi, \cos\varphi\sin\psi, \sin\varphi\cos\psi, \sin\varphi\sin\psi$ ($\varphi=2\pi ux, \psi=2\pi vy$) nad nekonečnou oblastí. Obdobně lze postupovat i v případě druhého vztahu.

1.2.4 Lineární sumace bodových zdrojů, konvoluce

Uvažujme funkci $f(x, y)$. Podle (1.47) můžeme $f(x, y)$ s využitím Diracova impulsu zapsat ve tvaru

$$f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(a, b) \delta(a - x, b - y) da db. \quad (1.49)$$

Protože se pro Diracův impuls používá také název bodový zdroj, říkáme, že jsme funkci $f(x, y)$ zapsali ve formě lineární sumace bodových zdrojů. Aby lépe vynikly vlastnosti uvedeného zápisu, zapíšeme integrál ze vztahu (1.49) na okamžik ve tvaru sumace. Dostaneme

$$f(x, y) = \sum_i \sum_j f(a_i, b_j) \delta(a_i - x, b_j - y) \Delta a_i \Delta b_j. \quad (1.50)$$

Je užitečné poznamenat, že hodnota $f(a_i, b_j)$ ve vztahu (1.50) je na každém jednom intervalu konstantní. Předpokládejme nyní, že na funkci $f(x, y)$ chceme aplikovat operátor \mathcal{O} . Budeme předpokládat, že operátor \mathcal{O} je lineární a invariantní vůči posuvu. Aplikací operátoru získáme výsledek $g(x, y)$. Máme

$$g(x, y) = \mathcal{O}\{f(x, y)\}. \quad (1.51)$$

Zapíšeme-li funkci $f(x, y)$ podle vztahu (1.49) jako lineární sumu bodových zdrojů, dostaneme

$$g(x, y) = \mathcal{O}\{f(x, y)\} = \mathcal{O}\left\{ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(a, b) \delta(a - x, b - y) da db \right\}. \quad (1.52)$$

Protože je podle předpokladu operátor \mathcal{O} lineární, dostáváme dále (viz též vztah (1.50))

$$g(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(a, b) \mathcal{O}\{\delta(a-x, b-y)\} da db \quad (1.53)$$

a podle (1.43) také

$$g(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(a, b) \mathcal{O}\{\delta(x-a, y-b)\} da db. \quad (1.54)$$

Předpokládejme dále, že podrobíme-li Diracův impulz $\delta(x, y)$ operaci \mathcal{O} , získáme funkci $h(x, y) = \mathcal{O}\{\delta(x, y)\}$ (funkce h se nazývá impulzová charakteristika operátoru \mathcal{O}). Protože je podle předpokladu operátor \mathcal{O} invariantní vůči posuvu, je $\mathcal{O}\{\delta(x-a, y-b)\} = h(x-a, y-b)$. S uvážením této skutečnosti je tedy

$$g(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(a, b) h(x-a, y-b) da db. \quad (1.55)$$

Poznamenejme, že úpravou vztahu (1.53) na vztah (1.54) jsme převedli úlohu do tvaru, v němž je možné využít důsledků invariance operátoru \mathcal{O} vůči posuvu (předpoklad invariance totiž zajišťuje platnost vztahu $\mathcal{O}\{f(x-a, y-b)\} = g(x-a, y-b)$, obecně nikoli platnost vztahu $\mathcal{O}\{f(a-x, b-y)\} = g(a-x, b-y)$). Rovnice (1.55) má zásadní význam. Říká toto: Máme-li lineární operátor \mathcal{O} invariantní vůči posuvu, pak odezvu $\mathcal{O}\{f(x, y)\}$ na funkci $f(x, y)$ lze spočítat integrací na základě znalosti odezvy $h(x, y) = \mathcal{O}\{\delta(x, y)\}$ na Diracův impulz. Výraz na pravé straně rovnice (1.55), se nazývá konvolucí funkcí f a h (stručně zapisujeme $f * h$). Je tedy

$$g(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(a, b) h(x-a, y-b) da db = f(x, y) * h(x, y). \quad (1.56)$$

Po záměně proměnných $a' = x-a$, $b' = y-b$ a jejich přejmenování $a' \rightarrow a$, $b' \rightarrow b$ máme také

$$g(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(a, b) f(x-a, y-b) da db = h(x, y) * f(x, y). \quad (1.57)$$

Je proto

$$f(x, y) * h(x, y) = h(x, y) * f(x, y). \quad (1.58)$$

Ze vztahu (1.56) vidíme, že konvoluce počítá výslednou hodnotu v bodě (x, y) jako „vážený průměr“ funkce f s váhovou funkcí h , která se často nazývá konvoluční jádro nebo filtr (přesněji řečeno se jedná o impulzovou charakteristiku filtru). Výpočet při tom probíhá tak, že se funkce h v rovině x, y otočí o 180° (důsledek členů se zápornými znaménky v (1.56)) a pak se posune o vektor (x, y) . Průběh výpočtu je ilustrován na obr. 1.1. Pro úplnost ještě poznamenejme, že vztahy pro jednorozměrnou konvoluci lze získat ze vztahů (1.56), (1.57) tak, že členy odpovídající jedné z proměnných jednoduše vypustíme. Máme tedy

$$f(x) * h(x) = h(x) * f(x) = \int_{-\infty}^{\infty} f(a) h(x-a) da = \int_{-\infty}^{\infty} f(x-a) h(a) da. \quad (1.59)$$

1.2.5 Diskrétní konvoluce

Uvažujme nyní prostor diskrétních obrazových signálů. Necht' je $\Omega = \{(m, n) \mid m = 0, 1, \dots, M-1; n = 0, 1, \dots, N-1\}$ a $f(m, n)$ a $h(m, n)$ necht' jsou signály z uvažovaného prostoru. Diskrétní konvoluci lze pak zapsat pomocí vztahů

$$f(m, n) * h(m, n) = \frac{1}{MN} \sum_{r=0}^{M-1} \sum_{s=0}^{N-1} f(r, s) h(m-r, n-s), \quad (1.60)$$

$$h(m, n) * f(m, n) = \frac{1}{MN} \sum_{r=0}^{M-1} \sum_{s=0}^{N-1} h(r, s) f(m-r, n-s). \quad (1.61)$$

Poznamenejme, že v literatuře lze pro konvoluci nalézt též vztahy bez členu $1/(MN)$. Tento rozdíl je vcelku nepodstatný, protože se jedná o pouhé násobení konstantou. Při odvozování navazujících výsledků je pouze zapotřebí mít na paměti, kterou definici používáme. Povšimněme si, že při výpočtu

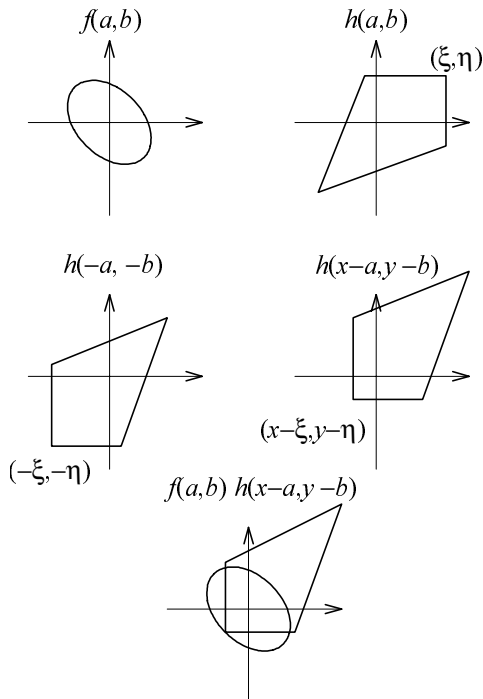
konvoluce budou pro funkce h popř. f ve vztazích (1.60), (1.61) vyžadovány také hodnoty vně rozsahu $0, 1, \dots, M-1$; $0, 1, \dots, N-1$. Tento problém lze vyřešit dvěma způsoby. V obou doplníme definici funkcí f, h i mimo oblast Ω . Při tzv. lineární konvoluci předpokládáme, že funkce f, h mimo oblast Ω nabývají hodnoty 0. Při tzv. cyklické konvoluci funkce f, h naopak považujeme za periodické s periodou M, N . Výpočet cyklické konvoluce lze pak provést na základě vztahů

$$f(m, n) * h(m, n) = \frac{1}{MN} \sum_{r=0}^{M-1} \sum_{s=0}^{N-1} f(r, s) h((m-r)_M, (n-s)_N) \quad (1.62)$$

nebo

$$h(m, n) * f(m, n) = \frac{1}{MN} \sum_{r=0}^{M-1} \sum_{s=0}^{N-1} h(r, s) f((m-r)_M, (n-s)_N), \quad (1.63)$$

kde zápis $(p)_q$ znamená p modulo q . V praxi jsou užívány oba způsoby výpočtu. Volba závisí na tom, zda v daném okamžiku více vyhovuje představa, že signál nabývá mimo oblast Ω nulových hodnot nebo představa, že je signál periodický. V kapitole 2 ukážeme, že představa periodického signálu je užitečná zejména v souvislosti s Fourierovou transformací.



Obr.1.1 Průběh výpočtu konvoluce.

Předpokládejme, že funkce h je reprezentována maticí. Při praktických výpočtech často bývá funkce h nenulová pouze v relativně malém počtu bodů (je-li to možné, pak takto funkci h často záměrně volíme, aby časová složitost výpočtu konvoluce byla přijatelná). Při použití dosud uvedených vztahů jsou nenulové hodnoty většinou bohužel soustředěny do rohů matice, což komplikuje algoritmizaci a snižuje efektivnost výpočtu. V uvedeném případě je proto užitečné členy v matici reprezentující funkci h přeskládat a odpovídajícím způsobem modifikovat vztahy pro výpočet konvoluce. Předpokládejme pro jednoduchost, že hodnoty M, N jsou liché a položíme $P=(M-1)/2$, $Q=(N-1)/2$. Zavedme funkci \tilde{h} , která je definována nad oblastí $\{(m, n) \mid m = -P, -P+1, \dots, P; n = -Q, -Q+1, \dots, Q\}$, pomocí vztahu

$$\tilde{h}(m, n) = \begin{cases} h(-m, -n), & m \leq 0, n \leq 0 \\ h(M-m, -n), & m > 0, n \leq 0 \\ h(-m, N-n), & m \leq 0, n > 0 \\ h(M-m, N-n), & m > 0, n > 0 \end{cases} \quad (1.64)$$

S využitím funkce \tilde{h} lze pak pro výpočet konvoluce použít modifikovaného vztahu

$$h(m, n) * f(m, n) = \frac{1}{MN} \sum_{r=-P}^P \sum_{s=-Q}^Q \tilde{h}(r, s) f(m+r, n+s). \quad (1.65)$$

Předpokládejme, že funkční hodnoty $\tilde{h}(m, n)$ jsou nenulové pouze nad oblastí $\{(m, n) \mid m = -R, \dots, R; n = -S, \dots, S\}$, kde $R < P$, $S < Q$. Ve vztahu (1.65) pak stačí sčítat v mezích $-R \leq r \leq R$, $-S \leq s \leq S$ a k praktické reprezentaci funkce \tilde{h} na počítači stačí matice rozměru $(2R+1) \times (2S+1)$ (pro tuto matici bývá často používáno termínu maska). Dále je zajímavé si povšimnout, že předpisem pro cyklickou konvoluci lze vypočítat též konvoluci lineární. Protože tato možnost má praktický význam, rozeberme ji podrobněji. Vyjděme ze vztahu (1.65), jímž se v praxi lineární konvoluce nejčastěji počítá. Uvažujme oblast $\tilde{\Omega} = \{(m, n) \mid m = 0, 1, \dots, M+R-1; n = 0, 1, \dots, N+S-1\}$ a rozšířme definici funkce f na celou oblast $\tilde{\Omega}$ tak, že funkci f položíme nad oblastí $\tilde{\Omega} \setminus \Omega$ rovnu nule. Nyní by již pro čtenáře mělo být snadným cvičením, aby ukázal, že lineární a cyklická konvoluce dávají shodný výsledek.

2 Transformace obrazových signálů

Při zpracování signálů, a to včetně signálů obrazových, se často uplatňují techniky opírající se o transformace signálu. V této kapitole se budeme problematice transformací signálu věnovat podrobněji. Nejprve se zaměříme na obecný pohled na transformace. Dále se pak budeme věnovat transformaci Fourierově, která mezi transformacemi zaujímá významné postavení a v praxi se často uplatňuje. Konečně uvedeme i některé další transformace. Transformací obecně rozumíme získání popisu signálu jako lineární kombinace funkcí jisté zvolené báze. Velmi často je vstupem (nebo naopak výstupem) transformace signál popsán obrazovou funkcí přímo udávající intenzitu jasu v jednotlivých bodech obrazu. Nejprve ukážeme, že i tento nejběžnější způsob reprezentace obrazu lze interpretovat jako lineární kombinaci jistých bázeových funkcí.

2.1 Báze δ

Uvažujme nejprve prostor \mathcal{S} diskrétních signálů, kde definičním oborem signálů je množina $\Omega = \{0, 1, \dots, M-1\}$. V tomto prostoru uvažujme bázi

$$\{\varphi_k(m) = \delta(k-m) \mid k, m = 0, 1, \dots, M-1\}. \quad (2.1)$$

Je zřejmé, že tato báze je ortonormální. V prostoru \mathcal{S} dále uvažujme signál popsáný funkcí $f(m)$ udávající intenzitu signálu v jednotlivých bodech. Ze vztahu (1.2) víme, že každý signál je možné vyjádřit jako lineární kombinaci bázeových funkcí. Je tedy

$$f(m) = \sum_{k=0}^{M-1} F(k) \delta(k-m). \quad (2.2)$$

Hledejme nyní funkci $F(k)$. S využitím vztahu (1.30) máme

$$F(k) = \sum_{m=0}^{M-1} f(m) \delta(k-m) = f(k), \quad k = 0, 1, \dots, M-1. \quad (2.3)$$

Na základě vztahu (2.3) zjišťujeme, že pro bázi dle (2.1) platí $F(k) = f(k)$. Tento závěr není překvapivý. Chtěli jsme zde pouze ukázat, že i běžný způsob praktické reprezentace diskrétního signálu pomocí M -tice $(f_0, f_1, \dots, f_{M-1})$ intenzit lze interpretovat z pohledu použití bázeových funkcí. Analogicky můžeme postupovat také v prostoru spojitých signálů. Zvolme bázeové jádro podle předpisu

$$\varphi(u, x) = \delta(u-x). \quad (2.4)$$

Opět je zřejmé, že takto zvolená báze je ortonormální. Uvažujme nyní v prostoru \mathcal{S} spojitý signál $f(x)$ definovaný nad oblastí Ω . Ze vztahu (1.3) víme, že tento signál je možné zapsat ve tvaru lineární kombinace

$$f(x) = \int_U F(u) \delta(u-x) du. \quad (2.5)$$

Na základě vztahu (1.33) pak máme

$$F(u) = \int_{\Omega} f(x) \delta(u-x) dx = f(u). \quad (2.6)$$

Vztah (2.6) ukazuje, že při volbě báze dle (2.4) je hledaná funkce F rovna přímo funkci f . I zde jsme chtěli ukázat, že i reprezentace signálu funkcí f má interpretaci z pohledu bázeových funkcí. Pro bázi dle vztahu (2.1), (2.4) budeme v dalším textu používat názvu báze δ .

2.2 Transformace signálu

Mějme unitární prostor signálů. V tomto prostoru zvolme bázi. Transformací rozumíme pochod, kdy pro signál popsáný pomocí obrazové funkce (tedy jako lineární kombinace funkcí báze δ) určujeme jeho popis pomocí zvolené báze. Zpětnou (inverzní) transformací pak nazýváme pochod opačný, kdy pro signál popsáný pomocí zvolené báze určujeme obrazovou funkci. Uvažujme nejprve diskrétní jed-

norozměrné signály. Transformací signálu $f(m)$ rozumíme nalezení funkce $F(k)$ ze vztahu (1.2), která signál reprezentuje s ohledem na zvolenou bázi $\{\varphi_k\}$. Je-li báze $\{\varphi_k\}$ ortonormální (což je výhodné a časté), pak lze pro výpočet koeficientů použít vztahu (1.30). Známe-li naopak funkci $F(k)$, pak lze obrazovou funkci $f(m)$ určit jako lineární kombinaci (1.2). Pro diskretní dvojrozměrné signály je odpovídající zobecnění popsáno vztahy (1.34), (1.4).

Transformaci diskretního dvojrozměrného signálu nazveme separabilní, jestliže platí

$$\varphi_{k,l}(m,n) = \varphi_k(m)\varphi_l(n). \quad (2.7)$$

Je-li transformace separabilní, pak ze vztahů (1.34) a (2.7) plyne

$$F(k,l) = \sum_{m=0}^{M-1} \left[\sum_{n=0}^{N-1} f(m,n)\varphi_l^*(n) \right] \varphi_k^*(m). \quad (2.8)$$

Vztah (2.8) ukazuje, že jestliže je transformace separabilní, pak ji lze provádět odděleně po řádcích a po sloupcích bodů oblasti Ω , nad níž je signál definován.

Předpokládejme nyní, že diskretní dvojrozměrný signál je reprezentován jednorozměrným vektorem \mathbf{f} (stačí zvolit pravidlo, podle kterého hodnoty $f(m,n)$ umístíme do jednorozměrného vektoru) a podobně předpokládejme, že i výsledek transformace je reprezentován jednorozměrným vektorem \mathbf{F} . Každý z těchto vektorů má MN prvků. Ze vztahu (1.34) vidíme, že každý prvek vektoru \mathbf{F} je lineární kombinací prvků z \mathbf{f} . Lze tedy transformaci (1.34) vyjádřit též maticovým zápisem

$$\mathbf{F} = \mathbf{\Phi}\mathbf{f}, \quad \mathbf{f} = \mathbf{\Phi}^{-1}\mathbf{F}, \quad (2.9, 10)$$

kde matice $\mathbf{\Phi}$ je rozměru $MN \times MN$. Poznamenejme, že je-li báze $\varphi_{k,l}(m,n)$ ortonormální, pak platí

$$\mathbf{\Phi}^{-1} = \mathbf{\Phi}^{*T}. \quad (2.11)$$

Jestliže je transformace separabilní (vztah (2.7)), pak lze nalézt maticový zápis, který je co do rozměru matic poněkud úspornější. Nechť je tentokrát signál před a po transformaci reprezentován maticemi \mathbf{f} a \mathbf{F} , které jsou rozměru $M \times N$. Separabilní transformaci lze pak vyjádřit ve tvaru (praktický příklad uvedeme později v souvislosti s výpočtem Fourierovy transformace)

$$\mathbf{F} = \mathbf{\Phi}_C \mathbf{f} \mathbf{\Phi}_R^T, \quad \mathbf{f} = \mathbf{\Phi}_C^{-1} \mathbf{F} \mathbf{\Phi}_R^{T-1}. \quad (2.12, 13)$$

Podle očekávání lze analogicky postupovat také v prostoru spojitých signálů. Nechť \mathcal{S} je takový prostor. V \mathcal{S} zvolme bázevé jádro $\varphi(u,x)$. Dále v \mathcal{S} uvažujme signál $f(x)$. Transformací signálu rozumíme stanovení funkce $F(u)$, která uvažovaný signál reprezentuje s ohledem na zvolené bázevé jádro. Splňuje-li bázevé jádro podmínku ortonormality (1.19), pak lze pro výpočet této funkce použít vztahu (1.33). Známe-li naopak funkci $F(u)$, lze podle vztahu (1.3) zapsat funkci $f(x)$ jako lineární kombinaci. Pro spojitě dvojrozměrné signály lze obdobně uplatnit vztahy (1.35), (1.5).

2.3 Fourierova transformace

Ústřední místo mezi transformacemi signálu zaujímá Fourierova transformace. V této podkapitole se zaměříme nejprve na spojitou Fourierovu transformaci, pak na transformaci diskretní. Konečně uvedeme také příklad algoritmu výpočtu rychlé Fourierovy transformace.

2.3.1 Spojitá Fourierova transformace

Uvažujme prostor signálů, kde oblast Ω definice signálu je dvojrozměrný euklidovský prostor E^2 . Fourierovu transformaci obdržíme, jestliže bázevé jádro volíme ve tvaru

$$\varphi(u, v, x, y) = \exp[j2\pi(ux + vy)]. \quad (2.14)$$

Ponecháváme na čtenáři, aby si ověřil, že uvedené bázevé jádro je ortonormální - tj., že ve shodě se vztahem (1.25) platí

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp[j2\pi(u_1x + v_1y)] \exp[-j2\pi(u_2x + v_2y)] dx dy = \delta(u_1 - u_2, v_1 - v_2). \quad (2.15)$$

(Návod: Použijte vlastnost (1.48) Diracova impulzu.) Na základě vztahů (1.35), (1.5) můžeme tedy pro spojitou Fourierovu transformaci dvojrozměrných signálů psát:

$$F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \exp[-j2\pi(ux + vy)] dx dy \quad (2.16)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) [\cos 2\pi(ux + vy) - j \sin 2\pi(ux + vy)] dx dy,$$

$$f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v) \exp[j2\pi(ux + vy)] du dv \quad (2.17)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v) [\cos 2\pi(ux + vy) + j \sin 2\pi(ux + vy)] du dv.$$

Ze vztahu (2.16) vyplývá, že obecně je $F(u, v)$ komplexní funkce. Rozepsáním na reálnou a imaginární složku máme

$$F(u, v) = R(u, v) + jI(u, v). \quad (2.18)$$

Pro amplitudu $|F(u, v)|$ a fázový posuv $\Phi(u, v)$ pak platí

$$|F(u, v)| = \sqrt{R^2(u, v) + I^2(u, v)}, \quad \Phi(u, v) = \arctg[I(u, v) / R(u, v)]. \quad (2.19)$$

Hodnota $|F(u, v)|$ vyjadřuje, jak je frekvence u, v obsažena v původní funkci $f(x, y)$. Funkce $|F(u, v)|^2$ se nazývá energetické spektrum signálu $f(x, y)$. Vztahy (2.16), (2.17) provádí zobrazení komplexního signálového prostoru na sebe a jsou tedy operátorem nad tímto prostorem. Říkáme, že funkce $F(u, v)$ je Fourierovou transformací funkce $f(x, y)$ a funkce $f(x, y)$ je zpětnou (inverzní) Fourierovou transformací funkce $F(u, v)$. Tyto skutečnosti zapisujeme takto:

$$F(u, v) = \mathcal{F}\{f(x, y)\}, \quad f(x, y) = \mathcal{F}^{-1}\{F(u, v)\}. \quad (2.20)$$

Poznamenejme ještě, že při zpracování obrazového signálu $f(x, y)$ chápeme frekvenci jako „frekvenci délkovou“ - tj. kolik period se vejde na jednotku délky obrazu. Je dále zajímavé poukázat na souvislost Fourierovy transformace s Fourierovými řadami, které jsou dobře známé ze základního kurzu matematiky. Nechť $f(x, y)$ je periodická funkce s periodou l_x ve směru osy x a periodou l_y ve směru osy y . Fourierovou řadou funkce $f(x, y)$ pak nazýváme trigonometrickou řadu

$$f(x, y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} c_{mn} \exp\left[+j2\pi\left(\frac{mx}{l_x} + \frac{ny}{l_y}\right)\right] \quad (2.21)$$

$$= \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} c_{mn} \left[\cos 2\pi\left(\frac{mx}{l_x} + \frac{ny}{l_y}\right) + j \sin 2\pi\left(\frac{mx}{l_x} + \frac{ny}{l_y}\right) \right],$$

kde

$$c_{mn} = \frac{1}{l_x l_y} \int_{-l_x/2}^{l_x/2} \int_{-l_y/2}^{l_y/2} f(x, y) \exp\left[-j2\pi\left(\frac{mx}{l_x} + \frac{ny}{l_y}\right)\right] dx dy. \quad (2.22)$$

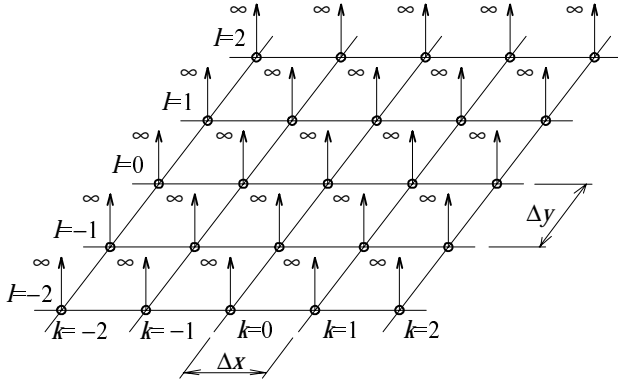
K přechodu od Fourierovy řady k Fourierově transformaci stačí obecnou (a tedy neperiodickou) funkci považovat za speciální případ funkce periodické, kdy je perioda nekonečně dlouhá, a dále vyjít z předpisu (2.21) pro Fourierovu řadu: Uvažujme periodickou funkci $f(x, y)$ s periodami l_x, l_y . Jestliže se periody l_x, l_y prodlužují, pak se podle vztahu (2.21) jednotlivé složky spektra navzájem přibližují až je nakonec při nekonečně dlouhých periodách vzdálenost mezi sousedními složkami spektra nekonečně malá. Diskrétní spektrum se tak změnilo ve spektrum spojité. Pro ilustraci spojité Fourierovy transformace uvedeme několik příkladů, jejichž výsledků později použijeme.

Příklad 2.1. Najdeme Fourierovu řadu pro funkci $f(x,y)$, která je tvořena nekonečným polem Diracových impulzů, které se opakují ve vzdálenostech Δx , Δy . Uvedená funkce má tvar

$$f(x, y) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \delta(x - k\Delta x, y - l\Delta y).$$

Funkce je vyobrazena na obr. 2.1. Podle vztahu (2.22) máme

$$c_{mn} = \frac{1}{\Delta x \Delta y} \int_{-\Delta x/2}^{\Delta x/2} \int_{-\Delta y/2}^{\Delta y/2} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \delta(x - k\Delta x, y - l\Delta y) \exp\left[-j2\pi\left(\frac{mx}{\Delta x} + \frac{ny}{\Delta y}\right)\right] dx dy = \frac{1}{\Delta x \Delta y}.$$



Vztah pro c_{mn} jsme obdrželi s uvážením skutečnosti, že integrujeme po délce jedné periody - v našem případě tedy od $-\Delta x/2$ do $\Delta x/2$ a od $-\Delta y/2$ do $\Delta y/2$, kde $\delta(x - k\Delta x, y - l\Delta y)$ nabývá nenulové hodnoty pouze pro $k=l=0$. Dále jsme využili vlastnost (1.46) Diracova impulzu. Pro hledanou řadu pak podle vztahu (2.21) vychází

$$f(x, y) = \frac{1}{\Delta x \Delta y} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \exp\left[j2\pi\left(\frac{mx}{\Delta x} + \frac{ny}{\Delta y}\right)\right].$$

Obr. 2.1. Pole Diracových impulzů.

Příklad 2.2. Nalezneme Fourierovu transformaci funkce $\delta_n(x,y)$ (odstavec 1.2.3). Na základě předpisu (2.16) postupně dostáváme:

$$\begin{aligned} F(u, v) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta_n(x, y) \exp[-j2\pi(ux + vy)] dx dy = n^2 \int_{-1/2n}^{1/2n} \int_{-1/2n}^{1/2n} \exp[-j2\pi(ux + vy)] dx dy \\ &= n^2 \int_{-1/2n}^{1/2n} \exp[-j2\pi ux] dx \int_{-1/2n}^{1/2n} \exp[-j2\pi vy] dy. \end{aligned}$$

S využitím Eulerova vztahu $e^{j\varphi} = \cos\varphi + j\sin\varphi$ a po integraci pak vyjde

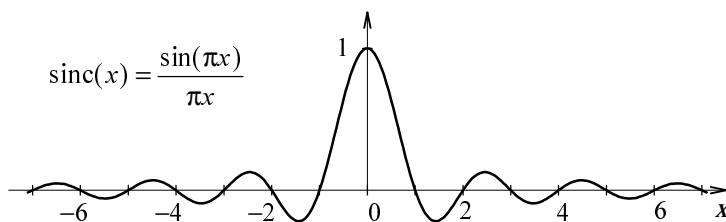
$$F(u, v) = n^2 \frac{\sin(\pi u/n)}{\pi u} \frac{\sin(\pi v/n)}{\pi v}.$$

Zavedeme funkci

$$\text{sinc}(u, v) = \frac{\sin(\pi u)}{\pi u} \frac{\sin(\pi v)}{\pi v}.$$

Pak vychází

$$F(u, v) = \mathcal{F}\{\delta_n(x, y)\} = \frac{\sin(\pi u/n)}{\pi u/n} \frac{\sin(\pi v/n)}{\pi v/n} = \text{sinc}\left(\frac{u}{n}, \frac{v}{n}\right). \quad (2.23)$$



Obr. 2.2. Funkce $\text{sinc}(x)$.

Konstatujeme, že obdélníkový impulz $\delta_n(x,y)$ má nekonečné spojité reálné frekvenční spektrum. Amplituda složek spektra je vyjádřena funkcí $\text{sinc}(u/n, v/n)$. Protože se s funkcí sinc budeme setkávat velmi často, uvádíme na obr. 2.2 graf funkce $\text{sinc}(x)$.

Příklad 2.3. Nalezneme Fourierovu transformaci Diracova impulsu $\delta(x,y)$. Využijeme výsledku z předchozího příkladu, kde položíme $n \rightarrow \infty$. Máme proto

$$\mathcal{F}\{\delta(x,y)\} = \lim_{n \rightarrow \infty} \left[\text{sinc}\left(\frac{u}{n}, \frac{v}{n}\right) \right] = 1.$$

Frekvenční spektrum Diracova impulsu $\delta(x,y)$ je opět spojité reálné a nekonečné. Amplituda všech složek spektra je konstantní a je rovna 1. •

Příklad 2.4. Nalezneme Fourierovu transformaci funkce $f(x,y)$, která je tvořena nekonečným polem Diracových impulsů opakujících se ve vzdálenostech $\Delta x, \Delta y$:

$$f(x,y) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \delta(x - k\Delta x, y - l\Delta y).$$

Uvedená funkce je vyobrazena na obr. 2.1. V příkladě 2.1 jsme pro tuto funkci již našli Fourierovu řadu (funkce je periodická). Abychom obdrželi „pěkný“ výsledek, provedeme nyní místo Fourierovy transformace původní funkce transformaci jejího Fourierova rozvoje, který jsme obdrželi v příkladu 2.1. Z příkladu 2.1 máme

$$f(x,y) = \frac{1}{\Delta x \Delta y} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \exp\left[j2\pi\left(\frac{mx}{\Delta x} + \frac{ny}{\Delta y}\right)\right].$$

Dosažením uvedeného vztahu do předpisu (2.16) získáme

$$F(u,v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{\Delta x \Delta y} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \exp\left[j2\pi\left(\frac{mx}{\Delta x} + \frac{ny}{\Delta y}\right)\right] \exp[-j2\pi(ux + vy)] dx dy.$$

Po záměně pořadí integrování a sumace a po úpravě dostaneme

$$F(u,v) = \frac{1}{\Delta x \Delta y} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\left\{-j2\pi\left[\left(u - \frac{m}{\Delta x}\right)x + \left(v - \frac{n}{\Delta y}\right)y\right]\right\} dx dy.$$

S uvážením vlastnosti (1.48) Diracova impulsu pak konečně obdržíme vztah

$$F(u,v) = \frac{1}{\Delta x \Delta y} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \delta\left(u - \frac{m}{\Delta x}, v - \frac{n}{\Delta y}\right).$$

Vidíme, že Fourierovým obrazem nekonečného pole Diracových impulsů vzdálených vzájemně o $\Delta x, \Delta y$ je opět nekonečné pole Diracových impulsů, kde jsou impulsy vzájemně vzdáleny o $1/\Delta x, 1/\Delta y$. •

Příklad 2.5. Ukažte, že Fourierovou transformací funkce $\cos(2\pi ax)$ je funkce $[\delta(u-a) + \delta(u+a)]/2$. Provedení příkladu ponecháváme čtenáři jako cvičení. Návod: použijte vztahu $\cos\varphi = (e^{j\varphi} + e^{-j\varphi})/2$ a dále použijte vlastnosti (1.48) Diracova impulsu. Z příkladu vidíme, že spektrum funkce $\cos(2\pi ax)$ obsahuje pouze složky $u = a, u = -a$. Takový výsledek jsme ovšem jistě intuitivně očekávali. •

2.3.2 Vlastnosti spojité Fourierovy transformace

V tomto odstavci uvedeme několik vlastností Fourierovy transformace, které se ve zpracování obrazu často uplatňují. Protože důkazy uvedených vlastností jsou vesměs jednoduché, ponecháváme je čtenáři jako cvičení (zpravidla však uvádíme stručný návod, jak při důkazu postupovat). Ve shodě s dříve používaným značením je i zde $F(u,v) = \mathcal{F}\{f(x,y)\}$ Fourierův obraz funkce $f(x,y)$, $G(u,v) = \mathcal{F}\{g(x,y)\}$ Fourierův obraz funkce $g(x,y)$; a, b jsou čísla. Obecně mohou být funkce $f(x,y), g(x,y)$ i čísla a, b komplexní.

a) Linearita: $\mathcal{F}\{af(x,y) + bg(x,y)\} = a\mathcal{F}\{f(x,y)\} + b\mathcal{F}\{g(x,y)\} = aF(u,v) + bG(u,v).$ (2.24)

K důkazu této vlastnosti použijte předpis (2.16) pro Fourierovu transformaci. Dále již vlastnost přímo vyplývá z vlastností integrování.

b) Měřítko: Necht' α, β jsou reálná čísla. Pak platí

$$\mathcal{F}\{f(\alpha x, \beta y)\} = F(u/\alpha, v/\beta)/|\alpha\beta|. \quad (2.25)$$

K důkazu této vlastnosti vyjděte z předpisu pro Fourierovu transformaci a zaveďte transformaci proměnných $x'=\alpha x, y'=\beta y$.

c) Posuv:
$$\mathcal{F}\{f(x-x_0, y-y_0)\} = F(u, v)\exp[-j2\pi(ux_0+vy_0)]. \quad (2.26a)$$

K důkazu této vlastnosti vyjděte z předpisu pro Fourierovu transformaci a zaveďte transformaci proměnných $x'=x-x_0, y'=y-y_0$. Podobnou vlastnost lze nalézt také ve frekvenční doméně:

$$\mathcal{F}^{-1}\{F(u-u_0, v-v_0)\} = f(x, y)\exp[j2\pi(u_0x+v_0y)]. \quad (2.26b)$$

Vlastnost (2.26b) bývá nazývána modulací. Ponecháváme na čtenáři, aby promyslel oprávněnost tohoto názvu.

d) Reflexe:
$$\mathcal{F}\{f(-x, y)\} = F(-u, v), \quad \mathcal{F}\{f(x, -y)\} = F(u, -v), \quad \mathcal{F}\{f(-x, -y)\} = F(-u, -v) \quad (2.27a)$$

a také
$$\mathcal{F}\{f^*(-x, -y)\} = F^*(u, v), \quad (2.27b)$$

kde * označuje komplexně sdruženou veličinu. K důkazu vlastností (2.27a) opět postačí substituce $x'=-x, y'=-y$. K důkazu vlastnosti (2.27b) použijte předpisu (2.16) pro $F(u, v)$, proveďte komplexní sdružení a výsledek dále upravte.

e) Konvoluční teorém:
$$\mathcal{F}\{f(x, y) * g(x, y)\} = F(u, v) G(u, v), \quad (2.28a)$$

$$\mathcal{F}\{f(x, y) g(x, y)\} = F(u, v) * G(u, v). \quad (2.28b)$$

Vlastnost (2.28a,b) je významná a později ji mnohokrát využijeme. Říká, že Fourierovým obrazem konvoluce funkcí f, g je součin Fourierových obrazů F, G obou funkcí. Podobně je Fourierovým obrazem součinu funkcí f, g konvoluce Fourierových obrazů F, G obou funkcí. S ohledem na mimořádný význam uvedených tvrzení provedeme tentokrát i důkaz.

Důkaz. Dokážeme vlastnost (2.28a). Důkaz druhé vlastnosti lze provést obdobně. Víme, že konvoluce funkcí je definována předpisem

$$f(x, y) * g(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(a, b)g(x-a, y-b)dadb.$$

Pro Fourierovu transformaci konvoluce pak podle předpisu (2.16) máme

$$\mathcal{F}\{f(x, y) * g(x, y)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(a, b)g(x-a, y-b)dadb \right] \exp[-j2\pi(ux+vy)]dxdy.$$

Záměnou pořadí integrace dostáváme

$$\mathcal{F}\{f(x, y) * g(x, y)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(a, b) \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x-a, y-b)\exp[-j2\pi(ux+vy)]dxdy \right] dadb.$$

Uplatněním vlastnosti (2.26a) získáme

$$\mathcal{F}\{f(x, y) * g(x, y)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(a, b)G(u, v)\exp[-j2\pi(ua+vb)]dadb.$$

Tento vztah lze pak již konečně upravit na tvar

$$\mathcal{F}\{f(x, y) * g(x, y)\} = G(u, v) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(a, b)\exp[-j2\pi(ua+vb)]dadb = F(u, v)G(u, v). \quad \bullet$$

f) Parsevalův teorém: Předpokládejme, že obecně jsou funkce $f(x, y), g(x, y)$ komplexní a použijme symbolu * pro označení funkce komplexně sdružené. Parsevalův teorém tvrdí:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) g^*(x, y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v) G^*(u, v) du dv. \quad (2.29)$$

Ve speciálním případě, kdy je $f(x, y) = g(x, y)$, přechází uvedený výraz na tvar

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f(x, y)|^2 dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |F(u, v)|^2 du dv. \quad (2.30)$$

Výraz na levé straně vztahu (2.30) může být interpretován jako celkový výkon signálu $f(x, y)$. Výraz $|F(u, v)|^2$ má proto význam výkonového spektra signálu $f(x, y)$.

g) Korelační teorém: Opět předpokládejme, že obecně jsou funkce $f(x, y)$, $g(x, y)$ komplexní. Křížová korelace R_{fg} a autokorelace R_{ff} deterministických funkcí $f(x, y)$ a $g(x, y)$ jsou definovány pomocí vztahů

$$R_{fg}(a, b) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f^*(x - a, y - b) g(x, y) dx dy, \quad (2.31)$$

$$R_{ff}(a, b) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f^*(x - a, y - b) f(x, y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x + a, y + b) f^*(x, y) dx dy. \quad (2.32)$$

Ne náhodou uvedené vztahy připomínají čtenáři vztah pro konvoluci. Skutečně, jediný rozdíl je ve znaménku členu $(x - a, y - b)$. Také tvrzení korelačního teorému koresponduje s tvrzením teorému konvolučního. Korelační teorém totiž tvrdí:

$$\mathcal{F}\{R_{fg}\} = F^*(u, v) G(u, v), \quad (2.33)$$

$$\mathcal{F}\{R_{ff}\} = F^*(u, v) F(u, v) = |F(u, v)|^2. \quad (2.34)$$

Důkaz vztahu (2.33) lze provést podobně jako u konvolučního teorému. Vztah (2.34) vyplývá ze vztahu (2.33).

h) Derivace: Pro Fourierův obraz derivace funkce platí vztahy

$$\mathcal{F}\left\{\frac{\partial f(x, y)}{\partial x}\right\} = j2\pi u F(u, v), \quad \mathcal{F}\left\{\frac{\partial f(x, y)}{\partial y}\right\} = j2\pi v F(u, v). \quad (2.35)$$

K důkazu této vlastnosti stačí vyjít z definice Fourierovy transformace a použít integrace per partes. Při tom předpokládáme, že platí $f(-\infty, y) = f(\infty, y) = f(x, -\infty) = f(x, \infty) = 0$. Analogicky lze také odvodit např. vztah pro Fourierův obraz laplaciánu funkce. Dostaneme

$$\mathcal{F}\left\{\frac{\partial^2 f(x, y)}{\partial x^2} + \frac{\partial^2 f(x, y)}{\partial y^2}\right\} = -4\pi^2 (u^2 + v^2) F(u, v). \quad (2.36)$$

2.3.3 Diskrétní Fourierova transformace

Uvažujme prostor signálů, kde je oblast Ω , nad níž jsou signály definovány, tvořena množinou bodů, které indexujeme pomocí dvojice indexů $m = 0, 1, \dots, M-1$, $n = 0, 1, \dots, N-1$. Diskrétní Fourierovu transformaci v tomto prostoru obdržíme, jestliže bázi volíme ve tvaru

$$\Phi_{k,l}(m, n) = \frac{1}{\sqrt{MN}} \exp\left[j2\pi\left(\frac{mk}{M} + \frac{nl}{N}\right)\right], \quad k = 0, 1, \dots, M-1, \quad l = 0, 1, \dots, N-1. \quad (2.37)$$

Opět ponecháváme na čtenáři, aby ověřil, že takto zvolená báze je ortonormální - tj., že platí

$$\frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \exp\left[j2\pi\left(\frac{mk_1}{M} + \frac{nl_1}{N}\right)\right] \exp\left[-j2\pi\left(\frac{mk_2}{M} + \frac{nl_2}{N}\right)\right] = \delta(k_1 - k_2, l_1 - l_2). \quad (2.38)$$

Podle vztahu (1.34) pak pro diskrétní Fourierovu transformaci platí

$$F(k, l) = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n) \exp\left[-j2\pi\left(\frac{mk}{M} + \frac{nl}{N}\right)\right], \quad k = 0, 1, \dots, M-1, \quad l = 0, 1, \dots, N-1. \quad (2.39)$$

V souladu se vztahem (1.4) pro zpětnou Fourierovu transformaci dostáváme

$$f(m,n) = \frac{1}{\sqrt{MN}} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} F(k,l) \exp \left[j2\pi \left(\frac{mk}{M} + \frac{nl}{N} \right) \right]. \quad (2.40)$$

Předpokládejme nyní, že signál je reprezentován maticí \mathbf{f} a jeho Fourierův obraz maticí \mathbf{F} . Obě matice mají rozměr $M \times N$. Protože Fourierova transformace je separabilní, lze ji zapsat podle vztahu (2.12) ve tvaru

$$\mathbf{F} = \Phi_C \mathbf{f} \Phi_R. \quad (2.41)$$

Zavedme označení

$$W_M = \exp \left[\frac{-j2\pi}{M} \right]. \quad (2.42)$$

Pro matice Φ_R, Φ_C pak platí

$$\Phi_R = \frac{1}{\sqrt{N}} \begin{bmatrix} W_N^0 & W_N^0 & W_N^0 & \dots & W_N^0 \\ W_N^0 & W_N^1 & W_N^2 & \dots & W_N^{N-1} \\ W_N^0 & W_N^2 & W_N^4 & \dots & W_N^{2(N-1)} \\ \dots & \dots & \dots & \dots & \dots \\ W_N^0 & W_N^{N-1} & W_N^{2(N-1)} & \dots & W_N^{(N-1)^2} \end{bmatrix}, \quad (2.43)$$

$$\Phi_C = \frac{1}{\sqrt{M}} \begin{bmatrix} W_M^0 & W_M^0 & W_M^0 & \dots & W_M^0 \\ W_M^0 & W_M^1 & W_M^2 & \dots & W_M^{M-1} \\ W_M^0 & W_M^2 & W_M^4 & \dots & W_M^{2(M-1)} \\ \dots & \dots & \dots & \dots & \dots \\ W_M^0 & W_M^{M-1} & W_M^{2(M-1)} & \dots & W_M^{(M-1)^2} \end{bmatrix}. \quad (2.44)$$

Zpětnou Fourierovu transformaci lze pak zapsat maticovým vztahem

$$\mathbf{f} = \Phi_C^* \mathbf{F} \Phi_R^*, \quad (2.45)$$

kde * znamená komplexně sdružené matice.

2.3.4 Vlastnosti diskretní Fourierovy transformace

V tomto odstavci uvedeme několik významnějších vlastností diskretní Fourierovy transformace. Jistě nepřekvapí, že většina z nich koresponduje s vlastnostmi Fourierovy transformace spojité. Vlastnosti proto uvedeme pouze ve stručném přehledu. Předpokládejme, že $f(m,n)$, $g(m,n)$ jsou diskretní signály, $F(k,l)$, $G(k,l)$ jsou Fourierovy obrazy těchto signálů, a, b jsou čísla (obecně komplexní). Vlastnosti diskretní Fourierovy transformace pak můžeme formulovat takto:

a) Linearita: $\mathcal{F}\{af(m,n)+bg(m,n)\} = a\mathcal{F}\{f(m,n)\} + b\mathcal{F}\{g(m,n)\} = aF(k,l) + bG(k,l).$ (2.46)

b) Periodicita: Ve výrazu (2.39) pro Fourierovu transformaci jsme předpokládali, že k, l jsou z rozsahu $k = 0, 1, 2, \dots, M-1$ a $l = 0, 1, 2, \dots, N-1$. Pokud bychom se pokoušeli vypočítat hodnoty Fourierova obrazu pro k, l mimo uvedené meze, pak analýzou vztahu (2.39) zjistíme, že platí

$$F(k,-l) = F(k,N-l), \quad F(-k,l) = F(M-k,l), \quad F(-k,-l) = F(M-k,N-l), \quad (2.47)$$

$$F(\alpha M+k, \beta N+l) = F(k,l), \quad \alpha, \beta = 0, \pm 1, \pm 2, \dots \quad (2.48)$$

Podobně také z předpisu (2.40) pro zpětnou transformaci vychází

$$f(-m,n) = f(M-m,n), \quad f(m,-n) = f(m,N-n), \quad f(-m,-n) = f(M-m,N-n), \quad (2.49)$$

$$f(\alpha M+m, \beta N+n) = f(m,n), \quad \alpha, \beta = 0, \pm 1, \pm 2, \dots \quad (2.50)$$

Uvedené vlastnosti vyplývají z periodicity exponenciálního členu ve výrazech (2.39), (2.40).

c) Kruhový posuv: Necht' $f(m,n)$ je periodická funkce s periodou M,N . Pro Fourierovu transformaci funkce $f(m-m_0, n-n_0)$ platí

$$\mathcal{F}\{f(m-m_0, n-n_0)\} = F(k,l) \exp\left[-j2\pi\left(\frac{m_0k}{M} + \frac{n_0l}{N}\right)\right]. \quad (2.51a)$$

Podobnou vlastnost lze nalézt také ve frekvenční doméně:

$$\mathcal{F}^{-1}\{F(k-k_0, l-l_0)\} = f(m,n) \exp\left[j2\pi\left(\frac{mk_0}{M} + \frac{nl_0}{N}\right)\right]. \quad (2.51b)$$

Poznamenejme, že předpoklad periodicity funkce $f(m,n)$ jsme zde zavedli k tomu, aby funkce $f(m-m_0, n-n_0)$ byla definována na celé oblasti $m = 0, 1, \dots, M-1$; $n = 0, 1, \dots, N-1$ pro každé m_0, n_0 .

d) Komplexně sdružená hodnota: Je-li funkce $f(m,n)$ reálná, pak platí

$$F(\alpha M - k, \beta N - l) = F^*(k, l), \quad \alpha, \beta = 1, 2, \dots \quad (2.52)$$

Vlastnost snadno dokážeme, jestliže levou i pravou stranu rovnice (2.52) rozepíšeme podle vztahu (2.39). Pro praktický výpočet diskretní Fourierovy transformace reálných signálů má vlastnost velký význam. S jejím využitím postačí hodnoty $F(k,l)$ počítat přibližně v polovině bodů.

e) Konvoluční teorém: Podobně jako pro spojitou Fourierovu transformaci platí i pro transformaci diskretní, že Fourierova transformace konvoluce funkcí v prostorové doméně je součinem Fourierových obrazů funkcí v doméně frekvenční a naopak součinu v prostorové doméně odpovídá konvoluce v doméně frekvenční. Konvoluce se v tomto případě uvažuje cyklická. Je tedy

$$\mathcal{F}\{f(m,n) * g(m,n)\} = \frac{1}{\sqrt{MN}} F(k,l) G(k,l), \quad (2.53a)$$

$$\mathcal{F}\{f(m,n) g(m,n)\} = \sqrt{MN} F(k,l) * G(k,l). \quad (2.53b)$$

f) Korelační teorém: Definujme nejprve křížovou korelaci R_{fg} a autokorelaci R_{ff} deterministických diskretních funkcí $f(m,n)$, $g(m,n)$:

$$R_{fg}(m,n) = \frac{1}{MN} \sum_{r=0}^{M-1} \sum_{s=0}^{N-1} f^*(r-m, s-n) g(r,s), \quad (2.54)$$

$$R_{ff}(m,n) = \frac{1}{MN} \sum_{r=0}^{M-1} \sum_{s=0}^{N-1} f^*(r-m, s-n) f(r,s) = \frac{1}{MN} \sum_{r=0}^{M-1} \sum_{s=0}^{N-1} f(r+m, s+n) f^*(r,s). \quad (2.55)$$

Korelační teorém tvrdí: $\mathcal{F}\{R_{fg}(m,n)\} = \frac{1}{\sqrt{MN}} F^*(k,l) G(k,l), \quad (2.56)$

$$\mathcal{F}\{R_{ff}(m,n)\} = \frac{1}{\sqrt{MN}} F^*(k,l) F(k,l) = \frac{1}{\sqrt{MN}} |F(k,l)|^2. \quad (2.57)$$

Poznamenejme, že podobně jako u konvoluce se i korelace a autokorelace uvažují cyklické (nebo se předpokládá, že funkce $f(m,n)$, $g(m,n)$ jsou periodické).

g) Parsevalův teorém: Předpokládejme, že obecně jsou signály $f(x,y)$, $g(x,y)$ komplexní a použijme symbolu $*$ pro označení komplexně sdružené funkce. Parsevalův teorém tvrdí:

$$\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m,n) g^*(m,n) = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} F(k,l) G^*(k,l). \quad (2.58)$$

Ve speciálním případě, kdy $f(x,y) = g(x,y)$, přechází uvedený výraz na tvar

$$\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |f(m,n)|^2 = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} |F(k,l)|^2. \quad (2.59)$$

2.3.5 Rychlá Fourierova transformace

Ze vztahu (2.39) vidíme, že algoritmus výpočtu dvojrozměrné diskretní Fourierovy transformace může ve své naivní podobě dosáhnout časové složitosti řádově M^2N^2 operací. Při tak vysoké složitosti by možnosti praktického použití Fourierovy transformace při číslicovém zpracování dvojrozměrných signálů byly velmi omezené. Naštěstí však byly nalezeny algoritmy, jejichž časová složitost je podstatně nižší. Ukážeme zde podrobněji jeden z nich. Efektivnost algoritmu je založena na využití techniky dělení na podproblémy poloviční velikosti, s níž se při konstrukci efektivních algoritmů setkáváme v programování dosti často. Algoritmus, který zde popíšeme vyžaduje, aby pole popisující signál bylo čtvercové ($M=N$) a aby jeho rozměr byl celočíselnou mocninou čísla 2. Pro úspornost zápisu zavedme

$$W_N = \exp\left[\frac{-j2\pi}{N}\right] \text{ a položíme } Q(k,l) = NF(k,l). \quad (2.60)$$

Vztah (2.39) můžeme pak přepsat takto

$$Q(k,l) = S_{00}(k,l) + S_{01}(k,l)W_N^l + S_{10}(k,l)W_N^k + S_{11}(k,l)W_N^{k+l}, \quad (2.61)$$

$$\text{kde } S_{00}(k,l) = \sum_{m=0}^{N/2-1} \sum_{n=0}^{N/2-1} f(2m,2n)W_N^{2mk+2nl}, \quad S_{01}(k,l) = \sum_{m=0}^{N/2-1} \sum_{n=0}^{N/2-1} f(2m,2n+1)W_N^{2mk+2nl}, \quad (2.62)$$

$$S_{10}(k,l) = \sum_{m=0}^{N/2-1} \sum_{n=0}^{N/2-1} f(2m+1,2n)W_N^{2mk+2nl}, \quad S_{11}(k,l) = \sum_{m=0}^{N/2-1} \sum_{n=0}^{N/2-1} f(2m+1,2n+1)W_N^{2mk+2nl}.$$

Z periodicity členů $W_N^{2mk+2nl}$ vyplývá, že pole $S_{00}, S_{01}, S_{10}, S_{11}$ jsou periodická s periodou $N/2$ v obou směrech. S využitím tohoto faktu a s využitím skutečnosti, že $W_N^{N/2} = -1$, pak ze vztahu (2.61) plyne

$$\begin{aligned} Q(k,l) &= S_{00}(k,l) + S_{01}(k,l)W_N^l + S_{10}(k,l)W_N^k + S_{11}(k,l)W_N^{k+l}, \quad (2.63) \\ Q\left(k, l + \frac{N}{2}\right) &= S_{00}(k,l) - S_{01}(k,l)W_N^l + S_{10}(k,l)W_N^k - S_{11}(k,l)W_N^{k+l}, \\ Q\left(k + \frac{N}{2}, l\right) &= S_{00}(k,l) + S_{01}(k,l)W_N^l - S_{10}(k,l)W_N^k - S_{11}(k,l)W_N^{k+l}, \\ Q\left(k + \frac{N}{2}, l + \frac{N}{2}\right) &= S_{00}(k,l) - S_{01}(k,l)W_N^l - S_{10}(k,l)W_N^k + S_{11}(k,l)W_N^{k+l}, \end{aligned}$$

$$\text{kde } k, l = 0, 1, \dots, N/2 - 1.$$

Uvedené vztahy již podávají návod na aplikaci techniky dělení na podproblémy poloviční velikosti. Hledané pole Q rozměru $N \times N$ lze podle vztahů (2.63) určit na základě znalosti čtyř dílčích polí $S_{00}, S_{01}, S_{10}, S_{11}$ rozměru $N/2 \times N/2$. Ze vztahů (2.62) vyplývá, že pole $S_{00}, S_{01}, S_{10}, S_{11}$ jsou diskretní Fourierovou transformací vybraných prvků vstupního pole f (např. pole S_{00} je Fourierovou transformací těch prvků pole f , které mají sudé indexy). K výpočtu polí $S_{00}, S_{01}, S_{10}, S_{11}$ proto můžeme rekurzivně použít právě popisovaného algoritmu. Rekurzivní dělení končí tehdy, jakmile dělením vzniknou pole obsahující právě jediný prvek. Pole obsahující jediný prvek je totiž samo svým Fourierovým obrazem.

Provedme ještě odhad časové složitosti popsáného algoritmu. Předpokládáme, že pro činnost algoritmu máme k dispozici tabulku hodnot W_N^k , které se vyskytují ve vztazích (2.63). Složitost algoritmu budeme hodnotit podle počtu násobení. Nechť C_N je počet komplexních násobení, které algoritmus vykoná při výpočtu transformace pole rozměru N . Platí

$$C_N = 4C_{N/2} + \frac{3}{4}N^2. \quad (2.64)$$

Člen $3N^2/4$ na pravé straně rovnice (2.64) vyjadřuje počet násobení, která je třeba provést při výpočtu pole Q o rozměru $N \times N$ podle vztahů (2.63) (jedná se o počet násobení $S_{01}(k,l)W_N^l, S_{10}(k,l)W_N^k, S_{11}(k,l)W_N^{k+l}$, pole $S_{00}, S_{01}, S_{10}, S_{11}$ mají rozměr $N/2$). Člen $4C_{N/2}$ pak vyjadřuje počet násobení, které-

ho je zapotřebí k výpočtu polí $S_{00}, S_{01}, S_{10}, S_{11}$. S ohledem na rekurzi můžeme vztah (2.64) dále rozepsat takto:

$$C_N = 4C_{N/2} + \frac{3}{4}N^2 = 4 \left[4C_{N/4} + \frac{3}{4} \left(\frac{N}{2} \right)^2 \right] + \frac{3}{4}N^2 = \dots \quad (2.65)$$

Protože hloubka rekurze je $\log_2 N$, vychází z výrazu (2.65) pro počet násobení vztah

$$C_N = \frac{3}{4}N^2 \log_2 N. \quad (2.66)$$

(0,7)	(4,7)	(2,7)	(6,7)	(1,7)	(5,7)	(3,7)	(7,7)
(0,3)	(4,3)	(2,3)	(6,3)	(1,3)	(5,3)	(3,3)	(7,3)
(0,5)	(4,5)	(2,5)	(6,5)	(1,5)	(5,5)	(3,5)	(7,5)
(0,1)	(4,1)	(2,1)	(6,1)	(1,1)	(5,1)	(3,1)	(7,1)
(0,6)	(4,6)	(2,6)	(6,6)	(1,6)	(5,6)	(3,6)	(7,6)
(0,2)	(4,2)	(2,2)	(6,2)	(1,2)	(5,2)	(3,2)	(7,2)
(0,4)	(4,4)	(2,4)	(6,4)	(1,4)	(5,4)	(3,4)	(7,4)
(0,0)	(4,0)	(2,0)	(6,0)	(1,0)	(5,0)	(3,0)	(7,0)

Počet komplexních sčítání je při promyšlené implementaci následující (všechny čtyři vztahy (2.63) počítáme najednou, stejné součty jednoduše nepočítáme dvakrát, takže pro jednu dvojici k, l je zapotřebí celkem 8 komplexních sčítání)

$$A_N = 2N^2 \log_2 N. \quad (2.67)$$

Obr. 2.3. Přeskládání pole $N=8$ pro výpočet rychlé Fourierovy transformace.

K praktické realizaci algoritmu ještě poznamenejme, že mají-li prvky polí vznikajících při rekurzivním dělení ležet vedle sebe, pak je nutné vstupní pole před provedením trans-

formace přeskládat. Pravidlo pro přeskládání je následující: Prvek vstupního pole na pozici m, n přemístíme na pozici m', n' , kde hodnoty m', n' získáme reverzí binární reprezentace hodnot m, n (tj. přečteme binární reprezentaci pozpátku). Ponecháme na čtenáři, aby toto pravidlo za pomoci vztahů (2.62) promyslel a zdůvodnil. Způsob přeskládání pole rozměru 8×8 ukazuje obr. 2.3.

Poznamenejme, že zde naznačeného postupu založeného na využití techniky dělení na podproblémy poloviční velikosti lze použít také pro konstrukci algoritmů rychlé Fourierovy transformace libovolné dimenze. Časová složitost algoritmů tohoto typu v d dimenzionálním prostoru je dána výrazy

$$C_N = \frac{2^d - 1}{2^d} N^d \log_2 N, \quad A_N = dN^d \log_2 N. \quad (2.68)$$

Např. pro jednorozměrnou Fourierovu transformaci uvedeným postupem snadno odvodíme vztahy

$$Q(k) = \sqrt{N} F(k), \quad (2.69)$$

$$Q(k) = S_0(k) + S_1(k) W_N^k, \quad (2.70)$$

$$S_0(k) = \sum_{m=0}^{N/2-1} f(2m) W_N^{2mk}, \quad S_1(k) = \sum_{m=0}^{N/2-1} f(2m+1) W_N^{2mk}, \quad (2.71)$$

$$Q(k) = S_0(k) + S_1(k) W_N^k, \quad Q\left(k + \frac{N}{2}\right) = S_0(k) - S_1(k) W_N^k, \quad (2.72)$$

$$C_N = \frac{1}{2} N \log_2 N, \quad A_N = N \log_2 N. \quad (2.73)$$

Na závěr tohoto odstavce o rychlé Fourierově transformaci naznačíme ještě jednu možnost, která je zajímavá zejména svojí jednoduchostí. Vztah (2.39) lze přepsat na tvar

$$F(k, l) = \frac{1}{\sqrt{M}} \sum_{m=0}^{M-1} \left\{ \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} f(m, n) \exp\left[-j2\pi \frac{nl}{N}\right] \right\} \exp\left[-j2\pi \frac{mk}{M}\right]. \quad (2.74)$$

Tento vztah ukazuje, že dvojrozměrnou Fourierovu transformaci pole f lze vypočítat tak, že provedeme nejprve jednorozměrnou Fourierovu transformaci všech řádků (výraz ve složené závorce) a pak jednorozměrnou Fourierovu transformaci všech sloupců. Poznamenejme, že se jedná o důsledek separability Fourierovy transformace. Vyhodnoňme časovou složitost tohoto postupu, kterou budeme opět měřit počtem násobení. Pokud bychom pro jednorozměrnou Fourierovu transformaci řádků a sloupců

použili naivního algoritmu s kvadratickou složitostí, vyšla by pro pole o rozměrech $M \times N$ celková složitost algoritmu $MN^2 + NM^2$. Zvolíme-li však pro transformaci řádků a sloupců rychlou Fourierovu transformaci se složitostí $C_N = 0.5M \log_2 N$ (vztah 2.73), pak je celková složitost algoritmu $0.5(MN \log_2 N + NM \log_2 M)$. Je-li pole f čtvercové o rozměru $N \times N$, pak pro celkovou složitost vyjádřenou počtem komplexních násobení vychází hodnota $N^2 \log_2 N$. Vidíme tedy, že tento algoritmus je jen o málo horší než algoritmus, který jsme popsali dříve.

2.4 Kosinová transformace

V kosinové transformaci se využívá skutečnosti, že Fourierova transformace funkce, která je symetrická podle obou souřadných os, má pouze reálnou složku - imaginární složka vyjde všude nulová. Přítomnost pouze reálné složky je výhodná z hlediska paměťové i časové složitosti výpočtů. Symetrii funkce lze zajistit postupem, který dále popíšeme. Mějme diskrétní funkci $f(m, n)$ definovanou nad oblastí $N \times N$ bodů ($m, n = 0, 1, \dots, N-1$). Tuto funkci doplníme tak, abychom vytvořili funkci symetrickou. To lze provést dvěma způsoby. První z nich vede na funkci definovanou nad oblastí $2N \times 2N$ bodů ($m, n = -N, -N+1, \dots, N-1$) a nazývá se proto sudou kosinovou transformací. Druhý vede na funkci definovanou nad oblastí $(2N-1) \times (2N-1)$ bodů ($m, n = -N+1, -N+2, \dots, N-1$) a nazývá se proto kosinovou transformací lichou. Podrobněji ukážeme nejprve druhý způsob - tj. lichou kosinovou transformaci. Označme symetrickou funkci $\tilde{f}(m, n)$ a definujme ji podle předpisu

$$\tilde{f}(m, n) = \begin{cases} f(m, n), & m \geq 0, n \geq 0 \\ f(-m, n), & m < 0, n \geq 0 \\ f(m, -n), & m \geq 0, n < 0 \\ f(-m, -n), & m < 0, n < 0 \end{cases} \quad (2.75)$$

Uplatníme-li na funkci $\tilde{f}(m, n)$ diskrétní Fourierovu transformaci, máme

$$F(k, l) = \frac{1}{2N-1} \sum_{m=-N+1}^{N-1} \sum_{n=-N+1}^{N-1} \tilde{f}(m, n) \exp \left[-j2\pi \left(\frac{mk}{2N-1} + \frac{nl}{2N-1} \right) \right], \quad (2.76)$$

$k, l = -N+1, \dots, N-1.$

Vezmeme-li v úvahu, že funkce $\tilde{f}(m, n)$ je symetrická, pak po úpravě dostaneme

$$F(k, l) = \frac{1}{2N-1} \sum_{m=-N+1}^{N-1} \sum_{n=-N+1}^{N-1} \tilde{f}(m, n) \cos \left(2\pi \frac{mk}{2N-1} \right) \cos \left(2\pi \frac{nl}{2N-1} \right). \quad (2.77)$$

S ohledem na symetrii dle (2.75) lze tento vztah dále přepsat do tvaru

$$F(k, l) = \frac{4}{2N-1} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} c(m)c(n) f(m, n) \cos \left(2\pi \frac{mk}{2N-1} \right) \cos \left(2\pi \frac{nl}{2N-1} \right), \quad (2.78)$$

kde,
$$c(p) = \begin{cases} 1/2, & p = 0 \\ 1, & \text{jinak} \end{cases} \quad (2.79)$$

Funkce c zohledňuje skutečnost, že ve čtyřech symetrických kvadrantech funkce $\tilde{f}(m, n)$ vytvořených podle vztahu (2.75) jsou hodnoty na osách symetrie sdíleny dvěma kvadranty a hodnota v počátku je sdílena čtyřmi kvadranty. Protože funkce $\cos(x)$ je sudá, zjišťujeme na základě vztahu (2.78), že platí $F(k, -l) = F(k, l)$, $F(-k, l) = F(k, l)$, $F(-k, -l) = F(k, l)$ a že tedy i frekvenční spektrum lze rozdělit na čtyři symetrické kvadranty. Hodnoty $F(k, l)$ proto postačí počítat a uchovávat pouze pro nezáporné indexy k, l . Symetrie lze využít i při provádění zpětné kosinové transformace. Z předpisu pro zpětnou Fourierovu transformaci obdržíme pro zpětnou transformaci kosinovou vztah

$$f(m, n) = \frac{4}{2N-1} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} c(k)c(l) F(k, l) \cos \left(2\pi \frac{mk}{2N-1} \right) \cos \left(2\pi \frac{nl}{2N-1} \right). \quad (2.80)$$

Vzorce pro sudou kosinovou transformaci lze odvodit analogicky. Symetrie je v tomto případě dosaženo zrcadlením původního obrazu kolem přímek $x = -1/2$, $y = -1/2$. Aby ve Fourierově transformaci vedla symetrie k anulování členů obsahujících funkci sinus, je zapotřebí posunout počátek souřadné soustavy do bodu o souřadnicích $(-1/2, -1/2)$. Původní souřadnice m, n se tak transformují na hodnoty $(m+1/2)$, $(n+1/2)$. V dopředné sudé kosinové transformaci neleží na osách symetrie žádné řady bodů obrazu. Také v případě sudé kosinové transformace platí $F(k, -l) = F(k, l)$, $F(-k, l) = F(k, l)$, $F(-k, -l) = F(k, l)$, a proto i zde postačí počítat a uchovávat pouze jeden kvadrant frekvenčního spektra. Hodnoty $F(k, l)$ ležící na osách symetrie jsou sdíleny sousedními kvadranty spektra. Při výpočtu zpětné transformace lze tuto skutečnost vzít opět v úvahu použitím funkce c . Zpravidla se polovina uvedené korekce realizuje už při dopředné transformaci a polovina při transformaci zpětné. Pro dopřednou a zpětnou sudou kosinovou transformaci tak máme

$$F(k, l) = \frac{4}{2N} c(k)c(l) \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} f(m, n) \cos \left[\pi \frac{(2m+1)k}{2N} \right] \cos \left[\pi \frac{(2n+1)l}{2N} \right], \quad (2.81)$$

$$f(m, n) = \frac{4}{2N} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} c(k)c(l) F(k, l) \cos \left[\pi \frac{(2m+1)k}{2N} \right] \cos \left[\pi \frac{(2n+1)l}{2N} \right], \quad (2.82)$$

kde,
$$c(p) = \begin{cases} 1/\sqrt{2}, & p = 0 \\ 1, & \text{jinak} \end{cases}. \quad (2.83)$$

Na závěr tohoto odstavce poznamenejme, že kosinová transformace nachází významné uplatnění při kompresi obrazových signálů. Využívá se v kompresních metodách JPEG a MPEG, kde se používá kosinové transformace sudé.

2.5 Waveletové transformace

V poslední době se díky některým svým výhodným vlastnostem uplatňují ve zpracování obrazu také tzv. waveletové transformace. Stejně jako ostatní transformace, je i transformace waveletová obecně popsána matematickým aparátem zavedeným v předchozí a dále rozváděným v této kapitole. Speciální je ovšem volba básových (souřadnicových) funkcí. Podrobnější informace uvedeme později. Pro ilustraci však již nyní poznamenejme, že významným rysem waveletové transformace je to, že jednotlivé básové funkce jsou lokalizované, čímž rozumíme že funkční hodnoty nad jistou „relativně malou“ oblastí dominují nad funkčními hodnotami ve zbytku definičního oboru (to je významný rozdíl ve srovnání např. s transformací Fourierovou). Z této skutečnosti pak vyplývá většina výhod waveletové transformace. Básové funkce nejsou ve waveletové transformaci pevně předspsány. Změnou průběhu básových funkcí můžeme měnit vlastnosti transformace podle požadavků, které jsou v řešené aplikaci kladeny. Na obr. 2.4, 2.5 je uveden příklad znázorňující typickou volbu básových funkcí. Waveletové transformace dnes již představují rozvinutý a mocný prostředek se širokým uplatněním. V tomto odstavci se proto z pochopitelných důvodů musíme omezit jen na základní informace.

2.5.1 Spojitá waveletová transformace

Začneme jednorozměrným spojitým případem. Nechť $\psi(x)$ je reálná nebo komplexní funkce $\psi(x) \in L^2(\mathbf{R})$ a $\Psi(u)$ nechť je její Fourierův obraz (zápis $L^2(\mathbf{R})$ označuje množinu funkcí integrovatelných se čtvercem, tj. funkcí, pro které platí $\int_{-\infty}^{\infty} |\psi(x)|^2 dx < \infty$). Je-li splněno tzv. kritérium přípustnosti (uvedené kritérium je postačující podmínkou pro invertibilitu waveletové transformace (Chui 92))

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\Psi(u)|^2}{|u|} du < \infty, \quad (2.84)$$

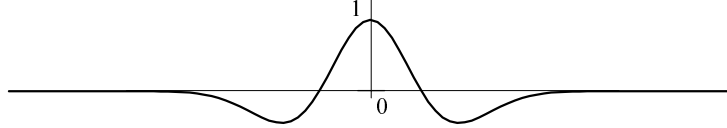
pak lze funkci $\psi(x)$ využít ke konstrukci waveletové báze. Funkci $\psi(x)$ budeme v tomto textu nazývat základním waveletem nebo stručně pouze waveletem. Protože se ve jmenovateli integrandu rovnice (2.84) nachází hodnota $|u|$, vyžaduje kritérium přípustnosti splnění podmínky

$$\Psi(0) = 0 \Rightarrow \int_{-\infty}^{\infty} \psi(x) dx = 0. \quad (2.85)$$

Intuitivně také dále očekáváme $\Psi(\infty)=0$. Vidíme tedy, že se průběh frekvenčního spektra přípustného waveletu podobá frekvenční charakteristice pásmové propusti. Opačně také impulzová charakteristika nějaké pásmové propusti, vyhovuje-li vztahům (2.84), (2.85), může být použita jako wavelet. Typickým příkladem základního waveletu je funkce ve tvaru „mexického klobouku“ popsaná následujícím předpisem (obr. 2.4)

$$\psi(x) = (1 - 2x^2) \exp(-x^2). \quad (2.86)$$

Množina $\{\psi_{a,b}(x)\}$ funkcí tvořících bázi waveletové transformace může být generována translací a změnou měřítka základního waveletu



Obr. 2.4 Wavelet $\psi(x)=(1-2x^2)\exp(-x^2)$.

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right), \quad (2.87)$$

kde $a > 0$, b jsou reálná čísla; a vyjadřuje měřítko, b vyjadřuje posuv podél osy x . Koeficient $1/\sqrt{a}$ na pr-

vé straně rovnice (2.87) zajišťuje, aby všechny funkce báze měly stejnou normu. Je totiž

$$\|\psi_{a,b}(x)\| = \frac{1}{\sqrt{a}} \sqrt{\int_{-\infty}^{\infty} \left| \psi\left(\frac{x-b}{a}\right) \right|^2 dx} = \|\psi(x)\|. \quad (2.88)$$

Ve vztahu (2.87) mohou a, b obecně nabývat libovolné hodnoty. Speciální volba hodnot a, b dává vzniknout speciálním variantám waveletové báze, které mohou být zajímavé z praktického hlediska. Takovým případům se budeme věnovat později. Při transformaci dvojrozměrných signálů lze postupovat analogicky jako v případě signálů jednorozměrných. Postačí předpokládat existenci přípustného základního waveletu $\psi(x, y)$. Funkce waveletové báze pak opět získáme translací a změnou měřítka

$$\Psi_{a,b_x,b_y}(x, y) = \frac{1}{|a|} \psi\left(\frac{x-b_x}{a}, \frac{y-b_y}{a}\right). \quad (2.89)$$

Waveletovou transformaci funkce f budeme v tomto textu označovat symbolem F_W . Ve shodě s (1.33), (1.35) je waveletová transformace definována vztahy

$$F_W(a, b) = \langle f, \Psi_{a,b} \rangle = \int_{-\infty}^{\infty} f(x) \Psi_{a,b}^*(x) dx, \quad (2.90)$$

$$F_W(a, b_x, b_y) = \langle f, \Psi_{a,b_x,b_y} \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \Psi_{a,b_x,b_y}^*(x, y) dx dy. \quad (2.91)$$

Je užitečné si povšimnout, že transformací funkce jedné proměnné jsme dle vztahu (2.90) obdrželi obraz, který je funkcí dvou proměnných. Podobně i v transformaci dle vzorce (2.91) vzrostl počet proměnných z 2 na 3. Poznamenáváme rovněž, že za waveletovou transformaci funkcí $f(x)$, $f(x, y)$ je považována hodnota skalárních součinů (2.90), (2.91) i v případě, že waveletové báze $\psi_{a,b}(x)$, $\Psi_{a,b_x,b_y}(x, y)$ nejsou ortonormální (při odvození vztahů (1.33), (1.35) jsme ovšem předpoklad ortonormality použili). Důsledkem z toho vyplývajícím je, že v takovém případě pak $F_W(a, b)$, $F_W(a, b_x, b_y)$ již nemají význam koeficientů v lineární kombinaci (1.3), (1.5). Lze však ukázat, že jestliže je splněna podmínka přípustnosti (2.84), pak pro zpětnou transformaci platí vztahy

$$f(x) = \frac{1}{C_\Psi} \int_0^{\infty} \int_{-\infty}^{\infty} F_W(a, b) \psi_{a,b}(x) db \frac{da}{a^2}, \quad (2.92)$$

$$f(x, y) = \frac{1}{C_\Psi} \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty F_W(a, b_x, b_y) \Psi_{a, b_x, b_y}(x, y) db_x db_y \frac{da}{a^3}. \quad (2.93)$$

Považujeme za poučné a praktické ukázat, že waveletová transformace souvisí s konvolucí (a tedy s filtrací signálu). Uvažujme transformaci jednorozměrného signálu dle rovnice (2.90). Předpokládejme nyní, že funkce $\psi(x)$ je reálná (tedy nikoli komplexní - splnění uvedeného předpokladu lze v praxi často očekávat). Z rovnic (2.87), (2.90) máme

$$\begin{aligned} F_W(a, b) &= \int_{-\infty}^\infty f(x) \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right) dx = \int_{-\infty}^\infty f(x) \frac{1}{\sqrt{a}} \psi\left[-\frac{1}{a}(b-x)\right] dx \\ &= \int_{-\infty}^\infty f(x) \tilde{\psi}_a(b-x) dx = f * \tilde{\psi}_a, \end{aligned} \quad (2.94)$$

kde
$$\tilde{\psi}_a = \frac{1}{\sqrt{a}} \psi\left(-\frac{x}{a}\right). \quad (2.95)$$

Vztah (2.94) ukazuje, že waveletovou transformaci vstupního signálu $f(x)$ lze vypočítat jako konvoluci funkce $f(x)$ s funkcemi $\tilde{\psi}_a(x)$. Funkce $\tilde{\psi}_a(x)$ získáme ze základního waveletu ψ zrcadlovým převrácením a změnou měřítka. Na základě uvedeného poznatku lze konstruovat efektivní algoritmy pro výpočet waveletové transformace.

2.5.2 Waveletová transformace s dyadickou bází

Waveletových bází, kde proměnné a, b, b_x, b_y ve vzorcích (2.87), (2.89) mohou nabývat libovolných hodnot, se v praxi nepoužívá příliš často. Spíše se používají báze s dalšími doplňujícími omezeními. Dyadická waveletová báze používá změny měřítka $a = 2^{-j}$ a translace $b = 2^{-j}k$. Bázové funkce jsou definovány předpisem

$$\Psi_{j,k}(x) = 2^{j/2} \psi(2^j x - k), \quad (2.96)$$

kde $-\infty < j, k < \infty$. Množina $\{\Psi_{j,k}\}$ tvoří ortonormální bázi v $L^2(\mathbf{R})$ jestliže

$$\langle \Psi_{j,k}, \Psi_{l,m} \rangle = \delta(j-l) \delta(k-m) \quad (2.97)$$

a jestliže libovolnou funkci $f(x) \in L^2(\mathbf{R})$ lze zapsat jako součet

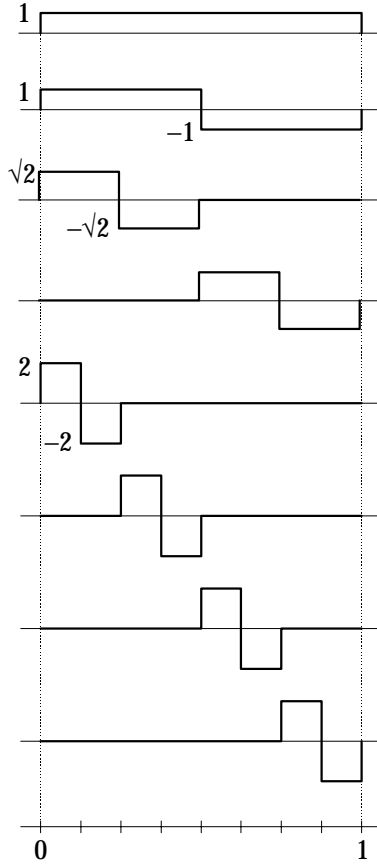
$$f(x) = \sum_{j=-\infty}^\infty \sum_{k=-\infty}^\infty c_{j,k} \Psi_{j,k}(x). \quad (2.98)$$

V případě ortonormální báze poskytuje waveletová transformace také hodnotu koeficientů $c_{j,k}$ ze vztahu (2.98). Je tedy

$$\begin{aligned} F_W(j, k) &= c_{j,k} = \langle f(x), \Psi_{j,k}(x) \rangle \\ &= \sqrt{2^j} \int_{-\infty}^\infty f(x) \psi^*(2^j x - k) dx. \end{aligned} \quad (2.99)$$

Vztahy (2.98), (2.99) popisují rozvoj funkce $f(x)$ vzhledem k waveletu $\psi(x)$. Stojí za povšimnutí, že ačkoli se jedná o rozvoj funkce jedné proměnné, je suma ve vztahu (2.98) dvojnásobná. Při praktické implementaci na počítači samozřejmě není možné pracovat s neomezeným počtem koeficientů. Využívá

Obr.2.5. Dyadická waveletová báze.



Obr.2.6. Bázové funkce Haarovy transformace.

se proto skutečnosti, že při vhodné volbě waveletu $\psi(x)$ jsou hodnoty koeficientů s velkým $|j|, |k|$ zanedbatelné.

Jestliže se omezíme na funkce $f(x)$ a základní wavelety, jejichž funkční hodnota je vně intervalu $\langle 0, 1 \rangle$ vždy 0, pak můžeme systém bázových funkcí specifikovat jediným indexem n . Uvedená báze, označovaná jako kompaktní dyadická báze, je popsána předpisem

$$\psi_n(x) = \sqrt{2^j} \psi(2^j x - k), \quad (2.100)$$

kde hodnoty j, k jsou funkcí n ($n=0, 1, \dots$). Pro $n \geq 1$ se j stanoví jako největší celé číslo, které vyhovuje vztahu $2^j \leq n$, hodnota k je $k = n - 2^j$; pro $n = 0$ je $j = -1, k = 0$. Podmínka ortonormality a waveletovský rozvoj pak mají tvar

$$\langle \psi_m, \psi_n \rangle = \delta(m - n), \quad (2.101)$$

$$f(x) = \sum_{n=0}^{\infty} c_n \psi_n(x), \quad (2.102)$$

$$F_W(n) = c_n = \langle f(x), \psi_n(x) \rangle = \sqrt{2^j} \int_{-\infty}^{\infty} f(x) \psi^*(2^j x - k) dx. \quad (2.103)$$

Příkladem kompaktní dyadické báze je báze Haarovy transformace (není bez zajímavosti, že Haarova transformace byla popsána již v roce 1910, tedy daleko dříve, než byla zpracována teorie waveletů). Bázové funkce Haarovy transformace jsou vyobrazeny na obr. 2.6.

Podmínka ortogonality představuje ve skutečnosti dosti silné omezení. Bylo ukázáno, že jediným reálným waveletem, který je symetrický a současně generuje ortogonální bázi (symetrií rozumíme, že pro některé q je funkce $\psi(x-q)$ je buď sudá nebo lichá), je Haarův wavelet. Aby byla zajištěna větší volnost při volbě bázových funkcí, je možné zavést tzv. biortogonální wavelety (Chui 92). Necht' $\psi, \tilde{\psi}$ jsou dva různé wavelety a $\{\psi_{j,k}\}, \{\tilde{\psi}_{j,k}\}$ necht' jsou báze jimi generované. Báze nazveme biortogonálními, jestliže je splněna následující podmínka

$$\langle \psi_{j,k}, \tilde{\psi}_{l,m} \rangle = \delta(j-l) \delta(k-m). \quad (2.104)$$

Pro dekompozici a zpětnou rekonstrukci pak platí (Chui 92, Castleman 96)

$$c_{j,k} = \langle f(x), \tilde{\psi}_{j,k}(x) \rangle, \quad d_{j,k} = \langle f(x), \psi_{j,k}(x) \rangle, \quad (2.105)$$

$$f(x) = \sum_{j,k} c_{j,k} \psi_{j,k}(x) = \sum_{j,k} d_{j,k} \tilde{\psi}_{j,k}(x). \quad (2.106)$$

Ve vztazích (2.99), (2.103) jsme zatím předpokládali, že $f(x)$ je spojitá funkce. Modifikace uvedených vztahů pro diskretní případ je přímočará. Dostáváme

$$F_W(j,k) = c_{j,k} = \sqrt{2^j} \sum_m f(m) \psi^*(2^j m - k). \quad (2.107)$$

V této souvislosti poznamenejme, že také transformace, v níž a, b v rovnici (2.87) nabývají pouze dyadických hodnot (tj. transformace, kde je báze popsána vztahem (2.96)), bývá někdy označována jako diskretní waveletová transformace. Jak vidíme, dochází v tomto případě ke dvojznačnosti, protože také funkce, která má být transformována, může být diskretní nebo spojitá.

3 Stochastický přístup k popisu obrazových signálů

V předchozích kapitolách jsme předpokládali, že obrazové signály popisujeme deterministicky jejich velikostí v jednotlivých bodech obrazu a zpracováváme jednotlivě, nezávisle na jiných obrazových signálech. Tato představa není ovšem vždy účelná. Často chceme například navrhnout nebo vyšetřit parametry systému zpracovávajícího nějakou třídu obrazových signálů, aniž bychom nějaký konkrétní jednotlivý signál měli předem na mysli. V takových případech lze použít přístupu stochastického. Při stochastickém přístupu je obraz charakterizován jistým souborem statistických veličin. V této kapitole popíšeme základní nástroje, které se při stochastickém přístupu používají. Ke studiu kapitoly jsou potřebné základní znalosti z počtu pravděpodobnosti (ve stručném přehledu je lze nalézt v dodatku A).

3.1 Náhodné pole

Při stochastickém přístupu se signál považuje za náhodný proces (jednorozměrné signály) nebo za náhodné pole (vícerozměrné signály). Zavedení uvedených pojmů je věnována tato podkapitola. Necht' \mathcal{S} je oblast v m -rozměrném euklidovském prostoru. Necht' R označuje bod této oblasti. Poloha bodu R je popsána vektorem \mathbf{r} . Definujme nyní množinu $\{\mathbf{f}_R\}$ náhodných proměnných tak, že pro každý bod oblasti \mathcal{S} zavedeme jednu náhodnou proměnnou. Je-li $m = 1$ (jednorozměrný signál), pak se takto vytvořená množina náhodných proměnných nazývá stochastickým (náhodným) procesem. Je-li $m > 1$, pak se používá termínu náhodné pole nebo termínu vícerozměrný stochastický proces. Necht' Ω označuje množinu všech elementárních jevů (dodatek A) a ω_i je i -tý elementární jev. Na náhodný proces (náhodné pole) lze nahlížet ze dvou pohledů: 1) Jedná se o množinu $\{\mathbf{f}_R(\omega_i) \mid R \in \mathcal{S}\}$ náhodných proměnných. 2) Jedná se o množinu $\{f_{\omega_i}(\mathbf{r}) \mid \omega_i \in \Omega\}$ funkcí nad oblastí \mathcal{S} - jedna funkce pro každý elementární jev ω_i . Pro náhodný proces (náhodné pole) zavedeme značení $\mathbf{f}(\mathbf{r}, \omega_i)$ nebo také pouze stručně $\mathbf{f}(\mathbf{r})$. Při zpracování obrazů pracujeme nejčastěji s dvojrozměrnými náhodnými poli $\mathbf{f}(x, y)$. Vezme-li v úvahu, že pro každou hodnotu \mathbf{r} je $\mathbf{f}(\mathbf{r})$ náhodnou proměnnou v obvyklém smyslu, pak není obtížné zavést distribuční funkci $P_f(z, \mathbf{r})$ a hustotu pravděpodobnosti $p_f(z, \mathbf{r})$ náhodného pole. Jak distribuční funkce, tak hustota pravděpodobnosti jsou v tomto případě také funkcí polohy \mathbf{r} . Distribuční funkce $P_f(z, \mathbf{r})$ je definována obvyklým způsobem (dodatek A) - tj. jako pravděpodobnost události $\{\mathbf{f}(\mathbf{r}) \leq z\}$. Je tedy

$$P_f(z, \mathbf{r}) = \mathcal{P}\{\mathbf{f}(\mathbf{r}) \leq z\}. \quad (3.1)$$

Hustotu pravděpodobnosti $p_f(z, \mathbf{r})$ získáme jako obvykle derivováním distribuční funkce:

$$p_f(z, \mathbf{r}) = dP_f(z, \mathbf{r})/dz. \quad (3.2)$$

Uvažujme nyní v oblasti \mathcal{S} n bodů $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n$. Podle dříve uvedeného odpovídají těmto bodům náhodné proměnné $\mathbf{f}(\mathbf{r}_1), \mathbf{f}(\mathbf{r}_2), \dots, \mathbf{f}(\mathbf{r}_n)$. Můžeme tedy pro náhodné pole zavést sdruženou distribuční funkci a sdruženou hustotu pravděpodobnosti n -tého řádu:

$$P_f(z_1, z_2, \dots, z_n, \mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n) = \mathcal{P}\{\mathbf{f}(\mathbf{r}_1) \leq z_1, \mathbf{f}(\mathbf{r}_2) \leq z_2, \dots, \mathbf{f}(\mathbf{r}_n) \leq z_n\}, \quad (3.3)$$

$$p_f(z_1, z_2, \dots, z_n, \mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n) = \frac{\partial^n P_f(z_1, z_2, \dots, z_n, \mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n)}{\partial z_1 \partial z_2 \dots \partial z_n}. \quad (3.4)$$

3.2 Momenty náhodného pole

Předpokládejme na okamžik, že signál reprezentujeme jeho n vzorky získanými v bodech $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n$. Pokud bychom za této situace měli k dispozici sdruženou hustotu pravděpodobnosti n -tého řádu, byl by takto signál vyčerpávajícím způsobem popsán. Bohužel jsou však postupy založené na sdružené pravděpodobnosti vyšších řádů teoreticky i prakticky obtížně zvládnutelné. Nejlépe jsou zatím pracovány postupy založené na sdružené pravděpodobnosti druhého řádu, které vyšetřují, jak jsou svázány vlastnosti obrazu ve dvou jeho různých bodech. Máme-li pro náhodné pole zavedenu sdruže-

nou hustotu pravděpodobnosti druhého řádu, můžeme definovat jeho střední hodnotu, korelaci a kovarianci. Střední hodnota $\mu(\mathbf{r})$ náhodného pole je definována vztahem

$$\mu(\mathbf{r}) = E\{\mathbf{f}(\mathbf{r})\} = \int_{-\infty}^{\infty} z p_f(z, \mathbf{r}) dz. \quad (3.5)$$

Autokorelace $R_{ff}(\mathbf{r}_1, \mathbf{r}_2)$ náhodného pole $\mathbf{f}(\mathbf{r})$ je definována jako korelace náhodných proměnných $\mathbf{f}(\mathbf{r}_1)$, $\mathbf{f}(\mathbf{r}_2)$. Je tedy

$$R_{ff}(\mathbf{r}_1, \mathbf{r}_2) = E\{\mathbf{f}(\mathbf{r}_1)\mathbf{f}(\mathbf{r}_2)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z_1 z_2 p_f(z_1, z_2, \mathbf{r}_1, \mathbf{r}_2) dz_1 dz_2. \quad (3.6)$$

Podobně autokovariance $C_{ff}(\mathbf{r}_1, \mathbf{r}_2)$ náhodného pole je definována jako kovariance náhodných proměnných $\mathbf{f}(\mathbf{r}_1)$, $\mathbf{f}(\mathbf{r}_2)$. Odtud máme

$$C_{ff}(\mathbf{r}_1, \mathbf{r}_2) = E\{[\mathbf{f}(\mathbf{r}_1) - \mu_f(\mathbf{r}_1)][\mathbf{f}(\mathbf{r}_2) - \mu_f(\mathbf{r}_2)]\}. \quad (3.7)$$

Roznásobením výrazů v hranatých závorkách a uplatněním operace střední hodnoty na každý z takto vzniklých členů lze snadno dokázat, že platí

$$C_{ff}(\mathbf{r}_1, \mathbf{r}_2) = R_{ff}(\mathbf{r}_1, \mathbf{r}_2) - \mu_f(\mathbf{r}_1)\mu_f(\mathbf{r}_2). \quad (3.8)$$

Analogicky můžeme zavést křížovou korelaci $R_{fg}(\mathbf{r}_1, \mathbf{r}_2)$ a křížovou kovarianci $C_{fg}(\mathbf{r}_1, \mathbf{r}_2)$ náhodných polí $\mathbf{f}(\mathbf{r})$ a $\mathbf{g}(\mathbf{r})$:

$$R_{fg}(\mathbf{r}_1, \mathbf{r}_2) = E\{\mathbf{f}(\mathbf{r}_1) \mathbf{g}(\mathbf{r}_2)\}, \quad (3.9)$$

$$C_{fg}(\mathbf{r}_1, \mathbf{r}_2) = E\{[\mathbf{f}(\mathbf{r}_1) - \mu_f(\mathbf{r}_1)][\mathbf{g}(\mathbf{r}_2) - \mu_g(\mathbf{r}_2)]\}. \quad (3.10)$$

Náhodné pole $\mathbf{f}(\mathbf{r})$ je nekorelované právě tehdy, jestliže pro každé dva různé vektory $\mathbf{r}_1, \mathbf{r}_2$ platí $C_{ff}(\mathbf{r}_1, \mathbf{r}_2) = 0$. Podobně jsou náhodná pole $\mathbf{f}(\mathbf{r})$ a $\mathbf{g}(\mathbf{r})$ vzájemně nekorelovaná právě tehdy, jestliže pro každé dva vektory $\mathbf{r}_1, \mathbf{r}_2$ platí $C_{fg}(\mathbf{r}_1, \mathbf{r}_2) = 0$.

3.3 Homogenní a ergodické náhodné pole

Kromě toho, že se v praxi omezujeme na sdružené pravděpodobnosti druhého řádu, zavádíme obvykle ještě další zjednodušující předpoklady. Často se předpokládá homogenita a ergodicita náhodného pole. Náhodné pole se nazývá homogenním, jestliže jeho střední hodnota $\mu(\mathbf{r})$ není závislá na \mathbf{r} (pro všechna \mathbf{r} je střední hodnota stejná) a jestliže jeho autokorelace $R_{ff}(\mathbf{r}_1, \mathbf{r}_2)$ závisí pouze na rozdílu $\mathbf{r}_2 - \mathbf{r}_1$. Uvedené podmínky můžeme zapsat pomocí vztahů

$$\mu_f(\mathbf{r}) = \mu_f = \text{konstanta}, \quad (3.11)$$

$$R_{ff}(\mathbf{r}_1, \mathbf{r}_2) = R_{ff}(\mathbf{r}_1 + \mathbf{r}_0, \mathbf{r}_2 + \mathbf{r}_0), \quad (3.12)$$

kde \mathbf{r}_0 je libovolný vektor. Jestliže ve vztahu (3.12) postupně položíme $\mathbf{r}_0 = -\mathbf{r}_2$ a $\mathbf{r}_0 = -\mathbf{r}_1$, získáme

$$R_{ff}(\mathbf{r}_1, \mathbf{r}_2) = R_{ff}(\mathbf{r}_1 - \mathbf{r}_2, 0) = R_{ff}(0, \mathbf{r}_2 - \mathbf{r}_1). \quad (3.13)$$

S ohledem na to, že z definice autokorelace vyplývá $R_{ff}(\mathbf{r}_1, \mathbf{r}_2) = R_{ff}(\mathbf{r}_2, \mathbf{r}_1)$, a zavedeme-li zápis $R_{ff}(\mathbf{r}, 0) \equiv R_{ff}(\mathbf{r})$, pak ze vztahu (3.13) dostáváme $R_{ff}(\mathbf{r}_1 - \mathbf{r}_2) = R_{ff}(\mathbf{r}_2 - \mathbf{r}_1)$. Proto platí

$$R_{ff}(\mathbf{r}) = R_{ff}(-\mathbf{r}). \quad (3.14)$$

Rozepíšeme-li vektor \mathbf{r} jako $\mathbf{r} = (a, b)$, pak pro výpočet autokorelace R_{ff} homogenního pole máme

$$R_{ff}(\mathbf{r}) = R_{ff}(a, b) = E\{\mathbf{f}(x + a, y + b)\mathbf{f}(x, y)\} = E\{\mathbf{f}(x, y)\mathbf{f}(x - a, y - b)\}. \quad (3.15)$$

Poznamenejme ještě, že pojem homogenity náhodného pole je totožný s pojmem stacionarita v širším smyslu, se kterým se můžeme v literatuře také setkat.

Pojem ergodicity souvisí s možností pohlížet na homogenní náhodné pole jako na množinu $\{f_{\omega_i}(\mathbf{r})\}$ funkcí, jak bylo vysvětleno v podkapitole 3.1. Nechť \mathcal{S} je oblast, nad níž náhodné pole (v našem případě obrazový signál) uvažujeme. Míra této oblasti nechť je S . Předpokládejme, že oblast \mathcal{S} je dostatečně velká - ideálně nekonečná. Stanovme střední hodnotu signálu nad oblastí \mathcal{S} . Máme

$$\mathbf{E} = \lim_{S \rightarrow \infty} \frac{1}{S} \iint_{\mathcal{S}} \mathbf{f}(x, y) dx dy . \quad (3.16)$$

Vztahu (3.16) je třeba rozumět tak, že integraci provádíme jednotlivě pro každý elementární jev ω_i . Pro každý elementární jev obdržíme hodnotu $\mathbf{E}(\omega_i)$. \mathbf{E} je zde tedy náhodnou proměnnou (nezaměňovat s operátorem střední hodnoty). Jestliže náhodná proměnná \mathbf{E} nabývá pro všechny elementární jevy stejné hodnoty, pak takové náhodné pole nazveme ergodickým vzhledem ke střední hodnotě. Podobně můžeme pro náhodné pole $\mathbf{f}(x, y)$ určit

$$\mathbf{R}(a, b) = \lim_{S \rightarrow \infty} \frac{1}{S} \iint_{\mathcal{S}} \mathbf{f}(x - a, y - b) \mathbf{f}(x, y) dx dy . \quad (3.17)$$

Stejně jako v předchozím případě, integrujeme i zde pro každý elementární jev jednotlivě a $\mathbf{R}(a, b)$ je proto náhodné pole. Jestliže pro libovolně zvolené a, b nabývá $\mathbf{R}(a, b)$ pro všechny elementární jevy téže hodnoty, pak náhodné pole $\mathbf{f}(x, y)$ nazveme ergodickým vzhledem k autokorelaci.

3.4 Výkonová spektrální hustota

Výkonová spektrální hustota $G_{\text{ff}}(u, v)$ homogenního náhodného pole je Fourierovou transformací jeho autokorelace. Je tedy

$$G_{\text{ff}}(u, v) = \mathcal{F}\{R_{\text{ff}}(a, b)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{\text{ff}}(a, b) \exp[-j2\pi(au + bv)] da db . \quad (3.18)$$

Na základě znalosti výkonové spektrální hustoty lze autokorelaci určit inverzní Fourierovou transformací pomocí vztahu

$$R_{\text{ff}}(a, b) = \mathcal{F}^{-1}\{G_{\text{ff}}(u, v)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G_{\text{ff}}(u, v) \exp[j2\pi(au + bv)] du dv . \quad (3.19)$$

Jestliže ve vztahu (3.19) položíme $a = b = 0$, můžeme dokázat následující vlastnost:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G_{\text{ff}}(u, v) du dv = R_{\text{ff}}(0, 0) = \mathbf{E}\left\{[\mathbf{f}(x, y)]^2\right\} \geq 0 . \quad (3.20)$$

3.5 Lineární operace nad náhodným polem

Necť $\mathbf{f}(x, y)$ je náhodné pole reprezentující obrazový signál. Náhodné pole je charakterizováno střední hodnotou $\mu_{\text{f}}(x, y)$, autokorelací $R_{\text{ff}}(a, b)$ a výkonovou spektrální hustotou $G_{\text{ff}}(u, v)$. Na náhodné pole $\mathbf{f}(x, y)$ aplikujeme operátor \mathcal{O} a získáme tak náhodné pole $\mathbf{g}(x, y)$. Je tedy

$$\mathbf{g}(x, y) = \mathcal{O}\{\mathbf{f}(x, y)\} . \quad (3.21)$$

Naším cílem nyní bude pro takto získané náhodné pole $\mathbf{g}(x, y)$ stanovit statistické veličiny, které jej popisují. Hledáme tedy předpis pro střední hodnotu, autokorelaci a výkonovou spektrální hustotu náhodného pole $\mathbf{g}(x, y)$, a to ve vztahu k odpovídajícím veličinám vstupního náhodného pole $\mathbf{f}(x, y)$. Pro stanovení uvedených charakteristik náhodného pole $\mathbf{g}(x, y)$ použijeme závěrů z odstavce 1.2.4. V tomto odstavci jsme odvodili, že je-li \mathcal{O} lineární operace invariantní vůči posuvu, pak operaci $\mathcal{O}\{\mathbf{f}(x, y)\}$ je možné provést jako konvoluci $\mathbf{f}(x, y) * h(x, y)$, kde $h(x, y) = \mathcal{O}\{\delta(x, y)\}$ je odezva operátoru na Diracův impuls (impulzová charakteristika operátoru). Chápejme podle podkapitoly 3.1 náhodné pole $\mathbf{f}(x, y)$ jako množinu $\{f_{\omega_i}(x, y)\}$ funkcí. Pro výstupní náhodné pole $\mathbf{g}(x, y)$ pak máme

$$\mathbf{g}(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbf{f}(x - a, y - b) h(a, b) da db . \quad (3.22)$$

Výrazu (3.22) je třeba rozumět tak, že pro každý elementární jev ω_i dává funkce $f_{\omega_i}(x, y)$ množiny $\mathbf{f}(x, y)$ vzniknout právě jedné funkci $g_{\omega_i}(x, y)$ množiny $\mathbf{g}(x, y)$. Je tedy $\mathbf{g}(x, y)$ náhodné pole, jak jsme již dříve předeslali. Lze navíc ukázat, že je-li náhodné pole $\mathbf{f}(x, y)$ homogenní, pak je homogenní i pole

$\mathbf{g}(x,y)$. Hledané vztahy pro veličiny charakterizující výstupní náhodné pole uvedeme ve formě tvrzení, která dokážeme. Předpokládáme, že vstupní náhodné pole \mathbf{f} je homogenní.

Pro střední hodnotu μ_g náhodného pole $\mathbf{g}(x,y)$ platí

$$\mu_g = \mu_f H(0,0), \quad (3.23)$$

kde $H(0,0)$ je hodnota Fourierova obrazu $H(u,v)$ funkce $h(x,y)$ v bodě $u = 0, v = 0$.

Důkaz. Tvrzení dokážeme tak, že na obě strany vztahu (3.22) uplatníme operaci střední hodnoty. Postupně dostaneme

$$\begin{aligned} \mu_g &= E \left\{ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbf{f}(x-a, y-b) h(a,b) da db \right\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E \{ \mathbf{f}(x-a, y-b) \} h(a,b) da db \\ &= \mu_f \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(a,b) da db = \mu_f \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(a,b) \exp[-j2\pi(0a+0b)] da db = \mu_f H(0,0). \end{aligned} \quad \bullet$$

Pro autokorelaci R_{gg} náhodného pole $\mathbf{g}(x,y)$ platí

$$R_{gg}(a,b) = R_{ff}(a,b) * h(-a,-b) * h(a,b). \quad (3.24)$$

Důkaz. Jestliže obě strany rovnice (3.22) násobíme $\mathbf{g}(x-\alpha, y-\beta)$ a uplatníme operaci střední hodnoty, získáme

$$R_{gg}(\alpha, \beta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{fg}(\alpha-a, \beta-b) h(a,b) da db = R_{fg}(\alpha, \beta) * h(\alpha, \beta). \quad (3.25)$$

Podobně násobíme obě strany rovnice (3.22) výrazem $\mathbf{f}(x+\alpha, y+\beta)$. Uplatněním operace střední hodnoty a substituce $a' = -a, b' = -b$ pak postupně máme:

$$R_{fg}(\alpha, \beta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{ff}(\alpha+a, \beta+b) h(a,b) da db = R_{ff}(\alpha, \beta) * h(-\alpha, -\beta). \quad (3.26)$$

Dosazením výsledku (3.26) do (3.25) již získáme dokazované tvrzení. •

Pro výkonovou spektrální hustotu náhodného pole $\mathbf{g}(x,y)$ platí

$$G_{gg}(u,v) = G_{ff}(u,v) |H(u,v)|^2. \quad (3.27)$$

Důkaz. Dokazované tvrzení vyplývá z tvrzení předchozího. Stačí provést Fourierovu transformaci rovnice (3.24), využít vlastností (2.27b) Fourierovy transformace a vzít v úvahu, že $H^*(u,v)H(u,v) = |H(u,v)|^2$. •

3.6 Statistická analýza obrazových transformací

V této podkapitole se budeme zabývat transformacemi diskrétního signálu z pohledu statistické analýzy. Předpokládáme, že vstupní signál je popsán náhodným polem $\mathbf{f}(m,n)$, kde $m = 0, 1, \dots, M-1, n = 0, 1, \dots, N-1$. Transformací obdržíme signál, který je popsán náhodným polem $\mathbf{F}(k,l)$, $k = 0, 1, \dots, M-1, l = 0, 1, \dots, N-1$. Ve shodě s výrazem (1.34) máme

$$\mathbf{F}(k, l) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \mathbf{f}(m, n) \phi_{k,l}^*(m, n). \quad (3.28)$$

Pro stručnost přepíšeme uvedený vztah do maticového tvaru (maticový zápis transformace signálu jsme zavedli již v podkapitole 2.2, na rozdíl od uvedené podkapitoly jsou zde však \mathbf{f}, \mathbf{F} náhodné vektory). Dostaneme

$$\mathbf{F} = \Phi \mathbf{f}, \quad (3.29)$$

kde \mathbf{f} , \mathbf{F} jsou sloupcové náhodné vektory rozměru MN a Φ je matice rozměru $MN \times MN$. Úlohou bude určit statistické charakteristiky (střední hodnotu, korelaci a kovarianci) výstupního náhodného pole \mathbf{F} na základě odpovídajících charakteristik vstupního náhodného pole \mathbf{f} . Výraz pro střední hodnotu lze získat uplatněním operace střední hodnoty na obě strany rovnice (3.29). Získáme

$$\boldsymbol{\mu}_F = E\{\mathbf{F}\} = \Phi E\{\mathbf{f}\} = \Phi \boldsymbol{\mu}_f, \quad (3.30)$$

kde $\boldsymbol{\mu}_f$, $\boldsymbol{\mu}_F$ jsou vektory středních hodnot rozměru MN . Pro autokorelaci výstupního náhodného pole dostáváme

$$\mathbf{R}_{FF} = E\{\mathbf{F}\mathbf{F}^{*T}\} = E\{\Phi \mathbf{f} \mathbf{f}^{*T} \Phi^{*T}\} = \Phi \mathbf{R}_{ff} \Phi^{*T}, \quad (3.31)$$

kde \mathbf{R}_{ff} , \mathbf{R}_{FF} jsou autokorelační matice vstupního a výstupního náhodného pole. Obě matice mají rozměr $MN \times MN$. Podobně pro autokovarianci odvodíme

$$\mathbf{C}_{FF} = E\{[\mathbf{F} - \boldsymbol{\mu}_F][\mathbf{F} - \boldsymbol{\mu}_F]^{*T}\} = E\{[\Phi \mathbf{f} - \Phi \boldsymbol{\mu}_f][\Phi \mathbf{f} - \Phi \boldsymbol{\mu}_f]^{*T}\} = \Phi \mathbf{C}_{ff} \Phi^{*T}, \quad (3.32)$$

kde \mathbf{C}_{ff} , \mathbf{C}_{FF} jsou autokovarianční matice vstupního a výstupního náhodného pole. Obě matice mají opět rozměr $MN \times MN$. Poznamenejme nakonec, že má-li transformace nějaké speciální vlastnosti (je-li např. separabilní), pak lze za jistých okolností také pro statistické charakteristiky transformovaného signálu odvodit speciální vztahy, které umožňují efektivnější výpočet než obecné vztahy, které jsme zde právě uvedli. Podrobnější výklad v tomto směru by však pravděpodobně přesáhl skromné možnosti tohoto učebního textu.

Příklad 3.1: Uvažujme homogenní diskrétní náhodné pole $\mathbf{f}(m,n)$. Ukážeme, že pro výkonovou spektrální hustotu platí vztah

$$G_{ff}(k,l) = \frac{1}{\sqrt{MN}} E\left\{|\mathbf{F}(k,l)|^2\right\}.$$

Z podkapitoly 3.4 víme, že výkonová spektrální hustota je Fourierovou transformací autokorelace homogenního náhodného pole. Je tedy

$$G_{ff}(k,l) = \mathcal{F}\{R_{ff}(m,n)\}.$$

Z podmínky homogenity plyne (přestože jsou hodnoty $\mathbf{f}(m,n)$ reálné, použijeme od této chvíle k jejich reprezentaci komplexních čísel - imaginární složka bude však nulová)

$$R_{ff}(m,n) = E\{\mathbf{f}^*(r-m, s-n)\mathbf{f}(r,s)\} = E\left\{\frac{1}{MN} \sum_{r=0}^{M-1} \sum_{s=0}^{N-1} \mathbf{f}^*(r-m, s-n)\mathbf{f}(r,s)\right\}.$$

Na základě linearit y Fourierovy transformace máme

$$\mathcal{F}\{R_{ff}(m,n)\} = E\left\{\mathcal{F}\left[\frac{1}{MN} \sum_{r=0}^{M-1} \sum_{s=0}^{N-1} \mathbf{f}^*(r-m, s-n)\mathbf{f}(r,s)\right]\right\}.$$

Uplatněním korelačního teorému (2.57) vychází

$$G_{ff}(k,l) = \mathcal{F}\{R_{ff}(m,n)\} = \frac{1}{\sqrt{MN}} E\{\mathbf{F}(k,l)\mathbf{F}^*(k,l)\} = \frac{1}{\sqrt{MN}} E\left\{|\mathbf{F}(k,l)|^2\right\}.$$

Jestliže by náhodné pole $\mathbf{f}(m,n)$ bylo navíc ergodické vzhledem k autokorelaci, pak by dále platilo

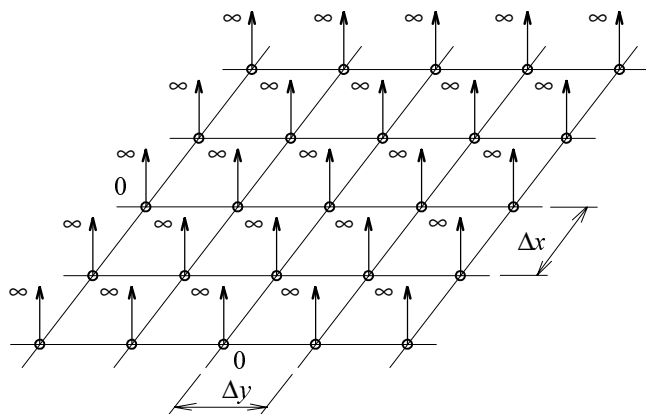
$$G_{ff}(k,l) = \frac{1}{\sqrt{MN}} |F(k,l)|^2. \quad \bullet$$

4 Digitalizace obrazových signálů

Při digitálním zpracování obrazu je dosti často cílem zpracovávat obrazy, které vznikají jako obrazy spojité. Spojitá reprezentace může být žádána také pro výstupy zpracování. Například tehdy, jestliže má být výsledek pozorován člověkem. Pro samotné digitální zpracování obrazu v počítači není však spojité reprezentace vhodná. Vhodnější je reprezentace diskrétní, kdy je obraz reprezentován spočetnou (v praxi konečnou) množinou nějakých hodnot. Z uvedeného důvodu provádíme před digitálním zpracováním převod obrazu do diskrétní podoby. Říkáme, že obraz vzorkujeme. Dále se při praktické reprezentaci obrazů v počítači vzorky někdy uchovávají jako celočíselné hodnoty. V takových případech pak skutečné úrovně signálu na tyto celočíselné hodnoty převádíme, což nazýváme kvantováním. Po zpracování v digitalizované podobě může být naopak požadována rekonstrukce obrazu do podoby spojité. Problematice vzorkování, kvantování a rekonstrukce je věnována tato kapitola.

4.1 Vzorkování deterministicky popsaných obrazů

Předpokládejme, že vzorkujeme vstupní obraz deterministicky popsaný spojitou obrazovou funkcí $f_1(x,y)$. Aby vztahy, ke kterým dospějeme, vyšly jednoduché, předpokládáme, že je obraz definován nad nekonečnou oblastí. Výsledek vzorkování obrazu je popsán funkcí $f_S(x,y)$.



Obr.4.1. Vzorkovací funkce ve tvaru nekonečného pole Diracových impulzů.

Jako matematický model vzorkování obrazu je zvolíme součin vstupního obrazu se vzorkovací funkcí $s(x,y)$. Je tedy

$$f_S(x,y) = f_1(x,y) s(x,y). \quad (4.1)$$

Jako vzorkovací funkci nejprve zvolíme nekonečné pole Diracových impulzů navzájem vzdálených o Δx , Δy (obr. 4.1). Tato funkce je popsána předpisem

$$s(x,y) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \delta(x - k\Delta x, y - l\Delta y). \quad (4.2)$$

Dosazením vztahu (4.2) do vztahu (4.1) obdržíme pro $f_S(x,y)$ následující předpis

$$f_S(x,y) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} f_1(x,y) \delta(x - k\Delta x, y - l\Delta y). \quad (4.3)$$

Je zřejmé, že ke stanovení funkce f_S podle předpisu (4.3) stačí znát funkční hodnoty f_1 pouze v bodech $(k\Delta x, l\Delta y)$. Při praktické realizaci vzorkování jsou proto právě hodnoty $f_1(k\Delta x, l\Delta y)$ uchovávány. Určeme nyní Fourierovu transformaci $F_S(u,v) = \mathcal{F}\{f_S(x,y)\}$ vzorkovaného obrazu $f_S(x,y)$. Necht' $F_1(u,v) = \mathcal{F}\{f_1(x,y)\}$, $S(u,v) = \mathcal{F}\{s(x,y)\}$ jsou Fourierovy transformace funkcí $f_1(x,y)$ a $s(x,y)$. Uplatníme-li na obě strany rovnice (4.1) Fourierovu transformaci a vezmeme-li v úvahu vlastnost (2.28b), pak získáme

$$F_S(u,v) = F_1(u,v) * S(u,v). \quad (4.4)$$

Fourierův obraz nekonečného pole Diracových impulzů jsme vypočetli již dříve v příkladě 2.4. Vyšlo

$$S(u,v) = \frac{1}{\Delta x \Delta y} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \delta\left(u - \frac{k}{\Delta x}, v - \frac{l}{\Delta y}\right). \quad (4.5)$$

Zavedme frekvence u_S , v_S vzorkování jako převrácené hodnoty vzdálenosti vzorků

$$u_S = 1 / \Delta x, \quad v_S = 1 / \Delta y. \quad (4.6)$$

Dosazením vztahů (4.5) a (4.6) do vztahu (4.4) a rozepsáním konvoluce máme

$$F_S(u, v) = u_S v_S \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_1(a, b) \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \delta(u - a - ku_S, v - b - lv_S) da db. \quad (4.7)$$

Vzájemnou záměnou pořadí integrování a sumace a po úpravě obdržíme

$$F_S(u, v) = u_S v_S \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} F_1(u - ku_S, v - lv_S). \quad (4.8)$$

Nyní jsme došli k významnému závěru: Výraz (4.8) říká, že frekvenční spektrum F_S vzorkovaného obrazu se skládá ze složek, kterými je spektrum F_1 obrazu původního. Složky se periodicky opakují ve vzdálenostech u_S, v_S (obr. 4.2) a sčítají. Jak uvidíme později, bude mít pro úspěšnou rekonstrukci obrazu zásadní význam požadavek, aby se jednotlivé složky spektra navzájem nepřekrývaly. Zkoumejme tuto situaci blíže. Předpokládejme, že spektrum vstupního obrazu je omezené - tj., že platí

$$F_1(u, v) = 0, \text{ jestliže } |u| > u_1 \cup |v| > v_1, \quad (4.9)$$

kde hodnoty u_1, v_1 jsou mezní frekvence vstupního obrazu. Z výrazu (4.8) vyplývá (viz. též obr. 4.2), že k překrývání složek spektra nedojde za podmínky, že platí nerovnosti

$$u_S \geq 2 u_1, \quad v_S \geq 2 v_1. \quad (4.10)$$

Nyní se budeme zabývat otázkou, za jakých podmínek lze ze vzorkovaného obrazu $f_S(x, y)$ vytvořit rekonstruovaný obraz $f_R(x, y)$ tak, aby byl shodný se vstupním obrazem $f_1(x, y)$. Předpokládáme, že rekonstrukci obrazu bude možné provést pomocí nějakého filtru, který proto budeme nazývat filtrem rekonstrukčním. Očekáváme, že rekonstrukční filtr bude lineární a invariantní vůči posunu. Filtr charakterizujeme odezvou $r(x, y)$ na Diracův impulz (impulzová charakteristika filtru) a Fourierovou transformací $R(u, v)$ této odezvy (frekvenční charakteristika). Tyto funkce však zatím neznáme a chceme je určit.

Z odstavce 1.2.4 víme, že obraz na výstupu rekonstrukčního filtru lze stanovit pomocí konvoluce vzorkovaného obrazu a odezvy filtru na Diracův impulz. Je tedy

$$f_R(x, y) = f_S(x, y) * r(x, y). \quad (4.11)$$

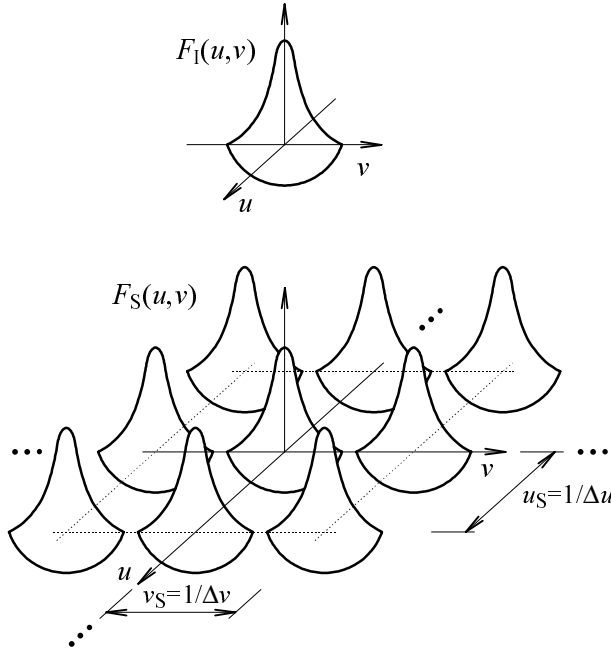
Ve frekvenční doméně pak s ohledem na vlastnost (2.28a) Fourierovy transformace máme

$$F_R(u, v) = F_S(u, v) R(u, v). \quad (4.12)$$

Dosadíme-li do vztahu (4.12) ze vztahu (4.8), dostaneme

$$F_R(u, v) = u_S v_S R(u, v) \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} F_1(u - ku_S, v - lv_S). \quad (4.13)$$

Předpokládejme, že nedojde k překrývání jednotlivých složek spektra, jak jsme již dříve diskutovali (vztahy (4.10)). Ze vztahu (4.13) pak vyplývá, že jestliže navrhne rekonstrukční filtr tak, aby odfiltroval ty složky spektra, pro které $k \neq 0, l \neq 0$ a složku $k = 0, l = 0$ naopak propustil, pak $F_R(u, v) = F_1(u, v)$ a tedy i $f_R(x, y) = f_1(x, y)$, jak jsme pro rekonstrukci požadovali. Docházíme tedy k závěru, že k optimální rekonstrukci by frekvenční charakteristika rekonstrukčního filtru mohla být např. (obr. 4.3)

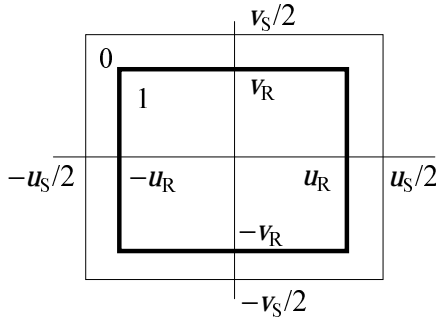


Obr.4.2. Spektrum původního a vzorkovaného obrazu.

$$R(u, v) = \begin{cases} 1, & (|u| \leq u_R) \cap (|v| \leq v_R) \\ 0, & \text{jinak} \end{cases}, \quad (4.14)$$

kde $u_1 \leq u_R \leq u_S/2$, $v_1 \leq v_R \leq v_S/2$. (Přesně vzato dává filtr (4.14) hodnotu $F_R(u, v) = u_S v_S F_1(u, v)$, což však nezmenšuje kvalitu rekonstrukce, protože $u_S v_S$ je známá konstanta.) Zdůrazněme ještě jednou skutečnost, že podmínkou úspěšné rekonstrukce obrazu je splnění požadavku, aby se jednotlivé složky spektra $F_S(u, v)$ vzorkovaného obrazu navzájem nepřekrývaly. Jestliže mezní frekvence u_1, v_1 vstupního obrazu považujeme za neměnné, pak lze splnění této podmínky zajistit vhodnou volbou frekvence vzorkování u_S, v_S tak, aby byly splněny nerovnosti (4.10). Uvedené nerovnosti jsou dvojrozměrnou variantou známého Nyquistova vzorkovacího teorému, který říká, že k tomu, aby mohl být signál úspěšně rekonstruován, je nutné, aby vzorkovací frekvence byla alespoň dvojnásobkem nejvyšší frekvence, která je v signálu obsažena.

Při provádění rekonstrukce obrazu může být výhodné vyhnout se Fourierově transformaci, která může být časově složitá. Vztah (4.11) napovídá, jak to udělat. Dosadíme-li do vztahu (4.11) vztah (4.3), pak po provedení konvoluce a úpravě obdržíme předpis



$$f_R(x, y) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} f_1(k\Delta x, l\Delta y) r(x - k\Delta x, y - l\Delta y). \quad (4.15)$$

Abychom mohli vztahu (4.15) prakticky využít, zbývá ještě určit impulzovou charakteristiku $r(x, y)$ rekonstrukčního filtru (4.14). Zpětnou Fourierovou transformací získáme ze vztahu (4.14) předpis

$$r(x, y) = 4u_R v_R \operatorname{sinc}(2xu_R) \operatorname{sinc}(2yv_R). \quad (4.16)$$

Obr.4.3. Frekvenční charakteristika obdélníkového rekonstrukčního filtru.

Vidíme tedy, že v prostorové doméně lze rekonstrukci obrazu provést podle předpisu (4.15) jako konvoluci vzorků $f_1(k\Delta x, l\Delta y)$ vstupního obrazu s funkcí sinc.

4.2 Zpřesněné modely vzorkování

Zjednodušující předpoklady, které jsme zavedli v podkapitole 4.1, směřovaly k tomu, abychom získali jednoduché a názorné výsledky. Přestože jsou takto získané výsledky užitečné, považujeme za nutné alespoň naznačit možné modifikace uvedeného postupu, jejichž cílem je věrnější popis skutečnosti.

Konečný obraz a konečný počet vzorků: V předchozí podkapitole jsme předpokládali, že obraz, který vzorkujeme, je definován nad nekonečnou oblastí a že vzorkováním získáme nekonečné množství vzorků. V praxi však pracujeme s obrazy, které jsou konečné a konečný musí být také počet vzorků. Zlepšení souhlasu teoretického modelu vzorkování se skutečností by v tomto ohledu bylo možné dosáhnout zavedením vzorkovací funkce ve tvaru konečného pole Diracových impulzů

$$s(x, y) = \sum_{k=-K}^K \sum_{l=-L}^L \delta(x - k\Delta x, y - l\Delta x), \quad (4.17)$$

kde K, L jsou meze pole. Tento postup však bohužel komplikuje další analýzu.

Obecný tvar vzorkovacího pulzu: Až doposud jsme předpokládali, že matematickým modelem získání vzorků je součin obrazu s vzorkovací funkcí. Jako vzorkovací funkci jsme používali funkci složenou z Diracových impulzů. Obecně však mohou mít vzorkovací pulzy i jiný tvar. Jestliže je tvar pulzu popsán funkcí $p(x, y)$, pak můžeme vzorkovací funkci popsat vztahem

$$s(x, y) = \sum_{k=-K}^K \sum_{l=-L}^L p(x - k\Delta x, y - l\Delta x). \quad (4.18)$$

Vzorkování integrací: Senzory, které lze prakticky realizovat, nedokáží sejmout hodnotu v nekonečně malém bodě. Místo toho je sejmuta jakási průměrná hodnota podél citlivé plošky senzoru. Tuto skutečnost lze zohlednit tak, že jako matematický model získávání vzorků volíme integraci

$$f_S(k\Delta x, l\Delta y) = \int_{k\Delta x - \alpha}^{k\Delta x + \alpha} \int_{l\Delta y - \beta}^{l\Delta y + \beta} f_1(x, y) p(x - k\Delta x, y - l\Delta y) dx dy, \quad (4.19)$$

kde α, β jsou poloviny délek stran snímací plošky senzoru, kterou si zde představujeme jako obdélník. Rozměry snímací plošky při tom předpokládáme menší, než je teoretická vzdálenost vzorkovacích bodů. Funkce $p(x, y)$ popisuje průběh citlivosti plošky. Vzorkování integrací má značný význam. To proto, že ať chceme či nechceme, vždy k němu v praxi do určité míry dochází. Ukažme proto další možnou interpretaci vztahu (4.19). Záměnou proměnných převedeme výraz (4.19) na tvar

$$f_S(k\Delta x, l\Delta y) = \int_{-\alpha}^{\alpha} \int_{-\beta}^{\beta} f_1(k\Delta x - a, l\Delta y - b) p(-a, -b) da db. \quad (4.20)$$

Vztah (4.20) je výpočtem konvoluce funkcí $f_1(x, y)$, $p(-x, -y)$. Výsledek konvoluce použijeme jen v bodech $k\Delta x$, $l\Delta y$ (mimo tyto body předpokládáme funkční hodnotu f_S rovnu 0). Můžeme proto psát

$$f_S(x, y) = [f_1(x, y) * p(-x, -y)] \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \delta(x - k\Delta x, y - l\Delta y). \quad (4.21)$$

Výraz (4.21) ukazuje, že vzorkování integrací má stejný účinek, jako kdybychom provedli konvoluci vstupního obrazu $f_1(x, y)$ se vzorkovací funkcí $p(-x, -y)$ a pak následně provedli vzorkování polem Diracových impulzů. Jak vyplývá z vlastnosti (2.28a) Fourierovy transformace, realizuje konvoluce obrazu se vzorkovací funkcí filtraci obrazu. Ta může být někdy nežádoucí, protože způsobuje degradaci obrazu. Jindy může být naopak vítaná - např. tehdy, jestliže vzorkování obrazu probíhá na nižší frekvenci, než odpovídá Nyquistovu teorému, a je proto nutné potlačit v obraze vyšší frekvenční složky.

4.3 Aliasing

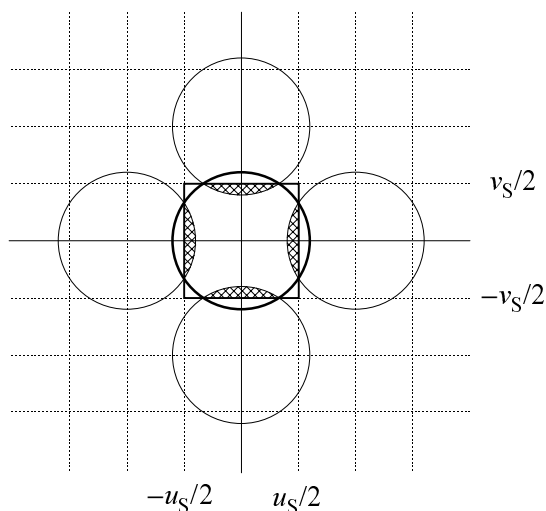
Ve vztahu (4.8) jsme ukázali, že frekvenční spektrum obrazu po vzorkování se skládá ze složek, které jsou tvořeny frekvenčním spektrem vstupního obrazu. Tyto složky se periodicky opakují ve vzdálenostech u_S, v_S a sčítají. Zjistili jsme také, že k tomu, aby mohl být obraz úspěšně rekonstruován, je nutné, aby se jednotlivé složky spektra navzájem nepřekrývaly. Dosud jsme však neřešili otázku, co se stane v případě, že tato podmínka není splněna. Věnujme se nyní tomuto problému podrobněji. Na základě vztahu (4.8) je možné spektrum obrazu po vzorkování zapsat ve tvaru

$$F_S(u, v) = u_S v_S [F_1(u, v) + F_Q(u, v)], \quad (4.22)$$

kde $F_1(u, v)$ je spektrum původního obrazu a $F_Q(u, v)$ zahrnuje všechny ty složky spektra $F_S(u, v)$ obrazu po vzorkování, pro které je ve výrazu (4.8) $k \neq 0$, $l \neq 0$ (složka $k = 0$, $l = 0$ je zastoupena členem $F_1(u, v)$). Předpokládejme nyní, že se jednotlivé složky spektra navzájem překrývají (obr. 4.4) a že na takto vzorkovaný obraz aplikujeme rekonstrukční filtr s frekvenční charakteristikou

$$R(u, v) = \begin{cases} 1, & (|u| \leq u_S / 2) \cap (|v| \leq v_S / 2) \\ 0, & \text{jinak} \end{cases}. \quad (4.23)$$

Zvolený rekonstrukční filtr propouští kromě složky $k = 0$, $l = 0$ (užitečná složka) frekvenčního spektra také určité části složek $k \neq 0$, $l \neq 0$ (nežádoucí složky). Na obr. 4.4 jsou přenášeny části nežádoucích složek vyšrafovány. Naopak si však také můžeme povšimnout, že nejsou přenášeny některé části užitečné složky. Chyba způsobená neúplným přenesením užitečné složky způsobí rozmazání obrazu (potlačení vyšších frekvencí). Naopak chyba způsobená přenosem částí nežádoucích složek dává v obraze vzniknout novým útvarům. Tento jev nazýváme aliasing. Necht' $f_{ID}(x, y)$ je obraz degradovaný neúplným přenosem užitečné složky. Vznik aliasingu můžeme matematicky modelovat tak, že na ob-



Obr. 4.4. Vznik aliasingu překrýváním složek spektra vzorkovaného obrazu.

raz $f_{ID}(x,y)$ superponujeme chybový obraz $a(x,y)$ vzniklý částečným přenosem nežádoucích složek. Je tedy

$$f_R(x,y) = f_{ID}(x,y) + a(x,y). \quad (4.24)$$

Chybový obraz $a(x,y)$ můžeme popsat vztahem

$$a(x,y) = \mathcal{F}^{-1}\{R(u,v)F_Q(u,v)\}. \quad (4.25)$$

Pokusme se nyní kvantifikovat míru degradace rekonstruovaného obrazu, kterou aliasing způsobí. Ke kvantifikaci použijeme stochastického přístupu. Nechť $f_I(x,y)$ je vstupní náhodné pole a $G_{f_I f_I}(x,y)$ nechť je jeho výkonová spektrální hustota. Na míru degradace rekonstruovaného obrazu budeme usuzovat na základě energetické bilance. Energie E_I vstupního obrazového pole je

$$E_I = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G_{f_I f_I}(u,v) du dv. \quad (4.26)$$

Z této energie je rekonstrukčním filtrem na výstup přenášeno množství E_{IR} , kde

$$E_{IR} = \int_{-u_S/2}^{u_S/2} \int_{-v_S/2}^{v_S/2} G_{f_I f_I}(u,v) du dv. \quad (4.27)$$

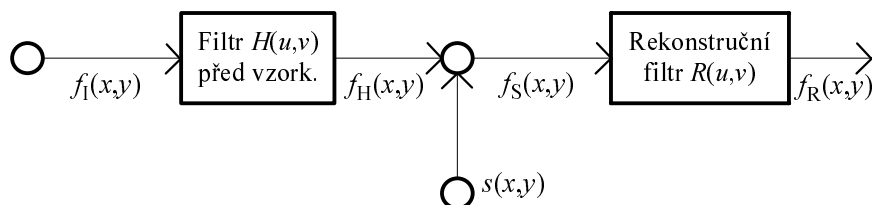
Rekonstrukčním filtrem je na výstup ovšem také částečně přenášena energie nežádoucích složek. K určení velikosti této energie bychom měli vyšetřovat, jak do oblasti $\langle -u_S/2, u_S/2 \rangle \times \langle -v_S/2, v_S/2 \rangle$ každá z těchto složek přispívá. Není těžké zdůvodnit, že energie nežádoucích složek přenášena v oblasti $\langle -u_S/2, u_S/2 \rangle \times \langle -v_S/2, v_S/2 \rangle$ je stejná jako energie, kterou přenáší užitečná složka mimo tuto oblast (prosíme čtenáře, aby platnost tohoto tvrzení promyslel). Energie E_A nežádoucích složek přenášena rekonstrukčním filtrem na výstup tedy je

$$E_A = E_I - E_{IR}. \quad (4.28)$$

Relativní chybu ϵ_A způsobenou aliasingem lze pak vyjádřit podílem

$$\epsilon_A = E_A / E_I. \quad (4.29)$$

K omezení jevů způsobených aliasingem můžeme použít filtrace, kterou provedeme ještě před vzorkováním. Cílem filtrace je potlačení vysokých frekvencí v obraze, čímž budou lépe zajištěny požadavky Nyquistova teoremu. Obr. 4.5 znázorňuje matematický model takového postupu: Před vzorkováním je zařazen filtr s frekvenční charakteristikou $H(u,v)$. Rekonstrukce je provedena rekonstrukčním filtrem dle vztahu (4.23). Zařazením filtru $H(u,v)$ do řetězce zpracování hrozí ovšem další degradace obrazu. Je tedy nutné zvolit takový filtr, kde omezení jevů aliasingu převažuje nad touto dodatečnou degradací.



Obr. 4.5. Matematický model redukce aliasingu.

Nechť $f_H(x,y)$ je náhodné pole vzniklé aplikací filtru $H(u,v)$ na vstupní náhodné obrazové pole $f_I(x,y)$ a pole $f_S(x,y)$ nechť je pole vzniklé vzorkováním pole $f_H(x,y)$. Označme E_H celkovou energii pole

\mathbf{f}_H , a dále označme E_{HR} energii té části užitečné složky pole \mathbf{f}_S , která je propuštěna filtrem $R(u,v)$ na výstup. S využitím vztahu (3.27) máme

$$E_H = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G_{\mathbf{f}_H} (u, v) |H(u, v)|^2 dudv, \quad E_{HR} = \int_{-u_S/2}^{u_S/2} \int_{-v_S/2}^{v_S/2} G_{\mathbf{f}_H} (u, v) |H(u, v)|^2 dudv. \quad (4.30, 31)$$

Chybu ε_A způsobenou aliasingem a chybu ε_H způsobenou zařazením filtru $H(u,v)$ lze vyjádřit vztahy

$$\varepsilon_A = (E_H - E_{HR}) / E_H, \quad \varepsilon_H = (E_{IR} - E_{HR}) / E_{IR}. \quad (4.32)$$

4.4 Praktické metody rekonstrukce obrazu

Úkolem rekonstrukce je získat hodnoty jasu v bodech ležících mimo rastr, v němž jsou k dispozici hodnoty vzorků. Uvedený problém vzniká např. při vykreslování obrazu nebo při provádění geometrických operací s obrazy (kapitola 6). V podkapitole 4.1 jsme ukázali, že rekonstrukci obrazu po vzorkování lze v prostorové doméně provést konvolucí vzorků obrazu s funkcí sinc (vztahy (4.15), (4.16)). Protože Fourierovým obrazem funkce sinc je ideální rekonstrukční filtr dle vztahu (4.14), je tento postup za zvolených předpokladů optimální z hlediska přesnosti rekonstrukce. Funkce sinc je však bohužel pro provádění konvoluce poněkud nepříjemná. Její hodnoty jsou téměř všude nenulové, a proto do každého bodu rekonstruovaného obrazu teoreticky přispívají všechny vzorky. Tato skutečnost vede k vysoké časové složitosti rekonstrukce. Navíc je přesné splnění předpokladů, za nichž byl ideální rekonstrukční filtr odvozen, prakticky sotva možné. V praxi se proto často používají filtry, které sice nejsou ideálními rekonstrukčními filtry, ale zato umožňují rychlý výpočet konvoluce. Rychlosti je dosahováno tím, že je impulzová charakteristika filtru nenulová jen na malé oblasti, takže do každého bodu rekonstruovaného obrazu přispívá pouze omezený počet vzorků. Frekvenční charakteristika takových filtrů se však zpravidla dosti odchyluje od charakteristiky rekonstrukčního filtru ideálního, což teoreticky vede k nižší kvalitě rekonstrukce. Uvedme několik příkladů impulzních a odpovídajících frekvenčních charakteristik častěji používaných rekonstrukčních filtrů:

1) sinc:
$$r(x, y) = u_S v_S \text{sinc}(u_S x) \text{sinc}(v_S y), \quad (4.33a)$$

$$R(u, v) = \begin{cases} 1, & (|u| \leq u_S / 2) \cap (|v| \leq v_S / 2); \\ 0, & \text{jinak} \end{cases}; \quad (4.33b)$$

2) čtverec:
$$r_0(x, y) = \begin{cases} 1, & (-\Delta x / 2 < x \leq \Delta x / 2) \cap (-\Delta y / 2 < y \leq \Delta y / 2); \\ 0, & \text{jinak} \end{cases}, \quad (4.34a)$$

$$R_0(u, v) = \Delta x \Delta y \text{sinc}(u \Delta x) \text{sinc}(v \Delta y); \quad (4.34b)$$

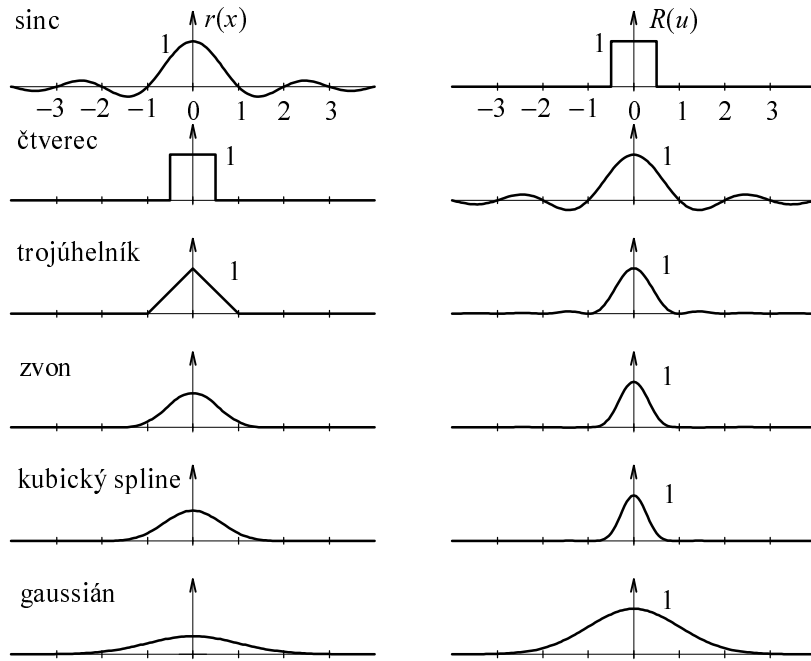
3) trojúhelník:
$$r_1(x, y) = r_0(x, y) * r_0(x, y), \quad R_1(u, v) = R_0^2(u, v); \quad (4.35a,b)$$

4) zvon:
$$r_2(x, y) = r_0(x, y) * r_1(x, y), \quad R_2(u, v) = R_0^3(u, v); \quad (4.36a,b)$$

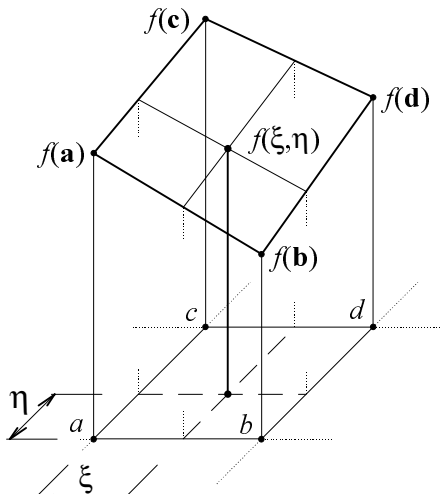
5) kubický B-spline:
$$r_3(x, y) = r_0(x, y) * r_2(x, y), \quad R_3(u, v) = R_0^4(u, v); \quad (4.37a,b)$$

6) gausián:
$$r(x, y) = \frac{1}{2\pi\sigma^2} \exp\left[-\frac{x^2 + y^2}{2\sigma^2}\right], \quad R(u, v) = \exp\left[-\frac{\sigma^2(u^2 + v^2)}{2}\right]. \quad (4.38a,b)$$

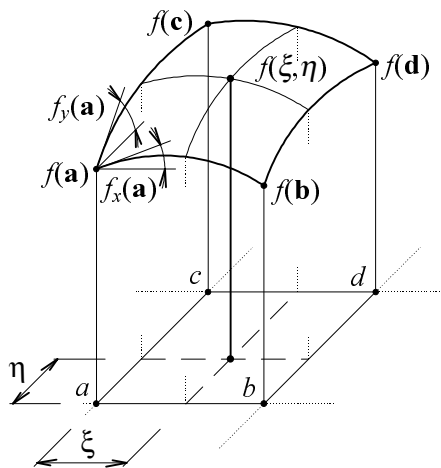
Průběhy uvedených funkcí (v jednorozměrné variantě) jsou uvedeny na obr. 4.6. Poznamenejme, že kvantifikaci věrnosti rekonstrukce, které lze zvoleným filtrem dosáhnout, je možné provést podobně jako v odstavci 4.3 energetickou bilancí. K hodnocení může sloužit množství energie, které zvolený rekonstrukční filtr přenáší v pásmu, kde filtr ideální má frekvenční charakteristiku rovnu nule, a také rozdíl v množství energií přenášených zvoleným a ideálním filtrem v pásmu, kde filtr ideální má frekvenční charakteristiku rovnu jedné.



Obr. 4.6. Průběhy častěji používaných rekonstrukčních filtrů ($\Delta x = \Delta y = 1$).



Obr. 4.7. Bilineární interpolace jasu.



Obr. 4.8. Bikubická interpolace jasu.

V praxi se rekonstrukce dosti často provádí interpolací. Nejčastěji se používá interpolace bilineární nebo bikubické. Řekněme, že hodnota jasové funkce má být stanovena v bodě o souřadnicích x, y a že $\Delta x = \Delta y = 1$. Zaveďme souřadnice ξ, η ($0 \leq \xi, \eta < 1$) pomocí vztahů

$$\xi = x - \lfloor x \rfloor, \quad \eta = y - \lfloor y \rfloor. \quad (4.39)$$

Dále pro stručnost zápisu položme

$$\begin{aligned} \mathbf{a} &= (\lfloor x \rfloor, \lfloor y \rfloor), & \mathbf{b} &= (\lfloor x \rfloor + 1, \lfloor y \rfloor), \\ \mathbf{c} &= (\lfloor x \rfloor, \lfloor y \rfloor + 1), & \mathbf{d} &= (\lfloor x \rfloor + 1, \lfloor y \rfloor + 1). \end{aligned} \quad (4.40)$$

Nechť symboly a, b, c, d označují body, které jsou reprezentovány vektory $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$. Bilineární interpolace je popsána následujícím vztahem (obr. 4.7)

$$\begin{aligned} f(x, y) &= f(\lfloor x \rfloor + \xi, \lfloor y \rfloor + \eta) \\ &= [(1-\xi), \xi] \begin{bmatrix} f(\mathbf{a}) & f(\mathbf{c}) \\ f(\mathbf{b}) & f(\mathbf{d}) \end{bmatrix} \begin{bmatrix} (1-\eta) \\ \eta \end{bmatrix} \\ &= (1-\xi)(1-\eta)f(\mathbf{a}) + \xi(1-\eta)f(\mathbf{b}) + (1-\xi)\eta f(\mathbf{c}) + \xi\eta f(\mathbf{d}). \end{aligned} \quad (4.41)$$

Poznamenejme, že vztah (4.41) dává pro hodnotu $f(x, y)$ tentýž výsledek jako konvoluce s filtrem $r_1(x, y)$ dle vztahu (4.35a). Vztah (4.41) má ovšem velmi názornou interpretaci a také v programech jej lze velmi snadno realizovat.

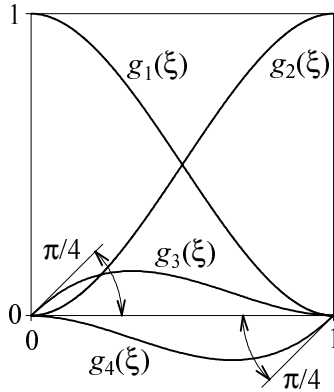
Při vyšších nárocích na kvalitu interpolace lze požit interpolace bikubické (obr. 4.8). Výpočet lze provést pomocí vztahu

$$f(\lfloor x \rfloor + \xi, \lfloor y \rfloor + \eta) = [g_1(\xi), g_2(\xi), g_3(\xi), g_4(\xi)] \begin{bmatrix} f(\mathbf{a}) & f(\mathbf{c}) & f_y(\mathbf{a}) & f_y(\mathbf{c}) \\ f(\mathbf{b}) & f(\mathbf{d}) & f_y(\mathbf{b}) & f_y(\mathbf{d}) \\ f_x(\mathbf{a}) & f_x(\mathbf{c}) & 0 & 0 \\ f_x(\mathbf{b}) & f_x(\mathbf{d}) & 0 & 0 \end{bmatrix} \begin{bmatrix} g_1(\eta) \\ g_2(\eta) \\ g_3(\eta) \\ g_4(\eta) \end{bmatrix}. \quad (4.42)$$

Funkce g_1, g_2, g_3, g_4 jsou definovány následujícími vztahy (obr. 4.9)

$$g_1(\xi) = 1 - 3\xi^2 + 2\xi^3, \quad g_2(\xi) = 3\xi^2 - 2\xi^3, \quad (4.43)$$

$$g_3(\xi) = \xi - 2\xi^2 + \xi^3, \quad g_4(\xi) = -\xi^2 + \xi^3.$$



Obr. 4.9. Funkce $g_1(\xi)$, $g_2(\xi)$, $g_3(\xi)$, $g_4(\xi)$.

Hodnoty f_x, f_y jsou hodnoty derivací (podle x resp. podle y) jasové funkce v odpovídajícím bodě. Hodnoty derivací nahradíme diferencemi. Např.

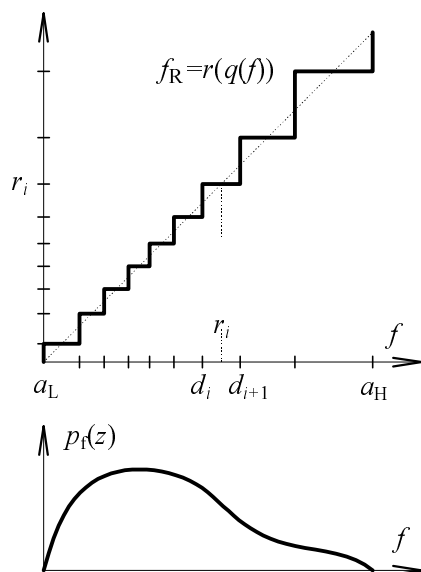
$$f_x(\mathbf{a}) = \frac{f(\lfloor x \rfloor + 1, \lfloor y \rfloor) - f(\lfloor x \rfloor - 1, \lfloor y \rfloor)}{2}, \quad (4.44)$$

$$f_y(\mathbf{a}) = \frac{f(\lfloor x \rfloor, \lfloor y \rfloor + 1) - f(\lfloor x \rfloor, \lfloor y \rfloor - 1)}{2}.$$

Na krajích obrazu, kde nejsou hodnoty požadované vztahem (4.44) k dispozici, lze použít nesymetrických diferencních vzorců.

4.5 Kvantování obrazových signálů

Při praktické reprezentaci signálů v počítači je často potřebné uchovávat jednotlivé vzorky signálu v proměnných s omezeným rozsahem (byte, slovo atd.). V takových případech provádíme převod skutečných úrovní signálu na hodnoty, které lze v proměnné zvoleného typu uchovávat, což nazýváme kvantováním. Na konci řetězce digitálního zpracování je naopak často potřebné obraz na základě jeho kvantované reprezentace opět zrekonstruovat tak, aby co nejlépe odpovídal obrazu teoretickému. V tomto odstavci formulujeme uvedenou úlohu poněkud podrobněji a naznačíme také její možné řešení.



Obr. 4.10. Kvantování jasu.

Zvolme v obraze nějaký libovolný bod. Nechť f je hodnota jasu v tomto bodě. Předpokládejme, že f vždy nabývá reálných hodnot $a_L \leq f < a_H$, kde a_L, a_H značí známou dolní a horní mez jasu. Proměnné, které používáme k uchovávání vzorků, nechť umožňují uložení celých čísel v rozsahu 0 až $I-1$. Přejít od reálné hodnoty f k celočíselné hodnotě $i \in \{0, 1, \dots, I-1\}$ lze uskutečnit podle předpisu $i = q(f)$, který nazýváme kvantováním. Problémem je nalézt optimální kvantovací funkci q . Kvantovací funkce bývá obvykle zkonstruována následujícím postupem: Zvolme $I-1$ hodnot $d_i, i=1, 2, \dots, I-1$, které interval $\langle a_L, a_H \rangle$ rozdělí na I podintervalů. Předpokládejme, že platí $d_i < d_{i+1}$ a zavedme dále označení $a_L = d_0, a_H = d_I$. Kvantovací funkce q zobrazuje všechny hodnoty jasu z intervalu $\langle d_i, d_{i+1} \rangle$ na celé číslo i . Problém nalezení kvantovací funkce se v tomto případě redukuje na nalezení $I-1$ hodnot d_1, d_2, \dots, d_{I-1} . Označme f_R hodnotu jasu po rekonstrukci. K rekonstrukci použijeme funkce $f_R = r(i)$. Rekonstrukční funkce r může být v praxi snadno realizována tabulkou. Funkce r zobrazuje celočíselnou hodnotu i na reálnou hodnotu r_i . V úloze o kvantování hledáme pro daný rozsah I

a pro jistou třídu předpokládaných obrazových signálů hodnoty d_i a r_i tak, aby rekonstruované hodnoty jasu v jistém smyslu co nejlépe odpovídaly jasu původnímu (obr. 4.10).

K řešení úlohy, kterou jsme právě formulovali, můžeme použít stochastického přístupu. Považujme vzorek jasu za náhodnou proměnnou s hustotou pravděpodobnosti $p_f(z)$ (obr. 4.10, připomeňme, že hodnota $p_f(z)dz$ udává pravděpodobnost jevu, že jas padne do intervalu $(z, z+dz)$). Označme \mathbf{f} , \mathbf{f}_R náhodné proměnné popisující velikost původního a rekonstruovaného jasu. Hodnoty d_i , r_i nalezneme tak, aby byla minimalizována kvantizační chyba $\varepsilon = E\{(\mathbf{f}_R - \mathbf{f})^2\}$. Uvedený vztah vychází z představy, že pro jasy, které se v obraze vyskytují málo, lze tolerovat větší diferenci $f_R - f$ než pro jasy, které se vyskytují často. Použitím operace střední hodnoty je zajištěno, že se difference $f_R - f$ na celkové chybě podílí úměrně hustotě pravděpodobnosti jasu p_f . Úpravami postupně dostáváme

$$\varepsilon = E\{[\mathbf{f}_R - \mathbf{f}]^2\} = E\left\{[r(q(\mathbf{f})) - \mathbf{f}]^2\right\} = \int_{a_L}^{a_H} [r(q(z)) - z]^2 p_f(z) dz = \sum_{i=0}^{I-1} \int_{d_i}^{d_{i+1}} (r_i - z)^2 p_f(z) dz. \quad (4.45)$$

Jestliže je počet I úrovní dostatečně velký, můžeme často hustotu pravděpodobnosti $p_f(z)$ uvnitř jednoho kvantovacího intervalu (d_i, d_{i+1}) považovat za konstantní, a to $p_f(r_i)$, protože zajisté bude $d_i \leq r_i \leq d_{i+1}$. Ze vztahu (4.45) pak dostáváme

$$\varepsilon = \sum_{i=0}^{I-1} p_f(r_i) \int_{d_i}^{d_{i+1}} (r_i - z)^2 dz = \frac{1}{3} \sum_{i=0}^{I-1} p_f(r_i) [(d_{i+1} - r_i)^3 - (d_i - r_i)^3]. \quad (4.46)$$

Z podmínky minima $\frac{\partial \varepsilon}{\partial r_i} = 0$ stanovíme $r_i = \frac{d_{i+1} + d_i}{2}$. (4.47)

Dosazením vztahu (4.47) do (4.46) vyjde

$$\varepsilon = \frac{1}{12} \sum_{i=0}^{I-1} p_f(r_i) (d_{i+1} - d_i)^3. \quad (4.48)$$

Optimální hodnoty úrovní d_i lze nalézt minimalizací hodnoty ε . Řešení této optimalizační úlohy je poněkud nepříjemné. V (Pratt 78) se jako přibližné řešení uvádí vztah

$$d_i = \frac{(a_H - a_L) \int_{a_L}^{a_i} [p_f(z)]^{-1/3} dz}{\int_{a_L}^{a_H} [p_f(z)]^{-1/3} dz} + a_L, \quad \text{kde} \quad a_i = \frac{i(a_H - a_L)}{I} + a_L. \quad (4.49)$$

Ze vztahu (4.49) je zřejmé, že jestliže je rozdělení pravděpodobnosti rovnoměrné, vychází také rozhodovací úrovně d_i rozmístěny pravidelně.

Komplikace nastávají v případě, kdy je počet kvantovacích úrovní malý, takže hustotu pravděpodobnosti na kvantovacím intervalu nelze považovat za konstantní. V takovém případě není možné provést zjednodušení použité ve výrazu (4.46), ale je nutné pracovat s výrazem (4.45). Za povšimnutí stojí také problém kvantování v barevných obrazech. Jak víme, je barva často reprezentována trojicí R, G, B intenzit základních barevných složek. Vlivem kvantování intenzity barevných složek může u barevných obrazů dojít k nepřesnostem nejen v intenzitě, ale i v barevném odstínu.

5 Filtrace obrazových signálů

Filtrace je jednou ze základních metod úpravy signálů (včetně signálů reprezentujících obrazy) spočívající v úpravě kmitočtového spektra. V principu lze filtraci provádět dvěma postupy: nerekurzivně a rekurzivně. Uvažujme prostor diskretních signálů. Nerekurzivní filtry vytvářejí výstupní signál v každém bodě jako lineární kombinaci vzorků signálu vstupního. Impulzová charakteristika nerekurzivních filtrů obsahuje vždy jen konečný počet nenulových hodnot, protože jinak by je nebylo možné prakticky realizovat. Nerekurzivní filtry jsou proto také označovány jako filtry FIR (finite-extent impulse response). Rekurzivní filtry naproti tomu předepisují vztah mezi vstupním a výstupním signálem obecněji: výstupní signál v každém bodě je lineární kombinací vzorků nejen signálu vstupního, ale také právě určovaného signálu výstupního. Jako rekurzivní lze realizovat také filtry, jejichž impulzová charakteristika obsahuje nekonečný počet nenulových hodnot. Rekurzivní filtry proto také bývají označovány jako filtry IIR (infinite-extent impulse response). Problematice nerekurzivní i rekurzivní filtrace je věnována tato kapitola. Kromě teoretických základů uvedeme také některé praktické metody návrhu filtrů, včetně metod využívajících stochastického přístupu.

5.1 Nerekurzivní filtrace

Nechť $f(m,n)$ je diskretní vstupní signál a $h(m,n)$ nechť je impulzová charakteristika filtru. Nerekurzivní filtraci rozumíme úpravu signálu podle předpisu

$$g(m,n) = f(m,n) * h(m,n). \quad (5.1)$$

Provedeme-li Fourierovu transformaci rovnice (5.1), získáváme vztah

$$G(k,l) = F(k,l)H(k,l). \quad (5.2)$$

Z rovnice (5.2) je patrné, že filtrace provádí úpravu frekvenčního spektra signálu. Tato představa bývá ostatně dosti často primární. Často požadujeme, aby filtr modifikoval frekvenční spektrum signálu jistým předepsaným způsobem, tj. aby měl jistou požadovanou frekvenční charakteristiku.

5.1.1 Realizace nerekurzivního filtru v prostorové a frekvenční doméně

V praxi se výpočet filtrace provádí jak v doméně prostorové podle vztahu (5.1), tak v doméně frekvenční podle vztahu (5.2). Jedním z faktorů, které ovlivňují volbu postupu, je časová složitost výpočtu. Uvedeme příklad časové rozvahy. Předpokládejme, že máme provést filtraci dle vztahu (5.1) a že funkce f, h jsou nenulové nad čtvercovými oblastmi o rozměrech $M \times M$ a $R \times R$ bodů. Předpokládejme, že vstupní signál f i impulzová charakteristika h filtru jsou reálné, což bývá v praxi často splněno. Časovou složitost měříme počtem reálných násobení. Protože při praktické realizaci konvoluce máme vypočítat M^2 hodnot a protože výpočet každé z nich vyžaduje R^2 násobení, dostáváme pro časovou složitost C_S filtrace realizované v prostorové doméně vztah

$$C_S = M^2 R^2. \quad (5.3)$$

Při provádění filtrace ve frekvenční doméně je třeba provést Fourierovu transformaci vstupního signálu, pak součin na pravé straně vztahu (5.2) a konečně zpětnou Fourierovu transformaci. Předpokládáme, že frekvenční charakteristiku $H(k,l)$ použitého filtru známe předem a že ji není nutné počítat. Protože přechod od vztahu (5.1) ke vztahu (5.2) předpokládá použití cyklické konvoluce, převedeme výpočet konvoluce lineární (v praktických aplikacích se konvoluce ve vztahu (5.1) dosti často uvažuje lineární) na výpočet konvoluce cyklické tak, jak jsme již dříve popsali v odstavci 1.2.5. Zavedme označení $Q=M+R-1$. Z odstavce 2.3.5 víme, že časová složitost rychlé Fourierovy transformace je $3Q^2 \log_2 Q / 4$ komplexních násobení - tj. $3Q^2 \log_2 Q$ násobení reálných (jedno násobení komplexní vyžaduje čtyři násobení reálná). Celková časová složitost C_F filtrace realizované ve frekvenční doméně měřená počtem reálných násobení tedy je

$$C_F = Q^2 (3 \log_2 Q + 4 + 3 \log_2 Q) = Q^2 (4 + 6 \log_2 Q). \quad (5.4)$$

Porovnáním vztahů (5.3) a (5.4) zjišťujeme, že provedení filtrace ve frekvenční doméně je výhodnější, jestliže platí

$$R > \sqrt{4 + 6 \log_2 Q}. \quad (5.5)$$

Např. pro $Q=1024$ vychází, že filtrace ve frekvenční doméně je při uvedených kriteriích výhodnější pro $R > 8$.

5.1.2 Filtry s nulovou fází

Úpravu frekvenčního spektra intuitivně chápeme tak, že se filrací mění amplituda jednotlivých frekvenčních složek signálu, nikoli však jejich fáze. Předpokládáme tedy, že frekvenční charakteristika filtru je reálná (obecněji můžeme na takový filtr pohlížet jako na filtr, jehož frekvenční charakteristika má nulové imaginární složky, a proto jej nazýváme filtrem s nulovou fází). Oprávněnost uvedeného požadavku je stvrzována praktickými zkušenostmi - při zpracování obrazu filtry, jejichž frekvenční charakteristika má nenulové imaginární složky, lze subjektivně zaznamenat nežádoucí degradace obrazu. Podmínku nulové fáze filtru ve frekvenční doméně můžeme zapsat pomocí vztahu

$$H(k, l) = H^*(k, l). \quad (5.6)$$

S využitím předpisu (2.39) pro Fourierovou transformaci odtud snadno získáváme podmínku

$$h(m, n) = h^*(-m, -n). \quad (5.7)$$

Vidíme tedy, že podmínka nulové fáze frekvenční charakteristiky filtru vyžaduje, aby impulzová charakteristika filtru byla centrovaná kolem počátku podle vztahu (5.7). Při filtraci obrazů je někdy možné volit impulzovou charakteristiku filtru na základě intuice. I v takovém případě lze ovšem z výše popsaných důvodů doporučit respektování vztahu (5.7).

5.1.3 Návrh filtru s využitím výřezové funkce

Často chceme v prostorové doméně realizovat filtr, jehož požadovanou impulzovou charakteristiku (označme ji např. $i(m, n)$) sice známe, avšak její praktické použití je nevýhodné, protože je nenulová nad příliš rozsáhlou oblastí, což by vedlo k vysoké časové složitosti výpočtu. V těchto případech lze vytvořit aproximaci $h(m, n)$ ideální impulzové charakteristiky $i(m, n)$ jejím zkrácením pomocí výřezové (okenní) funkce $w(m, n)$ podle předpisu

$$h(m, n) = i(m, n)w(m, n). \quad (5.8)$$

Návrh filtru nyní spočívá ve volbě vhodného typu výřezové funkce w a ve volbě velikosti výřezu tak, aby se vlastnosti filtru dostatečně shodovaly s vlastnostmi požadovaného filtru ideálního. Při volbě výřezové funkce se zpravidla požaduje splnění následujících vlastností:

$$w(m, n) = w^*(-m, -n), \quad (5.9)$$

$$w(m, n) = w_1(m)w_2(n). \quad (5.10)$$

Místo separability dle vztahu (5.10) může být někdy požadována funkce rotačně symetrická, což je zajištěno předpisem

$$w(m, n) = w_0\left(\sqrt{m^2 + n^2}\right). \quad (5.11)$$

Splnění uvedených vlastností většinou usnadňuje realizaci následného výpočtu. Splnění vlastnosti (5.9) je předpokladem pro vytvoření filtru s nulovou fází. Separabilita funkce podle (5.10) ulehčuje např. výpočet Fourierovy transformace. Je totiž

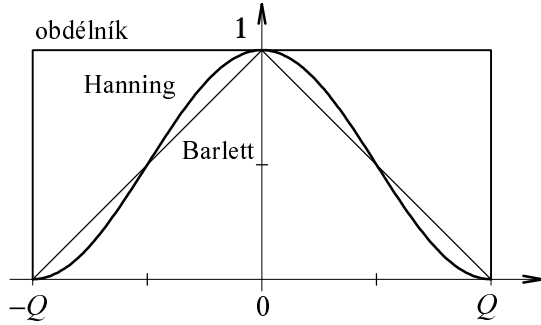
$$\mathcal{F}\{w(m, n)\} = \mathcal{F}\{w_1(m)w_2(n)\} = \mathcal{F}\{w_1(m)\}\mathcal{F}\{w_2(n)\} = W_1(k)W_2(l). \quad (5.12)$$

V praxi se používá mnoha typů výřezových funkcí. Zde uvedeme pouze některé z nich (dále uvedené funkce mohou být použity na místě funkcí w_1, w_2, w_0 , Q označuje velikost výřezu):

obdélník:
$$w(n) = \begin{cases} 1, & |n| < Q \\ 0, & \text{jinak} \end{cases}, \quad (5.13)$$

Barlettova funkce:
$$w(n) = \begin{cases} 1 - \frac{|n|}{Q}, & |n| < Q, \\ 0, & \text{jinak} \end{cases}, \quad (5.14)$$

Hanningova funkce:
$$w(n) = \begin{cases} \frac{1}{2} \left[1 + \cos\left(\pi \frac{n}{Q}\right) \right], & |n| < Q, \\ 0, & \text{jinak} \end{cases}. \quad (5.15)$$



Obr. 5.1. Průběhy některých výřezových funkcí.

Průběhy uvedených funkcí jsou znázorněny na obr. 5.1. Průběh frekvenční charakteristiky $\mathcal{F}\{i(m,n)w(m,n)\}$ lze použít pro posouzení kvality filtru. Uplatníme-li na obě strany rovnice (5.8) Fourierovu transformaci, pak pro frekvenční charakteristiku získáme vztah

$$H(k,l) = I(k,l) * W(k,l), \quad (5.16)$$

kde I je frekvenční charakteristika ideálního filtru a W je Fourierův obraz výřezové funkce.

5.1.4 Návrh optimálního filtru

V tomto odstavci se budeme věnovat praktické metodě návrhu impulzové charakteristiky $h(m,n)$ filtru. Uvažujme obraz rozměru $M \times N$. Budeme požadovat, aby hodnoty $h(m,n)$ byly nenulové jen nad zvolenou a relativně malou oblastí (označme její rozměry $R \times S$), což je nezbytné z hlediska dosažení přijatelné časové složitosti výpočtu v prostorové doméně. Dále předpokládáme, že známe požadovanou frekvenční charakteristiku filtru. Protože očekáváme, že této frekvenční charakteristiky nebude možné dosáhnout přesně (s ohledem na omezení rozměrů R, S), budeme tuto frekvenční charakteristiku nazývat charakteristikou ideální. Necht' I je ideální a H skutečná frekvenční charakteristika filtru (obě charakteristiky předpokládáme reálné). Nalezený filtr bude tím lepší, čím menší bude chyba

$$\varepsilon = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} W_{k,l} [H(k,l) - I(k,l)]^2, \quad (5.17)$$

kde $W_{k,l}$ jsou váhové konstanty zohledňující význam dílčí chyby pro jednotlivé frekvenční složky. Tyto konstanty volíme. Vždy však musí být $W_{k,l} > 0$. Za optimální považujeme takový filtr, který minimalizuje hodnotu chyby ε . Rozvedme nyní naznačený postup poněkud podrobněji. Fourierův obraz impulzové charakteristiky filtru (tj. skutečnou frekvenční charakteristiku) zapíšeme ve tvaru

$$H(k,l) = \sum_{r=0}^{R-1} \sum_{s=0}^{S-1} h(r,s) \Phi_{k,l}(r,s), \quad \text{kde } \Phi_{k,l}(r,s) = \exp\left[-j2\pi\left(\frac{rk}{M} + \frac{sl}{N}\right)\right]. \quad (5.18, 19)$$

Dosažením vztahu (5.18) do (5.17) dostáváme

$$\varepsilon = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} W_{k,l} \left[\sum_{r=0}^{R-1} \sum_{s=0}^{S-1} h(r,s) \Phi_{k,l}(r,s) - I(k,l) \right]^2. \quad (5.20)$$

Pro nalezení extrému položíme:

$$\frac{\partial \varepsilon}{\partial h(m,n)} = 0 \quad (m = 0, 1, \dots, R-1, \quad n = 0, 1, \dots, S-1). \quad (5.21)$$

Po úpravách zjišťujeme, že hledané hodnoty $h(r,s)$ jsou řešením soustavy lineárních rovnic

$$\sum_{r=0}^{R-1} \sum_{s=0}^{S-1} A_{m,n}(r,s) h(r,s) = B_{m,n}, \quad (5.22)$$

kde
$$A_{m,n}(r,s) = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} W_{k,l} \Phi_{k,l}(r,s) \Phi_{k,l}(m,n), \quad B_{m,n} = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} W_{k,l} I(k,l) \Phi_{k,l}(m,n).$$

5.2 Rekurzivní filtrace

V případě rekurzivní filtrace je výstupní signál g svázán se vstupním signálem f prostřednictvím vztahu

$$b(m,n)*g(m,n) = a(m,n)*f(m,n), \quad (5.23)$$

kde b, a jsou diskrétní funkce, které rekurzivní filtr popisují. Necht' je hodnota $b(0,0)$ nenulová. Beze ztráty na obecnosti předpokládáme, že $b(0,0)=1$. Rozepsáním konvoluce ze vztahu (5.23) snadno ověříme, že platí

$$g(m,n) = a(m,n)*f(m,n) + c(m,n)*g(m,n), \quad (5.24)$$

kde funkce c je definována takto

$$c(m,n) = \begin{cases} 0, & m=0, n=0 \\ -b(m,n), & \text{jinak} \end{cases}. \quad (5.25)$$

Fourierovou transformací rovnice (5.23) obdržíme vztah

$$G(k,l) = \frac{A(k,l)}{B(k,l)} F(k,l), \quad (5.26)$$

kde A, B, F, G jsou Fourierovy obrazy funkcí a, b, f, g . Vidíme, že frekvenční charakteristika rekurzivního filtru je $H(k,l) = A(k,l) / B(k,l)$.

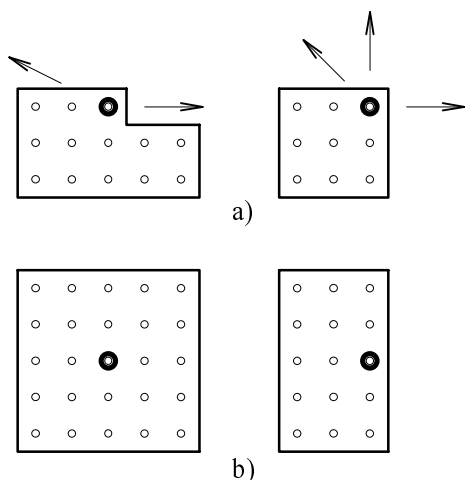
Motivací pro použití rekurzivní filtrace je zpravidla to, že rekurzivní filtr může mít nižší časovou složitost než nerekurzivní filtr obdobných vlastností. Necht' N_a, N_b jsou rozměry oblastí, nad nimiž funkce f, a, b nabývají nenulových hodnot (pro jednoduchost předpokládáme, že všechny oblasti mají čtvercový tvar). Z výrazu (5.24) je zřejmé, že časová složitost rekurzivní filtrace při přímém výpočtu je (detailněji je postup výpočtu popsán v následující podkapitole)

$$C_R = N^2(N_a^2 + N_b^2). \quad (5.27)$$

Rekurzivní filtr je tedy výhodný tehdy, jestliže se funkce a, b podaří navrhnout tak, aby rozměry N_a, N_b byly malé (menší než rozměr oblasti, nad níž nabývá nenulových hodnot funkce h nerekurzivního filtru srovnatelných vlastností). Poznamenejme, že hodnoty N_a, N_b mohou být u rekurzivního filtru konečné i tehdy, jestliže je impulzová charakteristika filtru nenulová nad nekonečným počtem bodů.

5.2.1 Realizace rekurzivního filtru přímým výpočtem

Ze vztahu (5.24) vidíme, že k tomu, aby bylo možné provést výpočet hodnoty $g(m,n)$ výstupního pole, je nutné dříve znát ty hodnoty pole g , pro které jsou odpovídající hodnoty funkce c nenulové. Výpočet



Obr.5.2. a) Funkce $c(m,n)$ umožňující přímý výpočet, b) neumožňující přímý výpočet.

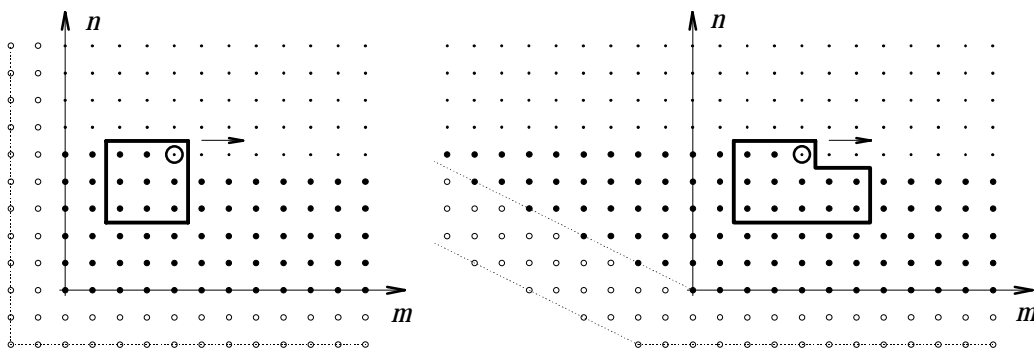
je proto žádoucí zorganizovat tak, aby hodnoty g byly získávány v potřebném pořadí. Je ovšem zřejmé, že pro některé funkce c nebude existovat žádné pořadí, které by přímý výpočet umožnilo (později však uvidíme že, takové filtry mohou být vyčíslitelné iteračním postupem). Na obr.5.2 jsou ilustrovány oba případy. Je zde vyznačen tvar oblastí, nad níž funkce c nabývá nenulových hodnot, a také možné směry postupu při výpočtu. Plným kruhem je vyznačeno místo, pro které je hodnota $g(m,n)$ výstupního pole v daném okamžiku určována.

Pro „nastartování“ výpočtu rekurzivního filtru je třeba mít předem k dispozici některé hodnoty výstupního obrazu g . Tyto hodnoty mají význam okrajových podmínek. Požadujeme-li filtr lineární, pak je nutné jako okrajové podmínky zadávat pouze hodnoty 0. Toto tvrzení vyplývá z jednoduché úvahy: Necht' \emptyset je uvažova-

ný rekurzivní filtr. Předpokládejme, že filtr \mathcal{O} je lineární. Platí proto

$$\mathcal{O}\{af(m,n)\} = a\mathcal{O}\{f(m,n)\}. \quad (5.28)$$

Položíme-li v (5.28) $a=0$, konstatujeme na základě uvedeného vztahu, že v případě lineárního filtru nulovému vstupnímu obrazu nutně odpovídá také nulový obraz výstupní. Při volbě nenulových okrajových podmínek by zde obecně došlo ke sporu: nulovému vstupnímu obrazu by odpovídal nenulový obraz výstupní, a proto by filtr \mathcal{O} nebyl lineární. Ani poloha bodů, ve kterých zadáváme okrajové podmínky, nemůže být volena libovolně. Okrajové podmínky, lze zadávat pouze v těch místech, kde hodnotu výstupního pole neurčuje výraz (5.24). Obr. 5.3 ukazuje příklady míst, v nichž se zadávají okrajové podmínky, pro různé tvary oblastí, nad níž je funkce c nenulová.



Obr.5.3. Příklady okrajových podmínek pro různé funkce $c(m,n)$.

Na rozdíl od nerekurzivních filtrů, které jsou vždy stabilní, je u filtrů rekurzivních nutno sledovat problémy stability. Jestliže je filtr nestabilní, pak se v průběhu výpočtu mohou šířit zaokrouhlovací chyby a případné šумы vstupního pole a dosahovat vysokých hodnot. Tento problém byl v minulosti detailně studován, ale jeho podrobnější diskuse by přesáhla rámec tohoto textu. V případě zájmu proto čtenáře odkazujeme např. na práci (Dudgeon 84).

5.2.2 Realizace rekurzivního filtru iteračním výpočtem

V předchozím odstavci jsme vysvětlili, že má-li oblast, nad níž funkce b, c nabývají nenulové hodnoty, vhodný tvar, pak lze rekurzivní filtraci realizovat přímým výpočtem. Ne všechny filtry jsou ovšem takto realizovatelné. Přímý výpočet nelze uplatnit tehdy, jestliže by pro výpočet hodnoty $g(m,n)$ bylo zapotřebí hodnot v těch místech výstupního pole, kde zatím nebyly vypočteny. V těchto případech lze však použít iteračního postupu. Na základě vztahu (5.24) se pro výpočet $g(m,n)$ nabízí iterační předpis

$$g_i(m,n) = a(m,n)*f(m,n) + c(m,n)*g_{i-1}(m,n), \quad (5.29)$$

kde $g_i(m,n)$ znamená hodnotu v i -tém iteračním kroku. Pro zahájení iteračního pochodu volíme $g_0(m,n) = 0$. Oprávněnost iteračního postupu popsaného vztahem (5.29) je ovšem samozřejmě nutné doložit studiem konvergence. Ačkoli předpokládáme, že samotný výpočet filtrace bude prováděn v prostorové doméně, analýzu konvergence provedeme v doméně frekvenční. Fourierovou transformací vztahu (5.24) obdržíme

$$G(k,l) = A(k,l)F(k,l) + C(k,l)G(k,l), \quad (5.30)$$

kde C je Fourierovou transformací funkce c . Z (5.30) dostaneme dále vztah

$$G(k,l) = \frac{A(k,l)}{1-C(k,l)}F(k,l). \quad (5.31)$$

Fourierovou transformací rovnice (5.29) získáme

$$G_i(k,l) = A(k,l)F(k,l) + C(k,l)G_{i-1}(k,l). \quad (5.32)$$

Odtud máme
$$G_1(k,l) = A(k,l)F(k,l), \quad G_2(k,l) = A(k,l)F(k,l)[1 + C(k,l)], \quad (5.33)$$

$$G_n(k, l) = A(k, l)F(k, l) \sum_{i=0}^{n-1} C^i(k, l).$$

S využitím identity

$$\sum_{i=0}^{n-1} C^i(k, l) = \frac{1 - C^n(k, l)}{1 - C(k, l)}, \quad (5.34)$$

za předpokladu že $|C(k, l)| < 1$ a s uplatněním vztahu (5.31) docházíme v případě nekonečného počtu iteračních kroků k následujícímu závěru:

$$\lim_{n \rightarrow \infty} G_n(k, l) = A(k, l)F(k, l) \lim_{n \rightarrow \infty} \frac{1 - C^n(k, l)}{1 - C(k, l)} = \frac{A(k, l)}{1 - C(k, l)} F(k, l) = G(k, l). \quad (5.35)$$

Vidíme tedy, že jestliže platí $|C(k, l)| < 1$, pak iterační proces konverguje k přesnému řešení. Poznamenejme však, že uvedená podmínka je zbytečně přísná. Podrobnější analýzou bylo totiž dokázáno, že iteračním postupem lze realizovat každý stabilní filtr (Dudgeon 84).

5.2.3 Návrh rekurzivního filtru

Návrhem rekurzivního filtru rozumíme stanovení funkcí $a(m, n)$, $b(m, n)$ tak, aby měl filtr požadované vlastnosti. Výchozími údaji pro návrh může být např. nějaký vstupní signál $f(m, n)$ a požadovaná odezva $d(m, n)$ filtru na tento signál (příkladem je realizace filtru se zadanou impulzovou charakteristikou). Naznačme postup takového návrhu. Aby časová složitost filtrace byla co nejmenší, požadujeme, aby funkce a, b měly nenulové funkční hodnoty jen v malém počtu bodů. Nechť $g(m, n)$ je skutečná odezva filtru na funkci $f(m, n)$. S ohledem na předchozí požadavek se bude skutečná odezva lišit od odezvy požadované. Úhrnná kvadratická chyba filtru v prostorové doméně pak je

$$\varepsilon = \sum_m \sum_n [g(m, n) - d(m, n)]^2. \quad (5.36)$$

Funkce $a(m, n)$, $b(m, n)$ stanovíme minimalizací funkcionálu (5.36). Analogický postup můžeme realizovat také ve frekvenční doméně. Úhrnná kvadratická chyba filtru ve frekvenční doméně je

$$\varepsilon = \sum_k \sum_l [G(k, l) - D(k, l)]^2 = \sum_k \sum_l \left[\frac{A(k, l)}{B(k, l)} F(k, l) - D(k, l) \right]^2. \quad (5.37)$$

Můžeme také zavést funkci $W_{k, l}$ vyjadřující váhu chyby pro jednotlivé frekvence. Máme pak

$$\varepsilon_w = \sum_k \sum_l W_{k, l} \left[\frac{A(k, l)}{B(k, l)} F(k, l) - D(k, l) \right]^2. \quad (5.38)$$

Poznamenejme na závěr, že ať již použijeme návrhu v doméně prostorové nebo frekvenční, jsou rezultující úlohy minimalizace nelineární, a proto je zde zpravidla zapotřebí použít numerického řešení.

5.3 Obecný model degradace a rekonstrukce obrazu

Při snímání obrazu často dochází k jeho různým poškozením (degradacím). Obraz může být např. zašumělý, neostrý atd. Pokud je to možné, bývá obvykle žádoucí tyto degradace korigovat. Korekce bývá dosti často realizována filtrací. Klíčovou otázkou je volba vhodného filtru. K jeho nalezení je nejprve vhodné zavést model degradace obrazu. Označme f_I vstupní nepoškozený obraz. Obraz f_I není ovšem k dispozici. Dostupný je místo něj pouze obraz poškozený, který označíme f_D . Představujeme si, že degradace je způsobena nějakým operátorem (označme jej \mathcal{D}) a dále si představujeme, že dochází k degradaci aditivním šumem (označme jej v). Je tedy

$$f_D = \mathcal{D}\{f_I\} + v. \quad (5.39)$$

Úkolem rekonstrukce obrazu je nalézt rekonstrukční operátor \mathcal{R} tak, abychom jeho aplikací na obraz f_D obdrželi obraz (označme jej f_R), který je co možná nejbližší obrazu f_I . Máme tedy

$$f_R = \mathcal{R}\{f_D\} = \mathcal{R}\{\mathcal{D}\{f_I\} + v\}. \quad (5.40)$$

Ze všech možných operátorů zvolíme za operátor \mathcal{R} ten, který dává minimální hodnotu chyby

$$\varepsilon = \|f_{\mathcal{R}} - f_{\mathcal{I}}\|. \quad (5.41)$$

Předpokládejme nyní, že degradační operátor \mathcal{D} je lineární a invariantní vůči posuvu. Jeho impulzovou a frekvenční charakteristiku označme $h_{\mathcal{D}}, H_{\mathcal{D}}$. Také rekonstrukční operátor \mathcal{R} předpokládáme lineární a invariantní vůči posuvu. Úkolem je nalézt impulzovou charakteristiku $h_{\mathcal{R}}$ nebo frekvenční charakteristiku $H_{\mathcal{R}}$ rekonstrukčního operátoru \mathcal{R} tak, aby byla minimalizována chyba ze vztahu (5.41). Zapišme rovnice (5.39), (5.40) pomocí konvoluce. Máme

$$f_{\mathcal{D}}(m, n) = f_{\mathcal{I}}(m, n) * h_{\mathcal{D}}(m, n) + v(m, n), \quad (5.42)$$

$$f_{\mathcal{R}}(m, n) = [f_{\mathcal{I}}(m, n) * h_{\mathcal{D}}(m, n) + v(m, n)] * h_{\mathcal{R}}(m, n). \quad (5.43)$$

Provedeme-li Fourierovu transformaci uvedených vztahů, dostaneme

$$F_{\mathcal{D}}(k, l) = F_{\mathcal{I}}(k, l)H_{\mathcal{D}}(k, l) + V(k, l), \quad (5.44)$$

$$F_{\mathcal{R}}(k, l) = [F_{\mathcal{I}}(k, l)H_{\mathcal{D}}(k, l) + V(k, l)]H_{\mathcal{R}}(k, l). \quad (5.45)$$

5.4 Inverzní filtrace

O inverzním filtru hovoříme tehdy, jestliže frekvenční charakteristiku rekonstrukčního operátoru \mathcal{R} zvolíme ve tvaru

$$H_{\mathcal{R}}(k, l) = \frac{1}{H_{\mathcal{D}}(k, l)}. \quad (5.46)$$

Dosazením vztahu (5.46) do (5.45) zjišťujeme, že platí

$$F_{\mathcal{R}}(k, l) = F_{\mathcal{I}}(k, l) + \frac{V(k, l)}{H_{\mathcal{D}}(k, l)}. \quad (5.47)$$

Inverzní Fourierovou transformací pak obdržíme

$$f_{\mathcal{R}}(m, n) = f_{\mathcal{I}}(m, n) + \mathcal{F}^{-1} \left\{ \frac{V(k, l)}{H_{\mathcal{D}}(k, l)} \right\}. \quad (5.48)$$

Ze vztahu (5.48) vidíme, že pokud by obraz $f_{\mathcal{D}}$ nebyl degradován šumem, pak by inverzní filtr poskytl obraz, který je shodný s nepoškozeným vstupním obrazem $f_{\mathcal{I}}$. Za přítomnosti šumu se ovšem obrazy $f_{\mathcal{R}}$ a $f_{\mathcal{I}}$ liší. Odchylka je popsána druhým členem na pravé straně rovnice (5.48). Je zřejmé, že vliv šumu roste s tím, jak klesají hodnoty $H_{\mathcal{D}}(k, l)$. Protože u většiny degradací jsou hodnoty $H_{\mathcal{D}}(k, l)$ odpovídající vyšším frekvenčním složkám, kde vliv šumu zejména očekáváme, malé (obraz je obvykle rozostřován tím, že jsou potlačeny vyšší frekvence), dochází použitím inverzního filtru často k nežádoucímu zdůraznění šumu. Jiným praktickým problémem je, že frekvenční charakteristiku degradačního operátoru mnohdy neznáme přesně. Často je nutné spokojit se s pouhým odhadem.

5.5 Wienerův filtr ve frekvenční doméně

V této podkapitole využijeme k nalezení rekonstrukčního filtru stochastického přístupu. Ukážeme nejprve odvození tzv. Wienerova filtru ve frekvenční doméně (v prostorové doméně jej odvodíme v následující podkapitole). Předpokládáme, že $\mathbf{f}_{\mathcal{I}}(m, n)$, $\mathbf{f}_{\mathcal{D}}(m, n)$, $\mathbf{f}_{\mathcal{R}}(m, n)$ jsou homogenní náhodná pole rozměru $M \times N$ reprezentující postupně ideální, degradovaný a rekonstruovaný obraz a že $\mathbf{v}(m, n)$ je náhodné pole rozměru $M \times N$, které reprezentuje aditivní šum. Také odpovídající Fourierovy obrazy $\mathbf{F}_{\mathcal{I}}(m, n)$, $\mathbf{F}_{\mathcal{D}}(m, n)$, $\mathbf{F}_{\mathcal{R}}(m, n)$, $\mathbf{V}(m, n)$ jsou náhodná pole rozměru $M \times N$. Pro zestručnění položíme

$$\mathbf{f}_{\Delta}(m, n) = \mathbf{f}_{\mathcal{R}}(m, n) - \mathbf{f}_{\mathcal{I}}(m, n), \quad (5.49)$$

$$\mathbf{F}_{\Delta}(k, l) = \mathcal{F}\{\mathbf{f}_{\Delta}(m, n)\} = \mathbf{F}_{\mathcal{R}}(k, l) - \mathbf{F}_{\mathcal{I}}(k, l). \quad (5.50)$$

Rekonstrukční filtr nalezneme minimalizací chyby

$$\varepsilon = E \left\{ \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |\mathbf{f}_\Delta(m, n)|^2 \right\}. \quad (5.51)$$

Podle Parsevalova teorému (2.59) ovšem platí

$$\varepsilon = E \left\{ \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |\mathbf{f}_\Delta(m, n)|^2 \right\} = E \left\{ \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} |\mathbf{F}_\Delta(k, l)|^2 \right\}. \quad (5.52)$$

$\mathbf{F}_\Delta(k, l)$ vyjádříme dosazením vztahu (5.45) do vztahu (5.50). Dostaneme

$$\mathbf{F}_\Delta(k, l) = [\mathbf{F}_I(k, l)H_D(k, l) + \mathbf{V}(k, l)]H_R(k, l) - \mathbf{F}_I(k, l). \quad (5.53)$$

Uvážíme-li, že $|\mathbf{F}_\Delta(k, l)|^2 = \mathbf{F}_\Delta(k, l)\mathbf{F}_\Delta^*(k, l)$, pak ze vztahu (5.52) máme

$$\varepsilon = E \left\{ \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} \mathbf{F}_\Delta(k, l)\mathbf{F}_\Delta^*(k, l) \right\}. \quad (5.54)$$

Frekvenční charakteristiku H_R rekonstrukčního filtru určíme z podmínky

$$\frac{\partial \varepsilon}{\partial H_R(r, s)} = 0, \quad r = 0, 1, \dots, M-1, \quad s = 0, 1, \dots, N-1. \quad (5.55)$$

Ze vztahu (5.54) máme

$$\begin{aligned} \frac{\partial \varepsilon}{\partial H_R(r, s)} &= E \left\{ \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} \left[\frac{\partial \mathbf{F}_\Delta(k, l)}{\partial H_R(r, s)} \mathbf{F}_\Delta^*(k, l) + \mathbf{F}_\Delta(k, l) \frac{\partial \mathbf{F}_\Delta^*(k, l)}{\partial H_R(r, s)} \right] \right\} \\ &= E \left\{ \frac{\partial \mathbf{F}_\Delta(r, s)}{\partial H_R(r, s)} \mathbf{F}_\Delta^*(r, s) + \mathbf{F}_\Delta(r, s) \frac{\partial \mathbf{F}_\Delta^*(r, s)}{\partial H_R(r, s)} \right\}. \end{aligned} \quad (5.56)$$

Předpokládejme, že náhodná pole \mathbf{f}_I, \mathbf{v} jsou nekorelovaná. Uplatněním výsledku z podkapitoly 3.6 lze snadno ukázat, že pole $\mathbf{F}_I(k, l)$ a $\mathbf{V}(k, l)$ jsou proto také nekorelovaná, z čehož vyplývá $E\{\mathbf{F}_I(k, l)\mathbf{V}(k, l)\} = E\{\mathbf{F}_I(k, l)\}E\{\mathbf{V}(k, l)\}$. Dále předpokládejme, že $E\{\mathbf{v}(m, n)\} = \mathbf{0}$. Odtud plyne, že $E\{\mathbf{V}(k, l)\} = \mathbf{0}$. Dosadíme-li do vztahu (5.56) vztah (5.53), položíme-li $E\{\mathbf{F}_I(k, l)\mathbf{F}_I^*(k, l)\} / \sqrt{MN} = G_{f_I f_I}(k, l)$, $E\{\mathbf{V}(k, l)\mathbf{V}^*(k, l)\} / \sqrt{MN} = G_{vv}(k, l)$, pak po úpravách dostaneme pro rekonstrukční filtr předpis

$$H_R(r, s) = H_D^*(r, s) \frac{G_{f_I f_I}(r, s)}{|H_D(r, s)|^2 G_{f_I f_I}(r, s) + G_{vv}(r, s)}. \quad (5.57)$$

Poznamenejme, že $G_{f_I f_I}$ je výkonová spektrální hustota vstupního nedegradovaného signálu a G_{vv} je výkonová spektrální hustota šumu (viz. výsledek příkladu 3.1).

5.6 Wienerův filtr v prostorové doméně

V této podkapitole ukážeme odvození Wienerova filtru v prostorové doméně. Opět předpokládáme, že $\mathbf{f}_I, \mathbf{f}_D, \mathbf{f}_R$ jsou homogenní náhodná pole reprezentující ideální, degradovaný a rekonstruovaný obraz a že \mathbf{v} je náhodné pole popisující aditivní šum. Stejně jako v předchozí podkapitole, je i zde cílem nalézt filtr, který na základě znalosti degradovaného obrazu provede jeho rekonstrukci, která je v jistém smyslu optimální. Abychom mohli použít přehledného maticového zápisu, reprezentujeme všechna uvedená pole pomocí sloupcových vektorů rozměru MN . Předpokládáme, že odhad bude možné vyjádřit ve tvaru

$$\mathbf{f}_R = \mathbf{W}\mathbf{f}_D + \mathbf{b}, \quad (5.58)$$

kde \mathbf{W} je zatím neznámá matice rozměru $MN \times MN$ a \mathbf{b} je zatím neznámý vektor rozměru MN , které musíme stanovit tak, aby byl odhad optimální. Požadujeme, aby střední hodnota odhadu byla stejná jako střední hodnota ideálního obrazu, tedy aby platilo

$$E\{\mathbf{f}_R\} = E\{\mathbf{f}_I\}. \quad (5.59)$$

Dále požadujeme, aby odhad \mathbf{f}_R minimalizoval chybu

$$\boldsymbol{\varepsilon} = E\left\{(\mathbf{f}_R - \mathbf{f}_I)^T (\mathbf{f}_R - \mathbf{f}_I)\right\}. \quad (5.60)$$

Ekvivalentním postupem k minimalizaci chyby ze vztahu (5.60) je aplikace principu ortogonalit. Podle tohoto principu můžeme hledanou matici \mathbf{W} a vektor \mathbf{b} nalézt pomocí následující podmínky

$$E\left\{(\mathbf{f}_R - \mathbf{f}_I)(\mathbf{f}_D - E\{\mathbf{f}_D\})^T\right\} = \mathbf{0}. \quad (5.61)$$

Protože jsme princip ortogonalit uvedli bez důkazu, pokusme se alespoň o jeho poněkud méně formální, avšak názornou interpretaci: Jestliže jsou chyba odhadu (rozdíl v první kulaté závorce) a odchylka pozorování od střední hodnoty (rozdíl v druhé kulaté závorce) nekorelované veličiny, pak již použitými prostředky (statistické momenty druhého řádu) nelze nic zlepšovat. Pro stručnost a přehlednost zápisu zavedme označení

$$E\{\mathbf{f}_D\} = \boldsymbol{\mu}_{f_D}, \quad E\{\mathbf{f}_I\} = \boldsymbol{\mu}_{f_I}. \quad (5.62)$$

Ze vztahu (5.58) a z podmínky (5.59) snadno získáme předpis

$$\mathbf{b} = \boldsymbol{\mu}_{f_I} - \mathbf{W}\boldsymbol{\mu}_{f_D}. \quad (5.63)$$

Dosadíme-li do podmínky ortogonalit (5.61) vztah (5.58) a (5.63), pak postupně dostaneme

$$\begin{aligned} \mathbf{0} &= E\left\{(\mathbf{f}_R - \mathbf{f}_I)(\mathbf{f}_D - \boldsymbol{\mu}_{f_D})^T\right\} = E\left\{(\mathbf{W}\mathbf{f}_D + \boldsymbol{\mu}_{f_I} - \mathbf{W}\boldsymbol{\mu}_{f_D} - \mathbf{f}_I)(\mathbf{f}_D - \boldsymbol{\mu}_{f_D})^T\right\} \\ &= E\left\{[\mathbf{W}(\mathbf{f}_D - \boldsymbol{\mu}_{f_D}) - (\mathbf{f}_I - \boldsymbol{\mu}_{f_I})][\mathbf{f}_D - \boldsymbol{\mu}_{f_D}]^T\right\} = \mathbf{W}\mathbf{C}_{f_D f_D} - \mathbf{C}_{f_I f_D}, \end{aligned} \quad (5.64)$$

kde $\mathbf{C}_{f_D f_D}$, $\mathbf{C}_{f_I f_D}$ jsou kovarianční matice. Ze vztahu (5.64) pak dále máme

$$\mathbf{W} = \mathbf{C}_{f_I f_D} \mathbf{C}_{f_D f_D}^{-1}. \quad (5.65)$$

Předpokládejme nyní, že vstupní obraz \mathbf{f}_I je degradován lineárním operátorem reprezentovaným maticí \mathbf{D} a šumem reprezentovaným vektorem \mathbf{v} . Pro degradovaný obraz \mathbf{f}_D proto můžeme psát

$$\mathbf{f}_D = \mathbf{D}\mathbf{f}_I + \mathbf{v}. \quad (5.66)$$

Uplatníme-li na vztah (5.66) operaci střední hodnoty a výsledek dosadíme do vztahu (5.63), máme

$$\mathbf{b} = \boldsymbol{\mu}_{f_I} - \mathbf{W}\mathbf{D}\boldsymbol{\mu}_{f_I} - \mathbf{W}\boldsymbol{\mu}_v, \quad (5.67)$$

kde $\boldsymbol{\mu}_v$ je vektor středních hodnot šumu. Dosazením vztahu (5.66) do (5.65) dále máme

$$\mathbf{W} = E\left\{[\mathbf{f}_I - \boldsymbol{\mu}_{f_I}][\mathbf{D}(\mathbf{f}_I - \boldsymbol{\mu}_{f_I}) + (\mathbf{v} - \boldsymbol{\mu}_v)]^T\right\} E\left\{[\mathbf{D}(\mathbf{f}_I - \boldsymbol{\mu}_{f_I}) + (\mathbf{v} - \boldsymbol{\mu}_v)][\mathbf{D}(\mathbf{f}_I - \boldsymbol{\mu}_{f_I}) + (\mathbf{v} - \boldsymbol{\mu}_v)]^T\right\}^{-1}. \quad (5.68)$$

Jestliže jsou náhodná pole \mathbf{f}_I a \mathbf{v} navzájem nekorelovaná, vychází

$$\mathbf{W} = \mathbf{C}_{f_I f_I} \mathbf{D}^T (\mathbf{D}\mathbf{C}_{f_I f_I} \mathbf{D}^T + \mathbf{C}_{v v})^{-1}. \quad (5.69)$$

Uvedený vztah můžeme dále zjednodušit za předpokladu, že pro každé dva různé body obrazu je ideální signál i šum nekorelovaný. V tomto případě jsou kovarianční matice $\mathbf{C}_{f_I f_I}$, $\mathbf{C}_{v v}$ diagonální. Zavedme další zjednodušení a předpokládejme kovarianční matice ve tvaru

$$\mathbf{C}_{f_I f_I} = \sigma_{f_I}^2 \mathbf{I}, \quad \mathbf{C}_{v v} = \sigma_v^2 \mathbf{I}, \quad (5.70)$$

kde \mathbf{I} je jednotková matice. Ze vztahu (5.69) pak dostaneme předpis

$$\mathbf{W} = \mathbf{D}^T \left[\mathbf{D}\mathbf{D}^T + \frac{\sigma_v^2}{\sigma_{f_I}^2} \mathbf{I} \right]^{-1}. \quad (5.71)$$

6 Bodové, algebraické a geometrické operace s obrazy

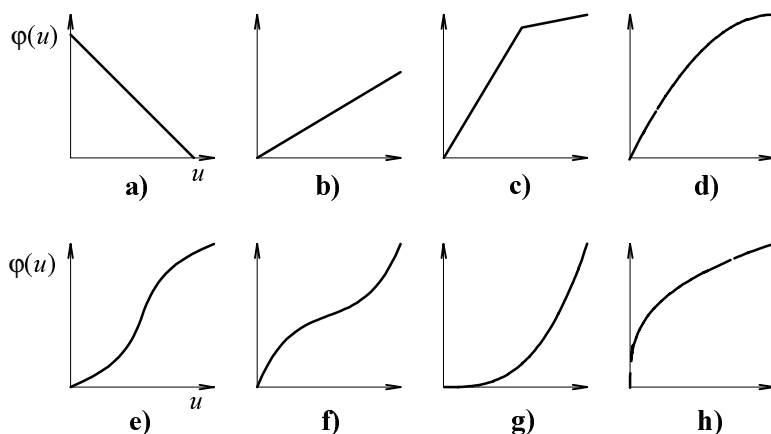
V praxi se často uplatňují operace, při nichž se modifikuje jas (případně intenzita barevných složek) tak, že jas výsledného obrazu v každém bodě je nějakou funkcí jasu původního obrazu v témže bodě. Takové operace nazýváme operacemi bodovými. Do operací algebraických vstupují zpravidla dva obrazy. Výsledný obraz se získává tak, že se jasy ve stejnohlých bodech vstupních obrazů sčítají, odčítají, násobí nebo dělí. Na rozdíl od bodových a algebraických operací mění operace geometrické tvar a rozměry obrazu. O uvedených třídách operací pojednáme v této kapitole podrobněji.

6.1 Bodové operace

Nechť $f(x,y)$ je obrazová funkce (jasová funkce) popisující vstupní obraz a $g(x,y)$ obrazová funkce popisující obraz výstupní. Bodovou operací rozumíme transformaci obrazu podle předpisu

$$g(x, y) = \varphi[f(x, y)]. \quad (6.1)$$

Ze vztahu (6.1) je zřejmé, že bodové operace vytvářejí výstupní obraz tak, že bodu o souřadnicích (x,y) přiřadí hodnotu jasu, která je nějakou funkcí φ jasu vstupního obrazu v bodě o stejných souřadnicích.



Obr. 6.1. Příklady funkcí realizujících bodové operace: **a), b)** lineární funkce dle vztahu (6.3), **c)** funkce po částech lineární, **d)** kvadratická parabola (6.4) ($c=1$), **e), f)** kubická parabola (6.5), **g), h)** exponenciální funkce (6.6) ($\gamma=3$, $\gamma=1/3$).

Důvody pro provádění bodových transformací mohou být rozmanité. Uvedme několik příkladů: Citlivost prvků, které se při snímání obrazu používají často není lineární. Při fotometrické kalibraci kompenzujeme bodovou operací nelinearitu snímacího prvku. Podobná je i situace na výstupu řetězce zpracovávajícího obraz, kde bodovou operací můžeme kompenzovat nelinearitu displeje. V systémech pro zpracování obrazu je jednou z nejpoužívanějších operací bodová operace provádějící změnu rozsahu nebo průběhu jasu.

Předpokládáme, že jas (označme jej u) vstupního obrazu může nabývat hodnot z intervalu $\langle u_{\min}, u_{\max} \rangle$ a požadujeme, aby hodnoty u_{\min}, u_{\max} byly transformovány na hodnoty $v_1 = \varphi(u_{\min}), v_2 = \varphi(u_{\max})$, které považujeme za známé. Zavedme hodnotu ξ vztahem

$$\xi = \frac{u - u_{\min}}{u_{\max} - u_{\min}}. \quad (6.2)$$

Jako příklady funkcí $\varphi(\xi)$ realizujících bodovou transformaci lze uvést např. následující funkce:

$$\varphi(\xi) = v_1 + \xi(v_2 - v_1), \quad (6.3)$$

$$\varphi(\xi) = v_1 + [\xi + c\xi(1-\xi)](v_2 - v_1), \quad (6.4)$$

$$\varphi(\xi) = (1 - 3\xi^2 + 2\xi^3)v_1 + (\xi - 2\xi^2 + \xi^3)v_1' + (3\xi^2 - 2\xi^3)v_2 + (-\xi^2 + \xi^3)v_2', \quad (6.5)$$

$$\varphi(\xi) = v_1 + \xi^\gamma(v_2 - v_1). \quad (6.6)$$

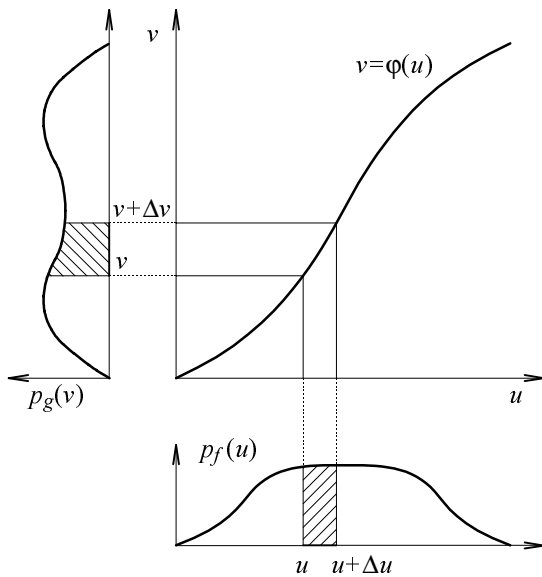
kde $v_1' = \varphi'(u)|_{u=u_{\min}}$, $v_2' = \varphi'(u)|_{u=u_{\max}}$ a c, γ jsou konstanty zvolené tak, aby bylo dosaženo žádaného průběhu funkce φ . Příklady průběhů funkcí jsou uvedeny na obr 6.1.

6.1.1 Vliv bodových operací na histogram jasu

Nejprve uvedeme obecné úvahy týkající se histogramu jasu. Uvažujme diskrétní obrazy s jasem popsaným celými čísly. Obraz je definován nad oblastí sestávající z N bodů. Nechť N_z označuje počet bodů, jejichž jas je právě z . Pro histogram (označme jej H) jasu pak máme $H(z) = N_z$. Probíráme postupně jednotlivé body obrazu a vyšetřujeme jejich jas. Na hodnotu jasu můžeme pohlížet jako na náhodnou proměnnou (označme ji \mathbf{b}). Hodnota $p(z) = H(z)/N = N_z/N$ udává pravděpodobnost jevu, že jas bodu je právě z . Je zřejmé, že $H(z) = Np(z)$. Poznamenejme, že pro hodnotu $p(z)$ bývá též někdy používáno názvu normalizovaný histogram. Pro teoretické úvahy je výhodné zavést odpovídající pojem také pro spojité obrazy s reálnými hodnotami jasu. Nechť $\mathcal{P}\{z < \mathbf{b} \leq z + \Delta z\}$ označuje pravděpodobnost jevu, že jas bodu nabývá hodnoty z z intervalu $(z, z + \Delta z)$, \mathcal{S} nechť je plocha celého obrazu a $\mathcal{S}\{z < \mathbf{b} \leq z + \Delta z\}$ nechť je plocha těch jeho částí, kde jas nabývá hodnoty z z intervalu $(z, z + \Delta z)$. Ve spojitém případě má $p(z)$ význam hustoty pravděpodobnosti

$$p(z) = \lim_{\Delta z \rightarrow 0} \frac{\mathcal{P}\{z < \mathbf{b} \leq z + \Delta z\}}{\Delta z} = \lim_{\Delta z \rightarrow 0} \frac{\mathcal{S}\{z < \mathbf{b} \leq z + \Delta z\}}{\mathcal{S}\Delta z}. \quad (6.7)$$

Vyšetřujeme nyní, jak se histogram jasu mění vlivem bodových operací. Uvažujme spojité obrazy s reálnými hodnotami jasu. Předpokládejme, že hustota pravděpodobnosti jasu vstupního obrazu f je známa (označme ji p_f) a že na obraz f byla aplikována bodová operace popsaná rovnicí (6.1). Ukážeme, jak lze stanovit hustotu pravděpodobnosti jasu výstupního obrazu g (označme ji p_g). Nechť u, v značí vstupní resp. výstupní jas. Je tedy $v = \varphi(u)$. Předpokládejme, že φ je bijekcí, a že lze proto konstruovat inverzní funkci $u = \varphi^{-1}(v)$. Uvažujme interval $\langle u, u + \Delta u \rangle$ vstupních jasů. Tomuto intervalu odpovídá interval $\langle v, v + \Delta v \rangle = \langle \varphi(u), \varphi(u + \Delta u) \rangle$ jasů výstupních (obr. 6.2). Řekněme, že ve vstupním obraze měla jistá plocha obrazu jas z intervalu $\langle u, u + \Delta u \rangle$. Po transformaci bude mít tatáž plocha jas z intervalu $\langle v, v + \Delta v \rangle$. Vzhledem ke vztahu (6.7) proto máme



Obr.6.2. Vliv bodových operací na histogram obrazu.

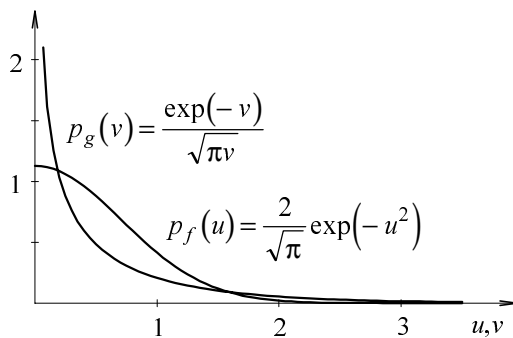
Jestliže jsou hodnoty $\Delta u, \Delta v$ dostatečně malé, pak můžeme na základě vztahu (6.8) také psát

$$\int_v^{v+\Delta v} p_g(z) dz = \int_u^{u+\Delta u} p_f(z) dz. \quad (6.8)$$

$$p_g(v) \Delta v = p_f(u) \Delta u. \quad (6.9)$$

Odtud získáme pro výpočet hustoty pravděpodobnosti $p_g(v)$ předpis

$$p_g(v) = \frac{p_f(u)}{\Delta v / \Delta u} = \frac{p_f(u)}{\varphi'(u)} = \frac{p_f[\varphi^{-1}(v)]}{\varphi'[\varphi^{-1}(v)]}. \quad (6.10)$$



Obr.6.3. Hustoty pravděpodobnosti jasu z příkladu 6.1.

Příklad 6.1. Předpokládejme, že jasy vstupního obrazu jsou nezáporné ($u \geq 0$) a že hustota pravděpodobnosti jasu vstupního obrazu je popsána výrazem

$$p_f(u) = \frac{2}{\sqrt{\pi}} \exp(-u^2).$$

Výstupní obraz je definován vztahem

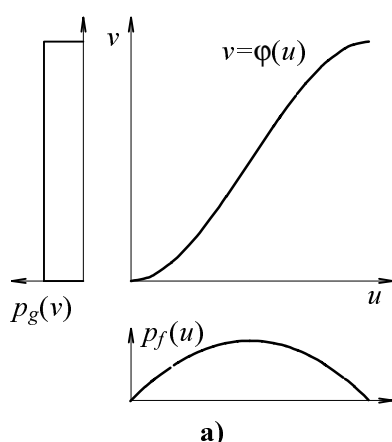
$$g(x, y) = f^2(x, y).$$

Určíme hustotu pravděpodobnosti jasu výstupního obrazu. Máme $v = \varphi(u) = u^2$, $u = \varphi^{-1}(v) = \sqrt{v}$. Podle vztahu (6.10) pak pro hledanou hustotu dostáváme

$$p_g(v) = \frac{\exp(-v)}{\sqrt{\pi v}}.$$

6.1.2 Vyrovnání histogramu

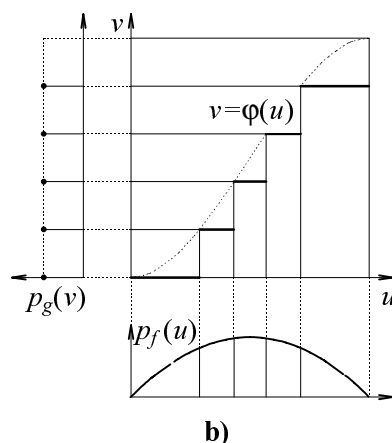
Vyrovnáním histogramu rozumíme provedení takové bodové operace, která zajistí rovnoměrné rozložení hustoty pravděpodobnosti jasu výsledného obrazu. Důvody k provedení operace mohou být následující: 1) Vyrovnání může vést ke zlepšení kvality obrazu subjektivně vnímané pozorovatelem. 2) Vyrovnání může být použito jako transformace zajišťující normalizované jednotné podmínky před dalším zpracováním obrazů (např. před segmentací a rozpoznáním). (Poznamenejme, že pro jistou třídu obrazů může však vyrovnání histogramu vést naopak i ke zhoršení subjektivně vnímané kvality. Jako příklad uvažte např. obraz vzniklý naskenováním výkresu obsahujícího černé čáry na bílém podkladě.) Zaměřme se nyní na problém stanovení funkce φ ze vztahu (6.1) tak, aby rozložení pravděpodobnosti p_g bylo rovnoměrné. Úvahy se zjednoduší, jestliže budeme předpokládat, že jas vstupního i výstupního obrazu je reprezentován reálnými čísly (obr.6.4a). Z podmínky rovnoměrného rozložení hustoty pravděpodobnosti máme $p_g(v) = 1/(v_{\max} - v_{\min})$. Použitím vztahu (6.10) pak dostaneme



$$\varphi'(u) = \frac{p_f(u)}{p_g(v)} = p_f(u)(v_{\max} - v_{\min}). \quad (6.11)$$

$$\text{Odtud } \varphi(u) = \int_{u_{\min}}^u \varphi'(z) dz = (v_{\max} - v_{\min}) \int_{u_{\min}}^u p_f(z) dz. \quad (6.12)$$

Protože je podle předpokladu hustota $p_f(z)$ známá, je vztah (6.12) předpisem pro výpočet hledané funkce φ . Jestliže jsou jasy reprezentovány celými čísly, je zapotřebí několik dalších upřesnění. Zaměřme se nejprve na případ, kdy je celými čísly reprezentován pouze jas výstupního obrazu. Máme tedy $v \in \{v_{\min}, v_{\min}+1, \dots, v_{\max}\}$. Jasy vstupního obrazu jsou reálné. Funkce φ má v tomto případě stupňovitý tvar (obr.6.4b). Pro $z \in (u_v, u_{v+1})$ platí $\varphi(z) = v$. Problémem je stanovit pro jednotlivé hodnoty v odpovídající hodnoty u_v . S přihlédnutím ke skutečnosti, že v tomto případě je $p_g(v) = 1/(v_{\max} - v_{\min} + 1)$, dostáváme z podmínky rovnoměrného rozdělení pravděpodobností pro hodnotu u_v ($v_{\min} \leq v \leq v_{\max}$) následující rovnici



$$\frac{v - v_{\min}}{(v_{\max} - v_{\min} + 1)} = \int_{u_{\min}}^{u_v} p_f(z) dz. \quad (6.13)$$

Jestliže je celými čísly reprezentován jas jak výstupního tak i vstupního obrazu, pak je funkce φ diskrétní. Stejně jako v předchozím případě, je i zde problémem stanovit pro jednotlivé hodnoty v odpovídající hodnoty u_v . Hledanou hodnotu u_v lze určit jako nejmenší hodnotu, pro kterou je splněna nerovnost

$$\frac{v - v_{\min}}{(v_{\max} - v_{\min} + 1)} \leq \sum_{z=u_{\min}}^{u_v} p_f(z). \quad (6.14)$$

Obr.6.4. Vyrovnání histogramu jasu.

Jak z uvedeného postupu vyplývá, nelze v případě, kdy je celými čísly reprezentován jas výstupního i vstupního obrazu, dosáhnout vyrovnání histogramu zcela přesně. Histogram je vyrovnán pouze přibližně.

6.2 Algebraické operace s obrazy

Nechť $f_1(x,y)$, $f_2(x,y)$ jsou známé obrazy. Algebraickou operací rozumíme postup, kdy je výsledný obraz $g(x,y)$ získáván pomocí některého z následujících vztahů:

$$g(x,y) = f_1(x,y) + f_2(x,y), \quad (6.15)$$

$$g(x, y) = f_1(x, y) - f_2(x, y), \quad (6.16)$$

$$g(x, y) = f_1(x, y) \times f_2(x, y), \quad (6.17)$$

$$g(x, y) = f_1(x, y) \div f_2(x, y). \quad (6.18)$$

Uvedme několik příkladů, kdy mohou být algebraické operace použity v praxi: Při snímání kamerou je někdy vlivem nedokonalosti objektivu obrázků uprostřed světlejší a na krajích (nejvíce v rozích) tmavší. Uvedenou vadu objektivu můžeme považovat za multiplikační chybu a korigovat ji pomocí vztahů (6.17) nebo (6.18). Obraz f_1 je obraz, který má být korigován. Korekční hodnoty obsažené v obrazu f_2 snadno vypočítáme např. z hodnot získaných sejmutím bílé rovnoměrně osvětlené plochy. Podobně lze korigovat systematickou chybu vznikající nerovnoměrným osvětlením předlohy při snímání kamerou. Při snímání scény, v níž se pohybují objekty, lze rozdíl po sobě jdoucích obrazů využít k detekci pohybu. Použití sčítání obrazu pro redukci šumu ukážeme podrobněji v samostatné kapitole.

6.2.1 Vliv algebraických operací na histogram jasu

Nechť $p_1(z)$, $p_2(z)$ jsou hustoty pravděpodobnosti jasu v obrazech $f_1(x, y)$, $f_2(x, y)$. Stanovíme hustotu pravděpodobnosti $p(z)$ jasu obrazu $g(x, y)$ vzniklého součtem dle vztahu (6.15). Označme $p_{1,2}(u, v)$ sdruženou hustotu pravděpodobnosti jasu v prvním a druhém obraze. Předpokládáme, že vstupní obrazy jsou nezávislé, a že proto platí $p_{1,2}(u, v) = p_1(u)p_2(v)$. Podle vztahu (6.15) je hodnoty z jasu v obraze $g(x, y)$ dosaženo tehdy, jestliže při jasu v v obraze f_2 je jas v obraze f_1 roven hodnotě $u = z - v$. Odtud vychází pro hledanou hustotu $p(z)$ vztah

$$p(z) = \int_{-\infty}^{\infty} p_{1,2}(z - v, v) dv = \int_{-\infty}^{\infty} p_1(z - v) p_2(v) dv = p_1(z) * p_2(z). \quad (6.19)$$

Vztah (6.19) ukazuje, že hustota pravděpodobnosti jasu výsledného obrazu je rovna konvoluci hustot pravděpodobností jasů obou obrazů vstupních. Získaný výsledek lze snadno rozšířit také na rozdíl obrazů dle vztahu (6.16). Postačí položit $\hat{f}_2(x, y) = -f_2(x, y)$, čímž převedeme rozdíl na součet. Pro hustotu pravděpodobnosti jasu obrazu \hat{f}_2 máme $\hat{p}_2(z) = p_2(-z)$.

6.2.2 Využití průměrování k redukci šumu

V tomto odstavci podrobněji popíšeme jednu z praktických aplikací sčítání obrazů. Nechť $f(x, y)$ je teoretický obraz neměnný v čase. Při snímání získáváme ovšem obraz, který je poškozen šumem. Obraz proto snímáme opakovaně M -krát. Při i -tém sejmutí obdržíme obraz $f(x, y) + n_i(x, y)$, kde n_i je šum. Za aproximaci teoretického obrazu můžeme intuitivně brát průměr

$$\hat{f}(x, y) = \frac{1}{M} \sum_{i=1}^M [f(x, y) + n_i(x, y)]. \quad (6.20)$$

Ukážeme, že použití průměru podle rovnice (6.20) objektivně vede ke zlepšení kvality sejmutého obrazu. K hodnocení kvality použijeme poměru signál-šum, který je definován vztahem

$$SNR(x, y) = \frac{f(x, y)}{\sqrt{E\{n^2(x, y)\}}}. \quad (6.21)$$

Pro hodnotu poměru signál-šum v obraze vzniklém průměrováním M obrazů odtud vychází

$$SNR_M(x, y) = \frac{f(x, y)}{\sqrt{E\left\{\left[\frac{1}{M} \sum_{i=1}^M n_i(x, y)\right]^2\right\}}}. \quad (6.22)$$

Výraz (6.22) dále upravíme. Při tom využijeme následujících předpokladů: 1) $n_i(x, y)$ je náhodná proměnná, pro níž platí $E\{n_i(x, y)\} = 0$; 2) náhodné proměnné $n_i(x, y)$, $n_j(x, y)$ jsou nekorelované, a proto

platí $E\{n_i(x,y) n_j(x,y)\} = E\{n_i(x,y)\} E\{n_j(x,y)\}$; 3) $E\{n_i^2(x,y)\} = E\{n_j^2(x,y)\} = E\{n^2(x,y)\}$. Postupně dostáváme

$$SNR_M(x,y) = \frac{Mf(x,y)}{\sqrt{E\left\{\sum_{i=1}^M n_i^2(x,y)\right\}}} = \sqrt{M} \frac{f(x,y)}{\sqrt{E\{n^2(x,y)\}}} = \sqrt{M} SNR(x,y). \quad (6.23)$$

Vztah (6.23) ukazuje, že průměrováním z M obrazů se poměr signál-šum zvětšil \sqrt{M} -krát.

6.3 Geometrické transformace obrazu

Geometrické transformace mění tvar nebo rozměry obrazu. Obecně jsou popsány vztahem

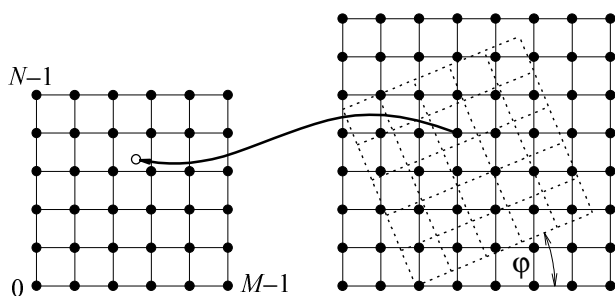
$$g(x,y) = f[\varphi(x,y), \psi(x,y)]. \quad (6.24)$$

Je zřejmé, že geometrická transformace je zcela zadána funkcemi $\varphi(x,y)$, $\psi(x,y)$. Vztah (6.24) navíc podává i praktický návod na to, jak transformaci provést. Předpokládejme, že obrazy $f(x,y)$ i $g(x,y)$ jsou diskrétní. Na základě vztahu (6.24) lze postupně probírat všechny body výstupního obrazu g a pro každý bod (x,y) určit odpovídající polohu $[\varphi(x,y), \psi(x,y)]$ ve vstupním obraze f . Hodnota jasu $f[\varphi(x,y), \psi(x,y)]$ je pak vzata jako hodnota $g(x,y)$. Zatímco x,y nabývají celočíselných hodnot (postupujeme po celočíselných souřadnicích pixelů výstupního obrazu), jsou hodnoty $\varphi(x,y)$, $\psi(x,y)$ obvykle reálné, nikoli celočíselné. Podle předpokladu je však i vstupní obraz f diskrétní, a proto i pro něj jsou hodnoty jasu známy jen v bodech o celočíselných souřadnicích. Pro získání hodnot v mezilehlých bodech o neceločíselných souřadnicích lze využít metod, které byly podrobněji popsány v podkapitole 4.4. V praxi se nejčastěji používá bilineární interpolace nebo interpolace vyššího řádu.

Zabývejme se nyní podrobněji funkcemi $\varphi(x,y)$, $\psi(x,y)$. Nejobecnějším požadavkem je, aby dvojice $[\varphi(x,y), \psi(x,y)]$ byla homeomorfismem (nesplnění tohoto požadavku by způsobilo „roztrhání“ obrazu). Kromě toho bývají obvykle požadovány i jiné vlastnosti. Jsou-li funkce $\varphi(x,y)$, $\psi(x,y)$ lineární, pak je transformace (6.24) transformací afinní. Afinní transformace je zadána šesti hodnotami a_{11} , a_{12} , a_{21} , a_{22} , b_1 , b_2 a lze ji zapsat následujícím vztahem

$$\begin{bmatrix} \varphi(x,y) \\ \psi(x,y) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}. \quad (6.25)$$

Příklad 6.2. Uvažujme diskrétní obraz rozměru $M \times N$ bodů popsáný obrazovou funkcí $f(m,n)$ (obr.6.5). Provedeme rotaci obrazu kolem levého dolního rohu o úhel φ ($0 \leq \alpha \leq \pi/2$). Rozměry obrazu vzniklého rotací budou $[(N-1)\sin \varphi + (M-1)\cos \varphi + 1] \times [(N-1)\cos \varphi + (M-1)\sin \varphi + 1]$ bodů. Počátek $(0,0)$ předpokládáme u obou obrazů v levém dolním rohu. Označme \tilde{x}, \tilde{y} souřadnice v původním a x,y souřadnice v novém obraze. Pro transformaci souřadnic z původního do nového obrazu snadno získáme předpis



Obr. 6.5. Afinní transformace obrazu - rotace.

$x = \tilde{x} \cos \varphi - \tilde{y} \sin \varphi + (N-1) \sin \varphi,$
 $y = \tilde{x} \sin \varphi + \tilde{y} \cos \varphi.$

Odtud máme

$$\begin{bmatrix} \varphi(x,y) \\ \psi(x,y) \end{bmatrix} = \begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix} = \begin{bmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + (N-1) \begin{bmatrix} -\sin \varphi \cos \varphi \\ \sin^2 \varphi \end{bmatrix}.$$

Prakticky provedeme transformaci tak, že probíráme všechny body výstupního obrazu jeden po druhém. Poloha bodu ve výstupním obraze je specifikována celočíselnou dvojicí x,y . Jestliže platí

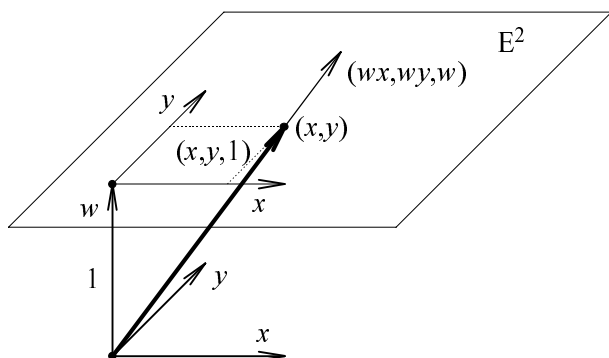
$[0 \leq \varphi(x,y) \leq M-1] \& [0 \leq \psi(x,y) \leq N-1]$, určíme hodnotu $g(x,y)$ jako hodnotu $f[\varphi(x,y), \psi(x,y)]$, kterou vypočítáme interpolací. V opačném případě položíme např. $g(x,y)=0$. •

Příklad 6.3. Geometrickou transformací, která se v praxi uplatňuje velmi často, je změna rozměru obrazu. Opět uvažujme diskrétní obraz rozměru $M \times N$ bodů. Tento obraz má být transformován tak, aby rozměry výstupního obrazu byly $P \times Q$. Snadno obdržíme vztah

$$\begin{bmatrix} \varphi(x,y) \\ \psi(x,y) \end{bmatrix} = \begin{bmatrix} \frac{P-1}{M-1} & 0 \\ 0 & \frac{Q-1}{N-1} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \quad \bullet$$

Obecnější než transformace afinní jsou transformace projektivní. Jak afinní tak projektivní transformace zachovávají linearitu transformovaných útvarů (přímky se transformují opět na přímky). Na rozdíl od afinních transformací však projektivní transformace nezachovávají rovnoběžnost. Projektivní transformace se uplatňují zejména v případech, které nějakým způsobem souvisí se středovým promítáním. Příkladem praktické úlohy je transformace, při níž má být obraz deformován tak, aby budil dojem, že je umístěn na nějaké obecně umístěné rovině v prostoru, kterou zobrazujeme středovým promítáním.

Při popisu projektivních transformací se zpravidla používá homogenních souřadnic. Homogenní souřadnice reprezentují bod v n -rozměrném afinním prostoru jako směr v jistém přidruženém $(n+1)$ -rozměrném vektorovém prostoru. Uvažujme dvojrozměrný afinní prostor a zavedme přidružený trojrozměrný vektorový prostor a souřadnou soustavu v něm tak, jak je znázorněno na obr.6.6. V afinním prostoru uvažujme bod, jehož afinní souřadnice jsou (x,y) . Homogenními souřadnicemi uvažovaného bodu je trojice (wx,wy,w) , kde w je libovolné reálné číslo $w \neq 0$. V homogenních souřadnicích je tedy každý bod reprezentován nekonečným počtem vektorů v přidruženém vektorovém prostoru, které jsou ovšem všechny navzájem kolinéární. Speciálně můžeme také volit $w=1$. Dostaneme tak trojici $(x,y,1)$. Zvláštní význam této volby je patrný z obr.6.6. Je-li naopak trojice (x,y,w) homogenními souřadnicemi



Obr.6.6. Zavedení homogenních souřadnic.

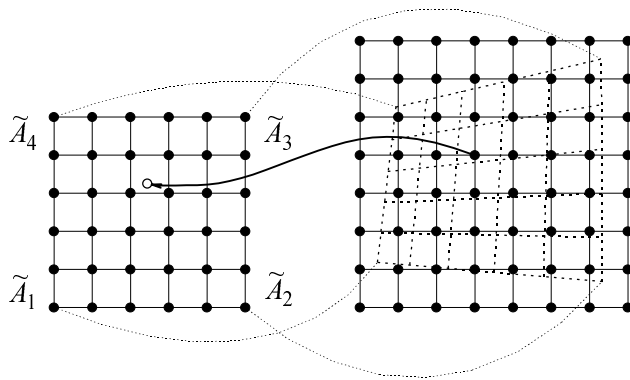
nějakého bodu, pak také trojice $(x/w,y/w,1)$ je homogenními souřadnicemi téhož bodu a dvojice $(x/w,y/w)$ je jeho souřadnicemi afinními. Projektivní transformace bodu o afinních souřadnicích (x,y) je popsána rovnicí

$$\begin{bmatrix} \varphi(x,y) \\ \psi(x,y) \\ \omega(x,y) \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}. \quad (6.26)$$

Homogenní souřadnice bodu po transformaci jsou $(\varphi(x,y), \psi(x,y), \omega(x,y))$, odkud pro afinní souřadnice plyne vztah $(\varphi(x,y)/\omega(x,y), \psi(x,y)/\omega(x,y))$. Ze vztahu (6.26) vyplývá, že ve dvojrozměrném prostoru je projektivní transformace je popsána devíti hodnotami $p_{11}, p_{12}, \dots, p_{33}$. Uvedené hodnoty lze ovšem násobit libovolnou nenulovou reálnou konstantou, aniž by se výsledek transformace změnil (násobením se sice změní délka vektoru na levé straně vztahu (6.26), ale bod, který je vektorem reprezentován, je stále tentýž).

Příklad 6.4. Uvažujme projektivní transformaci zadanou tak, že pro čtveřici bodů $\tilde{A}_1, \tilde{A}_2, \tilde{A}_3, \tilde{A}_4$ ve vstupním obraze jsou předepsány jim odpovídající body A_1, A_2, A_3, A_4 v obraze výstupním (na obr. 6.7 jsou jako body $\tilde{A}_1, \tilde{A}_2, \tilde{A}_3, \tilde{A}_4$ zvoleny rohy obrazu). Naznačíme postup odvození transformační matice ze vztahu (6.26). Souřadnice bodu \tilde{A}_i v původním obraze označíme \tilde{x}_i, \tilde{y}_i . Poloha korespondujícího bodu A_i ve výstupním obraze je předepsána souřadnicemi x_i, y_i . Dále zavedeme označení $\mathbf{p}_1 =$

$(p_{11}, p_{12}, p_{13})^T$, $\mathbf{p}_2 = (p_{21}, p_{22}, p_{23})^T$, $\mathbf{p}_3 = (p_{31}, p_{32}, p_{33})^T$, $\mathbf{x}_i = (x_i, y_i, 1)^T$. Na základě vztahu (6.26) nyní pro každý ze zadaných bodů dostáváme



Obr. 6.7. Projektivní transformace obrazu.

$$\begin{bmatrix} \tilde{w}_i \tilde{x}_i \\ \tilde{w}_i \tilde{y}_i \\ \tilde{w}_i \end{bmatrix} = \begin{bmatrix} \mathbf{p}_1^T \\ \mathbf{p}_2^T \\ \mathbf{p}_3^T \end{bmatrix} \mathbf{x}_i.$$

Ze třetí rovnice v uvedeném vztahu máme $\tilde{w}_i = \mathbf{p}_3^T \mathbf{x}_i$. Dosazením do prvních dvou rovnic po úpravě dostaneme

$$\begin{bmatrix} \mathbf{x}_i^T & \mathbf{0} & -\tilde{x}_i \mathbf{x}_i^T \\ \mathbf{0} & \mathbf{x}_i^T & -\tilde{y}_i \mathbf{x}_i^T \end{bmatrix} \begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \mathbf{p}_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Vidíme, že každý ze zadaných bodů přispívá k nalezení transformační matice dvěma rovnicemi, což dává celkem osm rovnic pro devět neznámých. Protože víme, že transformační matici je možné násobit libovolným reálným číslem, aniž by se výsledek transformace změnil, zvolíme k zajištění jednoznačnosti řešení doplňující podmínku. Nejjednodušší se prakticky realizuje postup, kdy některý z prvků transformační matice položíme roven předem zvolené hodnotě (prosíme čtenáře, aby promyslel úskalí tohoto řešení).

Nelineární geometrickou transformací obrazu rozumíme transformaci, kde $\varphi(x, y)$, $\psi(x, y)$ jsou nelineární funkce. Nečastěji se jedná o polynomy. Můžeme je volit např. ve tvaru

$$\varphi(x, y) = x + \Delta x(x, y), \quad \psi(x, y) = y + \Delta y(x, y), \quad (6.27)$$

kde

$$\Delta x(x, y) = \sum_{r=0}^m \sum_{s=0}^{m-r} a_{rs} x^r y^s, \quad \Delta y(x, y) = \sum_{r=0}^m \sum_{s=0}^{m-r} b_{rs} x^r y^s \quad (6.28)$$

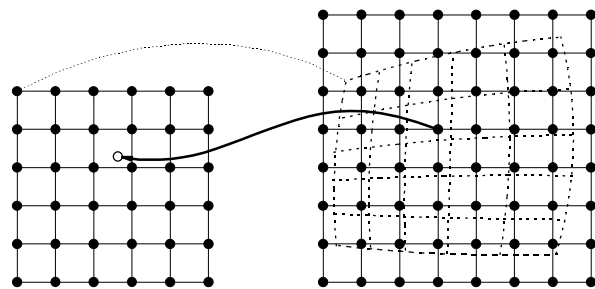
nebo také

$$\Delta x(x, y) = \sum_{r=0}^m \sum_{s=0}^m a_{rs} x^r y^s, \quad \Delta y(x, y) = \sum_{r=0}^m \sum_{s=0}^m b_{rs} x^r y^s. \quad (6.29)$$

Ze vztahu (6.27) je zřejmé, že funkce $\Delta x(x, y)$, $\Delta y(x, y)$ popisují „přemístění“ bodu, ke kterému během transformace dojde. Zavedení funkcí $\Delta x(x, y)$, $\Delta y(x, y)$ je názorné a také výhodné pro praktickou implementaci. Polynomy, které se pro transformaci používají, v praxi zpravidla nebyvají příliš vysokého stupně (obvykle druhého nebo třetího). Neznámé hodnoty koeficientů a_{rs} , b_{rs} je možné stanovit tak, že pro jistý počet bodů (x_i, y_i) předepíšeme hodnoty $\varphi(x_i, y_i)$, $\psi(x_i, y_i)$. To dává dvě soustavy rovnic

$$\varphi(x_i, y_i) = x_i + \sum_{r=0}^m \sum_{s=0}^{m-r} a_{rs} x_i^r y_i^s, \quad \psi(x_i, y_i) = y_i + \sum_{r=0}^m \sum_{s=0}^{m-r} b_{rs} x_i^r y_i^s, \quad (6.30)$$

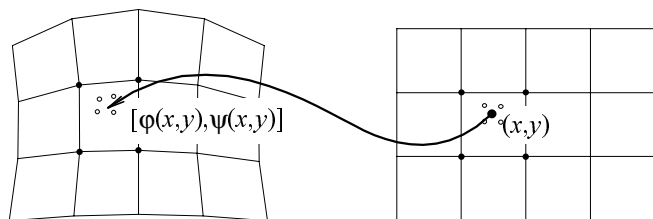
z nichž lze hodnoty a_{rs} , b_{rs} stanovit (v případě polynomu (6.29) lze postupovat analogicky). Předpisy (6.28), (6.29) obsahují $(m+2)(m+1)$ resp. $2(m+1)^2$ neznámých koeficientů. Každá dvojice $[(x_i, y_i), (\varphi(x_i, y_i), \psi(x_i, y_i))]$ dává vzniknout dvěma rovnicím. Je proto zapotřebí nejméně $(m+2)(m+1)/2$ resp.



Obr. 6.8. Nelineární geometrická transformace.

$(m+1)^2$ takových dvojic. Je-li počet dvojic právě roven uvedené hodnotě, obdržíme soustavu lineárních rovnic, kterou lze řešit eliminací. Je-li počet zadaných dvojic větší, obdržíme soustavu, která je předeterminovaná (a dost možná také nekonzistentní). O řešení předeterminovaných soustav pojednává dodatek B.

Je zřejmé, že čím je požadovaná geometrická transformace komplikovanější, tím vyšší stu-



Obr. 6.9. Geometrická transformace s rozdělením obrazu na oblasti.

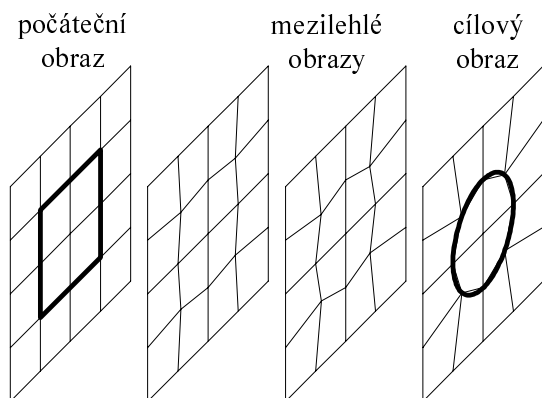
vají polynomy bilineární. Transformaci využívající dělení na podoblasti lze zadat tak, že výstupní obraz pokryjeme čtyřúhelníkovou sítí a ve vstupním obraze definujeme odpovídající síť transformovanou (obr.6.9). Síť zadáváme polohou uzlových kontrolních bodů ve vstupním i výstupním obraze. Nejnázne se implementuje případ, kdy je síť ve výstupním obraze obdélníková. Při praktické realizaci transformace se opět probírají všechny body výstupního obrazu. Podle toho, do které oblasti výstupního obrazu bod padl, se použije odpovídající transformační vztah. Pro *i*tou oblast má vztah tvar

$$\varphi_i(x, y) = x + (a_i x + b_i y + c_i x y + d_i), \quad (6.31)$$

$$\psi_i(x, y) = y + (e_i x + f_i y + g_i x y + h_i). \quad (6.32)$$

Neznámé hodnoty $a_i, b_i, c_i, d_i, e_i, f_i, g_i, h_i$ určíme pro každou oblast z podmínky, aby se vrcholy oblastí (uzlové kontrolní body) transformovaly tak, jak bylo předepsáno. Pro každou oblast máme čtyři vrcholy po dvou souřadnicích, což dává celkem osm rovnic potřebných pro určení hledaných hodnot.

Pro nelineární geometrické operace bývá někdy používáno termínu warping (borcení, kroucení). V televizní tvorbě se v poslední době stala populární také technika nazývaná morphing, která provádí plynulou změnu jednoho objektu v jiný. Předpokládáme, že je zadán počáteční a cílový obraz



Obr. 6.10. Přeměna čtverce na kružnici pomocí morphingu.

čáteční i cílový obraz. Polohu kontrolních bodů předepíšeme tak, aby bylo dosaženo požadované deformace obrazu (obr.6.10). Síť je nutné pokrýt i obrazy mezilehlé, v nichž lze však síť generovat automaticky pomocí interpolace. Ke generování *i*-tého mezilehlého obrazu použijeme vztahu

$$f_i(x, y) = (1 - \lambda) f_s[\varphi_{i,s}(x, y), \psi_{i,s}(x, y)] + \lambda f_d[\varphi_{i,d}(x, y), \psi_{i,d}(x, y)], \quad (6.34)$$

kde $\varphi_{i,s}, \psi_{i,s}, \varphi_{i,d}, \psi_{i,d}$ jsou po částech bilineární funkce transformující souřadnice z *i*-tého mezilehlého obrazu do obrazu počátečního resp. cílového. Uvedené funkce jsou určeny zadanou sítí.

peň polynomů je nutné v rovnicích (6.28), (6.29) zvolit. Jestliže by stupeň polynomů přesáhl rozumnou mez (řekněme např. 3 až 5), pak je lepší použít postupu, kdy vstupní i výstupní obraz rozdělíme na oblasti a transformační polynomy stanovíme zvlášť pro každou dílčí oblast. Stupeň polynomů může být v tomto případě nízký, protože s ohledem na složitost transformace můžeme volit jemnost dělení.

Nejčastěji se proto v tomto případě použí-

jeví se proto v tomto případě použí-
vají polynomy bilineární. Transformaci využívající dělení na podoblasti lze zadat tak, že výstupní obraz pokryjeme čtyřúhelníkovou sítí a ve vstupním obraze definujeme odpovídající síť transformovanou (obr.6.9). Síť zadáváme polohou uzlových kontrolních bodů ve vstupním i výstupním obraze. Nejnázne se implementuje případ, kdy je síť ve výstupním obraze obdélníková. Při praktické realizaci transformace se opět probírají všechny body výstupního obrazu. Podle toho, do které oblasti výstupního obrazu bod padl, se použije odpovídající transformační vztah. Pro *i*tou oblast má vztah tvar

$$f_i(x, y) = (1 - \lambda) f_s(x, y) + \lambda f_d(x, y) \quad (6.33)$$

při hodnotě λ postupně se zvyšující od 0 do 1. Výsledky samotného prolínání však zpravidla nejsou přesvědčivé. Lepšího dojmu lze dosáhnout kombinací prolínání s nelineární geometrickou transformací. K definování transformace lze použít síť kontrolních bodů, kterou pokryjeme po-

7 Komprese obrazu

Pro obrazy reprezentované maticí hodnot je charakteristická vysoká spotřeba paměti. Spotřeba paměti limituje možnosti archivace obrazů i možnosti jejich transportu v počítačových sítích. Je samozřejmě možné obrazy komprimovat běžnými metodami bezeztrátové komprese. Kompresní poměry, kterých by tak bylo možné dosáhnout (např. 2:1 až 4:1), jsou ale obvykle nižší, než by bylo žádoucí. Velká pozornost byla proto v posledním desetiletí věnována vývoji kompresních metod založených na předpokladu, že v mnoha aplikacích příliš nevedí, dojde-li během komprese k malé změně obrazu. Tento postup je označován za kompresi se ztrátou informace. Výhodou komprese se ztrátou informace jsou vyšší dosahované kompresní poměry, které se v závislosti na požadované kvalitě obrazu pohybují v rozmezí přibližně 7:1 až 30:1. Protože metody bezeztrátové komprese jsou všeobecně známy, zaměříme se v této kapitole pouze na kompresi ztrátovou. Popíšeme princip komprese JPEG, MPEG a komprese fraktální.

7.1 Redukce objemu barvonosné informace

Reprezentace barev v prostoru RGB není z hlediska komprese obrazů a videosekvencí příliš výhodná. Je dobře známo, že lidské oko je citlivější na změny jasu než na změny barvy. Nabízí se myšlenka využít této nedokonalosti lidského zraku i při kompresi obrazů. K redukci barvonosné informace se barva reprezentuje nečastěji v některém z prostorů YC_bC_r , YUV , YIQ . Složka Y je jas, zbývající dvě složky jsou barvonosné. Převodní vztahy mezi uvedenými prostory a prostorem RGB jsou následující:

$$Y = 0.299R + 0.587G + 0.114B, \quad C_b = B - Y, \quad C_r = R - Y, \quad (7.1)$$

$$U = 0.565C_b, \quad V = 0.713C_r, \quad (7.2)$$

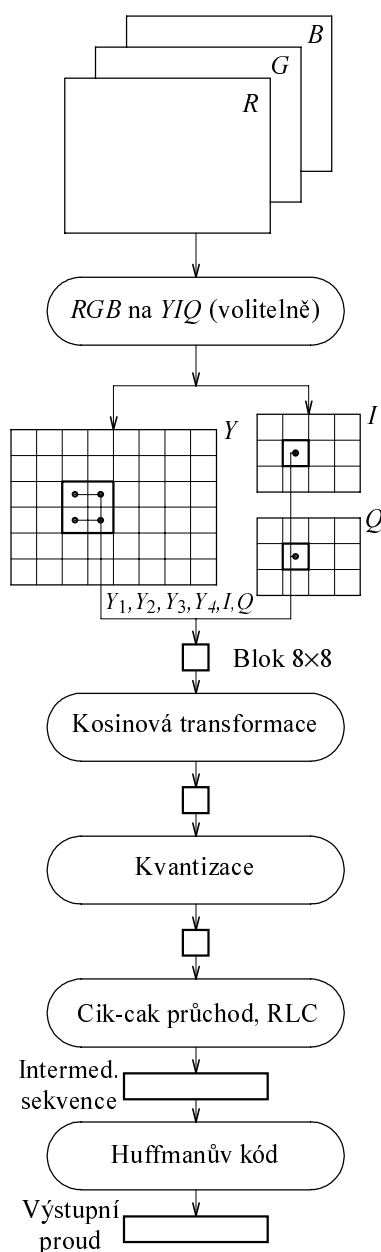
$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \quad \begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.528 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \quad (7.3)$$

Aby se dosáhlo redukce objemu barvonosné informace, bývají složky C_b, U, I, C_r, V, Q vzorkovány s nižší frekvencí než složka Y . Nejčastěji se používá formátů označovaných kódy 4:2:0 nebo 4:2:2. Formát 4:2:0 používá pro barvonosnou informaci polovičního počtu vzorků na řádku a polovičního počtu řádků. Čtveřice 2×2 pixelů obrazu je proto reprezentována čtyřmi hodnotami Y , jednou hodnotou $C_b(U, I)$ a jednou hodnotou $C_r(V, Q)$. Formát 4:2:2 používá polovičního počtu vzorků na řádku a plného počtu řádků. Čtveřice 2×2 pixelů obrazu je proto reprezentována čtyřmi hodnotami Y , dvěma hodnotami $C_b(U, I)$ a dvěma hodnotami $C_r(V, Q)$. Dodejme, že 4:4:4 je formát, který používá plného počtu vzorků i pro složky barvonosné. U JPEG komprese je redukce objemu barvonosné informace nepovinná, ale obvyklá. U MPEG-1 komprese se redukce provádí vždy. Používá se barevného prostoru YC_bC_r a formátu 4:2:0. Komprese MPEG-2 podporuje i formát 4:2:2. Pro úplnost poznamenejme, že myšlenka redukce objemu barvonosné informace není nová. Byla využita i při návrhu dnes běžně používaných televizních soustav. Televizní soustava PAL používá prostoru YUV . Složka Y je přenášena v pásmu širokém 5 MHz, zatímco složky U, V v pásmech širokých pouze 1.3 MHz. Americká televizní soustava NTSC používá prostoru YIQ . Jednotlivé složky jsou v tomto případě přenášeny v pásmech širokých 4, 1.5, 0.6 MHz.

7.2 Komprese JPEG

JPEG je zkratkou pro Joint Photographic Expert Group. Výsledek práce této skupiny byl v roce 1992 přijat jako standard pro kompresi nepohyblivých obrazů. Principy JPEG komprese popíšeme v této podkapitole podrobněji. Protože se jedná o rozsáhlý standard, který na mnoha místech umožňuje volbu z více variant, zaměříme se z pochopitelných důvodů v tomto textu vždy pouze na variantu nejtýpčtější, a to bez toho, abychom na to dále upozorňovali a ostatní varianty systematicky vyjmenovávali. Pro vyčerpávající informaci odkazujeme eventuální zájemce na normu samotnou (CCITT 92).

Komprese JPEG je obvykle prováděna v následujících krocích: Obraz se převede do prostoru umožňujícího redukci objemu barvosné informace (např. do prostoru YIQ). Provede se snížení počtu vzorků v barvosných rovinách (např. formátem 4:2:0). Rovina Y i roviny barvosné se rozdělí na bloky 8×8 bodů. Další zpracování probíhá po tzv. makroblocích. Makroblok se skládá ze čtyř bloků v rovině Y , kde tyto čtyři bloky vytvářejí oblast 16×16 pixelů, a dále z bloků obsahujících barvosné složky pro tutéž oblast. Při kódování 4:2:0 je pro každou složku zapotřebí po jednom bloku, při kódování 4:2:2 jsou to pro každou složku dva bloky. V dalším výkladu se omezíme na kódování 4:2:0, kdy pro jeden makroblok máme celkem šest bloků, které jsou na obr. 7.1 označeny jako bloky Y_1, Y_2, Y_3, Y_4, I, Q . Zpracování jednotlivých bloků v makrobloku probíhá v tom pořadí, jak byly právě uvedeny, tj. v pořadí Y_1, Y_2, Y_3, Y_4, I, Q . Předpokládejme, že hodnoty v bloku jsou popsány celými čísly $\langle 0, 1, \dots, 2^P - 1 \rangle$. Pro každý blok se provede následující: Provede se posunutí hodnot tak, aby byly z intervalu $\langle -2^{P-1}, \dots, 2^{P-1} - 1 \rangle$. Dále se provede diskrétní kosinová transformace bloku. Dopředná kosinová transformace je definována předpisem



Obr. 7.1. Schéma komprese JPEG.

$$F(k, l) = \frac{1}{4} c(k) c(l) \sum_{m=0}^7 \sum_{n=0}^7 f(m, n) \cos \frac{(2m+1)k\pi}{16} \cos \frac{(2n+1)l\pi}{16}, \quad (7.4)$$

$$\text{kde,} \quad c(k), c(l) = \begin{cases} 1/\sqrt{2}, & k, l = 0 \\ 1, & \text{jinak} \end{cases}$$

Pro dekódování je zapotřebí i předpis pro zpětnou kosinovou transformaci

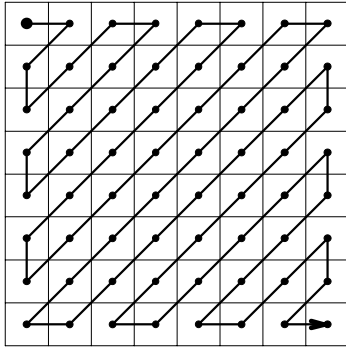
$$f(m, n) = \frac{1}{4} \sum_{k=0}^7 \sum_{l=0}^7 c(k) c(l) F(k, l) \cos \frac{(2m+1)k\pi}{16} \cos \frac{(2n+1)l\pi}{16}. \quad (7.5)$$

Je samozřejmé, že při praktické implementaci enkodéru i dekodéru je žádoucí použít pro výpočet kosinové transformace některého z rychlých algoritmů, kterých bylo vyvinuto nepřehledné množství. V závislosti na tom, jak přesně jsou reprezentovány kosinové členy a desetinné části mezivýsledků, se jednotlivé algoritmy mohou dosti lišit svoji přesností. Norma JPEG konkrétní algoritmus nepředepisuje, ale předepisuje přesnost, které musí být při transformaci dosaženo.

Po provedení kosinové transformace se pro každý zpracovávaný blok provede kvantování podle předpisu

$$F^Q(k, l) = \text{Round} \left[\frac{F(k, l)}{Q(k, l)} \right], \quad (7.6)$$

kde $Q(k, l)$ jsou prvky kvantizační tabulky. Rozměr kvantizační tabulky je 8×8 . Kvantizace se opírá o pozorování, že některé frekvence v obraze (zpravidla vysoké) lze reprezentovat dosti hrubě (tj. na malém počtu bitů), aniž by se obraz pozorovatelně poškodil. Pro mnoho hodnot k, l často dokonce vychází $F^Q(k, l) = 0$. Konkrétní kvantizační tabulku norma JPEG nepředepisuje. Teoreticky může být s ohledem na charakter obrazu stanovena speciální kvantizační tabulka. Kvantizační tabulka se stává součástí komprimovaného obrazu. Pro informaci však norma uvádí tabulky, které byly experimentálně stanoveny v rámci doporučení CCIR-601 a které jsou nyní v mnoha případech používány. Obr. 7.3c ukazuje kvantizační tabulku určenou pro složku Y (prvek $Q(0, 0)$ v levém horním rohu je dělitel pro stejnosměrnou složku, prvek $Q(7, 7)$ je dělitel pro nejvyšší frekvenční složku). Pro barvosné složky se používá tabulky uve-



Obr.7.2. Postup při kódování frekvenčního spektra bloku.

údaj (AMPLITUDE). Význam údajů je následující: Hodnota AMPLITUDE je nalezená nenulová hodnota $F^Q(k,l)$. Jedná se o celočíselnou hodnotu se znaménkem. Hodnota RUNLENGTH je počet nulových hodnot, které nalezenou nenulovou hodnotu $F^Q(k,l)$ předcházely. SIZE udává počet bitů, které jsou zapotřebí k reprezentaci hodnoty AMPLITUDE.

Předpokládejme, že platí $-2^P \leq f(m,n) \leq 2^P - 1$. Analýzou činnosti kosinové transformace lze ukázat, že pro hodnoty $F(k,l)$ platí $-2^{P+3} \leq F(k,l) \leq 2^{P+3} - 1$. Odtud např. plyne, že pro osmibitovou reprezentaci vstupního obrazu ($P=7$) platí $1 \leq \text{SIZE} \leq 10$ (vzhledem ke kvantování je hodnota SIZE zpravidla ještě menší). Hodnota RUNLENGTH je maximálně 15. Protože počet po sobě jdoucích nul může být větší, je zavedena speciální hodnota Symbol-1 = (15,0), která říká, že po průchodu patnácti nulovými hodnotami nebyla zatím nenulová hodnota $F^Q(k,l)$ nalezena. Hodnota Symbol-1 = (0,0) říká, že předchozí nenulová hodnota zapsaná do intermediální sekvence byla poslední a že žádné další nenulové hodnoty již následovat nebudou. Jedná se tedy o příznak konce bloku. Stejnoseměrná složka je kódována poněkud odlišně. V tomto případě je Symbol-1 = (SIZE), Symbol-2 = (AMPLITUDE). Hodnota AMPLITUDE je hodnotou $F^Q(0,0)$, která je však vztažena relativně k hodnotě $F^Q(0,0)$ předchozího zpracovávaného bloku. Hodnota SIZE opět udává počet bitů, které jsou zapotřebí k reprezentaci hodnoty AMPLITUDE. Pro osmibitovou reprezentaci vstupního obrazu vychází $\text{SIZE} \leq 11$.

Poslední fází komprese je zakódování intermediální sekvence. Hodnoty Symbol-1 jsou zakódovány Huffmanovým kódováním. Připomeňme, že v Huffmanově kódování jsou často používaným symbolům přiřazeny krátké kódy a symbolům používaným méně často kódy delší. Nalezení optimálního kódu závisí na povaze komprimovaných obrazů, a proto norma JPEG tabulku kódů pevně nepředepisuje. Mnoho enkodérů však pracuje s kódy uvedenými v informativní nezávazné příloze normy. Kódovací tabulka se stává součástí komprimovaného obrazu. Symbol-2 je číslo zakódované na proměnném počtu bitů. V tomto případě je způsob kódování normou pevně stanoven. Např. čísla -1,1 se kódují na jednom bitu takto (-1)→0, (1)→1. Čísla -3,-2,2,3 se kódují na dvou bitech: (-3)→00, (-2)→01, (2)→10, (3)→11 atd.

Ilustrujme postup provádění komprese JPEG na příkladě. Pro jednoduchost uvažujeme pouze jediný blok roviny Y . Obr.7.3a ukazuje hodnoty jasu ve zpracovávaném bloku (funkce $f(m,n)$). Obr.7.3b je výsledkem kosinové transformace (funkce $F(k,l)$). Obr. 7.3c je kvantizační matice $Q(k,l)$. Obr.7.3d ukazuje spektrum po provedení kvantování (funkce $F^Q(k,l)$). Intermediální sekvence má tvar (předpokládáme, že v předchozím bloku byla stejnoseměrná složka 12): (2)(3), (1,2)(-2), (0,1)(-1), (0,1)(-1), (0,1)(-1), (2,1)(-1), (0,0). Kódování hodnot Symbol-1 provedeme podle přílohy normy JPEG. V příloze nalezneme: (2)→011, (0,0)→1010, (0,1)→00, (1,2)→11011, (2,1)→11100. Způsob kódování potřebných hodnot Symbol-2 jsme uvedli již dříve. Výsledný kód pro uvažovaný blok je tedy: 011 11 11011 01 00 0 00 0 00 0 11100 0 1010. Nyní můžeme vyhodnotit efektivnost komprese. Informaci o hodnotách 64 vzorků se podařilo zkomprimovat do 31 bitů. Po komprimaci je tedy zapotřebí méně než 0.5 bitů na vzorek. Rekonstruovaný obraz je na obr.7.3f.

dené v tab. 7.1. Mnoho enkodérů umožňuje zadat požadovanou kvalitu obrazu. Ve většině implementací je problém jednoduše vyřešen tak, že se kvantizační tabulka násobí nějakou hodnotou. Při dekódování obrazu se provádí dekvantizace podle předpisu

$$F(k,l) \approx F^Q(k,l)Q(k,l). \quad (7.7)$$

Po provedení kvantizace pokračuje zpracování každého bloku převodem výsledku kvantizace do intermediální sekvence. Stejnoseměrná a střídavé složky se kódují odlišně. Popíšeme nejprve kódování složek střídavých. Pole $F^Q(k,l)$ se prochází postupně tak, jak je naznačeno na obr. 7.2. Nalezení nenulové hodnoty vede k vytvoření nové položky v intermediální sekvenci. Položka má tvar dvojice Symbol-1, Symbol-2. Symbol-1 obsahuje dva údaje (RUNLENGTH, SIZE), Symbol-2 obsahuje pouze jediný

139	144	149	153	155	155	155	155
144	151	153	156	159	156	156	156
150	155	160	163	158	156	156	156
159	161	162	160	160	159	159	159
159	160	161	162	162	155	155	155
161	161	161	161	160	157	157	157
162	162	161	163	162	157	157	157
162	162	161	161	163	158	158	158

a) Blok obsahující vzorky jasů.

235.6	-1.0	-12.1	-5.2	2.1	-1.7	-2.7	1.3
-22.6	-17.5	-6.2	-3.2	-2.9	-0.1	0.4	-1.2
-10.9	-9.3	-1.6	1.5	0.2	-0.9	-0.6	-0.1
-7.1	-1.9	0.2	1.5	0.9	-0.1	0.0	0.3
-0.6	-0.8	1.5	1.6	-0.1	-0.7	0.6	1.3
1.8	-0.2	1.6	-0.3	-0.8	1.5	1.0	-1.0
-1.3	-0.4	-0.3	-1.5	-0.5	1.7	1.1	-0.8
-2.6	1.6	-3.8	-1.8	1.9	1.2	-0.6	-0.4

b) Výsledek kosinové transformace.

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

c) Kvantizační matice pro složku Y.

15	0	-1	0	0	0	0	0
-2	-1	0	0	0	0	0	0
-1	-1	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

d) Spektrum po kvantizaci.

240	0	-10	0	0	0	0	0
-24	-12	0	0	0	0	0	0
-14	-13	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

e) Spektrum po dekvantizaci.

144	146	149	152	154	156	156	156
148	150	152	154	156	156	156	156
155	156	157	158	158	157	156	155
160	161	161	162	161	159	157	155
163	163	164	163	162	160	158	156
163	164	164	164	162	160	158	157
160	161	162	162	162	161	159	158
158	159	161	161	162	161	159	158

f) Rekonstruovaný obraz.

Intermediální sekvence: (2)(3), (1,2)(-2), (0,1)(-1), (0,1)(-1), (0,1)(-1), (2,1)(-1), (0,0)
(předpokládáme, že v předchozím bloku byla stejnosměrná složka 12).

Kódování hodnot Symbol-1: (2)→011, (0,0)→1010, (0,1)→00, (1,2)→11011, (2,1)→11100.

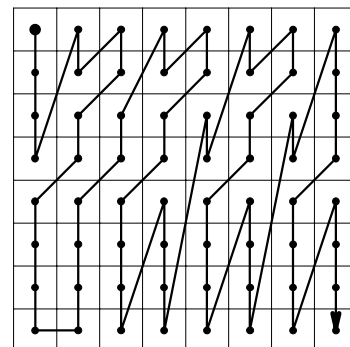
Kódování hodnot Symbol-2: (-1)→0, (1)→1, (-3)→00, (-2)→01, (2)→10, (3)→11.

Výsledný kód: 011 11 11011 01 00 0 00 0 00 0 11100 0 1010.

Obr.7.3. Příklad postupu komprese JPEG.

17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

Tab.7.1. Kvantizační matice pro barevné složky.



Obr. 7.4. Alternativní postup při kódování bloku v MPEG-2.

7.3 Komprese MPEG

MPEG je zkratkou pro Moving Picture Expert Group. Cílem práce této skupiny bylo standardizovat metody komprese videosignálu. Standard MPEG-1 byl dokončen v roce 1991 a jako norma přijat v roce 1992. Byl navržen zejména pro práci s obrazy o rozměru 352×288 pixelů, 25 rámců/s (odvozeno od televizní normy PAL) nebo 352×240, 30 rámců/s (odvozeno od normy NTSC) při datovém toku přibližně 1.5Mbit/s. Uvedené hodnoty rozměrů a datového toku nebyly sice maximální možné, ale ve standardu MPEG-1 byly považovány za optimální. Standard MPEG-2 byl dokončen v roce 1994. Tento standard je koncipován mnohem velkoryseji a snaží se být standardem co nejuniverzálnějším. Zavádí několik tzv. profilů a úrovní. Profil označuje jistou podmnožinu z nejširší možné syntaxe. Úroveň definuje parametry v rámci jednoho profilu. Příklady jsou uvedeny v tabulce 7.2 (oproti originální tabulce uvedené v normě MPEG je tabulka poněkud zjednodušena). Standardy MPEG-1, MPEG-2 jsou si přístupem k řešení komprese navzájem velmi podobné. Rozdíly mezi nimi jsou mnohdy pod rozlišovací úrovní tohoto textu. Od metod použitých ve standardech MPEG-1,2 se dosti

Profil Úroveň	Simple	Main	High
Low		4:2:0 352×288 4 Mb/s I,P,B	
Main	4:2:0 720×576 15 Mb/s I,P	4:2:0 720×576 15 Mb/s I,P,B	4:2:0, 4:2:2 720×576 20 Mb/s I,P,B
High		4:2:0 1920×1152 80 Mb/s I,P,B	4:2:0, 4:2:2 1920×1152 100 Mb/s I,P,B

Tab.7.2. MPEG-2: profily a úrovně.

odlišují metody navrhované standardem MPEG-4, který je určen pro extrémně nízké datové toky (méně než 64kb/s). Standardem MPEG-4 se v tomto textu podrobněji zabývat nebudeme. Pro úplnost poznamenejme, že práce na standardu označovaném jako MPEG-3 byly zastaveny. Tento standard měl původně sloužit pro HDTV, ale v průběhu času se ukázalo, že potřeby stačí pokrýt standard MPEG-2.

Standard MPEG používá tři typů rámců I,P,B. Rámce typu I jsou kódovány každý zvlášť, bez vazby na rámce předcházející či následující. Princip kódování rámců I je shodný jako u standardu JPEG, i když v detailech existují některé odlišnosti. Jiná je např. vzorová kvantovací tabulka (tab. 7.3), poněkud jiná je struktura a způsob kódování intermediální sekvence (tab. 7.5). Dále lze např. při kódování podle normy MPEG-2 kromě cik-cak postupu dle obr. 7.2 zvolit ještě postup dle obr. 7.4. Měřítko kvantizační matice může být jiné pro každý makroblok. Změnou měřítka lze řídit tok dat. Mnoho aplikací totiž vyžaduje, aby byl tok dat přibližně konstantní. Aby nedošlo k přetečení nebo naopak k vyprázdňení bufferů, lze tok měnit buď změnou měřítka kvantizační matice nebo dokonce záměnou kvantizační matice. Přesný předpis pro řešení uvedené situace však norma nepředepisuje. Konkrétní řešení je ponecháno na tvůrci enkodéru.

8	16	19	22	26	27	29	34
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	38	46	56	69
27	29	35	38	46	56	69	83

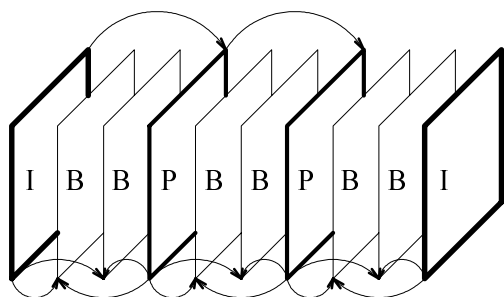
Tab.7.3. Vzorová kvantovací tabulka pro složku Y a rámce I v MPEG.

8	4	2	1	0	0	0	0
4	2	1	0	0	0	0	0
2	1	0	0	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Tab.7.4. Příklad bloku po DCT a kvantizaci.

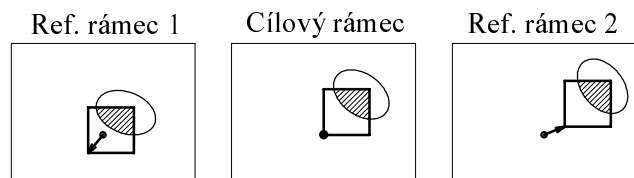
Počet nul před	Amplituda	MPEG kódování
-	8	110 1000
0	4	0000 1100
0	4	0000 1100
0	2	0 1000
0	2	0 1000
0	2	0 1000
0	1	110
0	1	110
0	1	110
0	1	110
12	1	0 0100 0100
EOB	EOB	10

Tab.7.5. Kódování bloku z tab. 7.4.



Obr.7.5. Sekvence IBBPBBPBB... v proudu MPEG rámců.

vaném) rámci jsou stanoveny vektory pohybu vzhledem k rámcům referenčním. Ke každému makrobloku v rámci P je tedy stanoven jeden vektor pohybu. Pro makroblok v rámci B jsou vektory pohybu dva. Vektor pohybu je definován takto: Jestliže o uvedený vektor posuneme kódovaný makroblok a porovnáme s odpovídající částí referenčního obrazu, pak je dosaženo dobré shody. Vektory posunutí se stávají součástí komprimované sekvence. Po nalezení vektoru posunutí jsou kódovány difference. Uvažujme nejprve rámce typu P. Nechť T označuje makroblok v cílovém rámci a R odpovídající oblast v rámci referenčním. Diferenci se rozumí rozdíl $T-R$. Jedná se o obraz obsahující čtyři bloky roviny Y a po jednom bloku od každé roviny barevné. Diference se kóduje stejným postupem, jako jsou kódovány makrobloky v rámcích typu I. Protože lze očekávat, že difference $T-R$ bude malá, vyjde po



Vektory pohybu:

Diference:

$$T - \frac{1}{2} [R_1 + R_2]$$

Obr.7.6. Kódování rámců typu B.

Kromě prostorové koherence (což je stejné jako u normy JPEG) využívá norma MPEG i koherence časové. Jinak řečeno: K dosažení vyšších stupňů komprese se předpokládá, že po sobě jdoucí rámce jsou s největší pravděpodobností dosti podobné. Počítá se ovšem s tím, že se jednotlivé části obrazu mohly přemístit a pozměnit. Využití časové koherence je dosaženo vložením rámců typu P a B.

Na rozdíl od rámců typu I nejsou rámce typu P a B kódovány nezávisle, nýbrž vzhledem k jednomu resp. dvěma jiným referenčním rámcům. Při kódování se využívá podobnosti rámce s rámci referenčními. Rámce P (predicted) jsou kódovány vzhledem k jedinému předcházejícímu rámci, kterým mohl být rámec typu I nebo P. Rámce typu B (interpolated bi-directionally) jsou kódovány vzhledem k nejbližšímu předchozímu a nejbližšímu budoucímu rámci typu I nebo P. Velmi často je např. používáno řazení rámců podle schématu IBBPBBPBB... (obr. 7.5). Použití rámců B je sice nepovinné, ale z hlediska dosahovaných kompresních poměrů výhodné. Typické hodnoty kompresních poměrů dosahované v rámcích typu I,P,B se totiž pohybují okolo hodnot 7:1, 20:1, 50:1. Komplikací vyplývající z použití rámců B je nutnost uchovávat v paměti dva kotevní obrazy. Dále je nevyhnutelné jisté časové zpoždění, protože nejprve musí být k dispozici obraz novější a teprve potom může být kódován obraz starší.

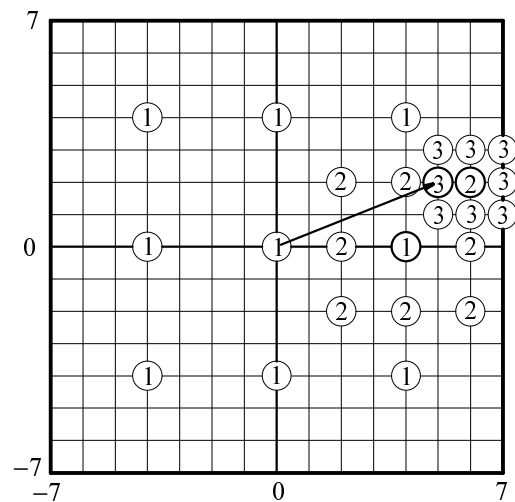
Také kódování rámců typu P a B probíhá po makroblocích. Pro každý makroblok v cílovém (tj. právě kódovaném) rámci jsou stanoveny vektory pohybu vzhledem k rámcům referenčním. Ke každému makrobloku v rámci P je tedy stanoven jeden vektor pohybu. Pro makroblok v rámci B jsou vektory pohybu dva. Vektor pohybu je definován takto: Jestliže o uvedený vektor posuneme kódovaný makroblok a porovnáme s odpovídající částí referenčního obrazu, pak je dosaženo dobré shody. Vektory posunutí se stávají součástí komprimované sekvence. Po nalezení vektoru posunutí jsou kódovány difference. Uvažujme nejprve rámce typu P. Nechť T označuje makroblok v cílovém rámci a R odpovídající oblast v rámci referenčním. Diferenci se rozumí rozdíl $T-R$. Jedná se o obraz obsahující čtyři bloky roviny Y a po jednom bloku od každé roviny barevné. Diference se kóduje stejným postupem, jako jsou kódovány makrobloky v rámcích typu I. Protože lze očekávat, že difference $T-R$ bude malá, vyjde po kosinové transformaci a kvantizaci v každém bloku většina členů nulových. Bloky tak ke svému zakódování vyžadují pouze krátké sekvence. Jako příklad kvantizační matice pro kódování diferencí uvádí norma MPEG matici, jejíž všechny prvky jsou rovny hodnotě 16. Kódování diferencí v rámcích typu B se provádí obdobně. Nechť T je opět makroblok v cílovém rámci a R_1, R_2 nechť jsou odpovídající oblasti v rámcích referenčních. V rámcích typu B se kóduje difference $T-0.5(R_1+R_2)$.

Postup při stanovení pohybového vektoru norma MPEG nepředepisuje. Při implementaci enkodéru se však ale jedná o jeden z nejobtížnějších problémů. Možný postup řešení je následující: K určení vektoru se využívá pouze jasové složky Y , barevné informace se nevyužívá. Necht' i, j jsou souřadnice levého dolního rohu kódovaného makrobloku v cílovém rámci a N necht' je velikost makrobloku ($N=16$). Pohybový vektor se stanoví tak, že se v referenčním rámci nalezne oblast velikosti $N \times N$, která se s makroblokem dobře shoduje. Předpokládejme, že souřadnice levého dolního rohu kandidující oblasti jsou $i+k, j+l$ a že $Y_T(r,s)$, $Y_R(r,s)$ označují jasovou složku cílového resp. referenčního rámce. Míru shody mezi makroblokem a kandidující oblastí lze měřit pomocí vztahu

$$\varepsilon(k, l) = \frac{1}{N^2} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} [Y_T(i+m, j+n) - Y_R(i+k+m, j+l+n)]^2. \quad (7.8)$$

Prakticky se hodnoty k, l uvažují z intervalu přibližně $-7 \leq k, l \leq 7$ až $-31 \leq k, l \leq 31$. Nízká hodnota $\varepsilon(k, l)$ signalizuje dobrou shodu; vektor (k, l) je pak v takovém případě pohybovým vektorem. Předpokládejme, že $-Q \leq k, l \leq Q$, kde Q je zvolená konstanta. Hledání nejnižší hodnoty $\varepsilon(k, l)$ systematickým prověřováním všech možných dvojic (k, l) by vedlo k neúnosně vysoké časové složitosti, protože vztah (7.8) by bylo nutné vyhodnocovat $(2Q+1)^2$ -krát. Systematického prohledávání se proto prakticky nepoužívá.

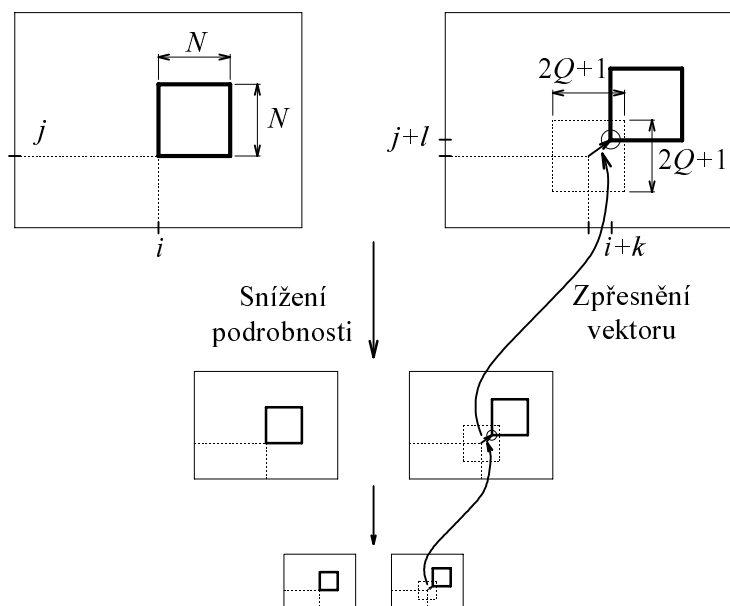
Jedním z postupů vykazujících přijatelnou časovou složitost je vyhledávání logaritmické. Obr. 7.7 znázorňuje jednu z jeho možných variant. V prvním kroku algoritmus testuje devět dvojic (k, l) . V každém dalším kroku se testuje vždy po osmi dvojicích. Na obr. 7.7 jsou dvojice znázorněny kroužky. Vepsané číslo udává pořadové číslo kroku, v němž je dvojice testována. Vzdálenost mezi body testovanými v jistém kroku je vždy poloviční ve srovnání se vzdáleností bodů testovaných v kroku předchozím. Po otestování bodů se vzdáleností 1 algoritmus končí. Při logaritmickém vyhledávání je vztah (7.8) zapotřebí vyhodnocovat méně než $(9 \log_2 Q)$ -krát, což je podstatné urychlení. Nevýhodou ovšem je, že nalezený pohybový vektor nemusí být nejlepší možný.



Obr. 7.7. Stanovení pohybového vektoru logaritmickým vyhledáváním.

Jinou možnou metodou hledání vektoru pohybu je využití postupného snižování podrobnosti obrazu. Převzorkováním referenčního i cílového obrazu vytvoříme obrazy nižší podrobnosti, které v obou směrech obsahují pouze poloviční počet vzorků. Tento krok dále rekurzivně opakujeme, takže dostaneme obrazy, v nichž je počet vzorků v obou směrech pouze čtvrtinový atd. Označme p počet rekurzivních opakování uvedeného kroku (typická hodnota je $p=2$). Hledání pohybového vektoru začíná stanovením odhadu vektoru v nejméně podrobném obraze. Protože v tomto obraze je velikost makrobloku $N/2^p$ a velikost prohledávaného okolí je $Q/2^p$, je časová složitost získání odhadu nízká. Necht' (k, l) je odhad pohybového vektoru získaný s využitím obrazů nejnižší podrobnosti. Tento odhad dále zpřesníme s využitím obrazů vyšší podrobnosti tak, že prověřujeme dvojice v jistém okolí hodnoty $(2k, 2l)$. Dvojice, která dává nejnižší hodnotu ε , je novým zpřesněným odhadem. Postup se dále opakuje s obrazy vyšší podrobnosti. Poslední zpřesnění se provede s využitím původních obrazů.

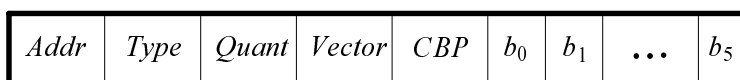
Kódování složek vektoru pohybu se provádí inkrementálně vzhledem k předchozí hodnotě vektoru pro tentýž makroblok. Hodnoty jsou kódovány jako celá čísla s proměnnou délkou. Poznamenejme, že je samozřejmě také možné, že vyhovující pohybový vektor nemusí být nalezen, ba dokonce vůbec ani nemusí existovat (termín vyhovující norma ovšem nijak podrobněji nespecifikuje; rozhodnutí je ponecháno na tvůrci enkodéru). Tento případ je ošetřen tak, že i v rámci typu P a B se mohou vyskytovat



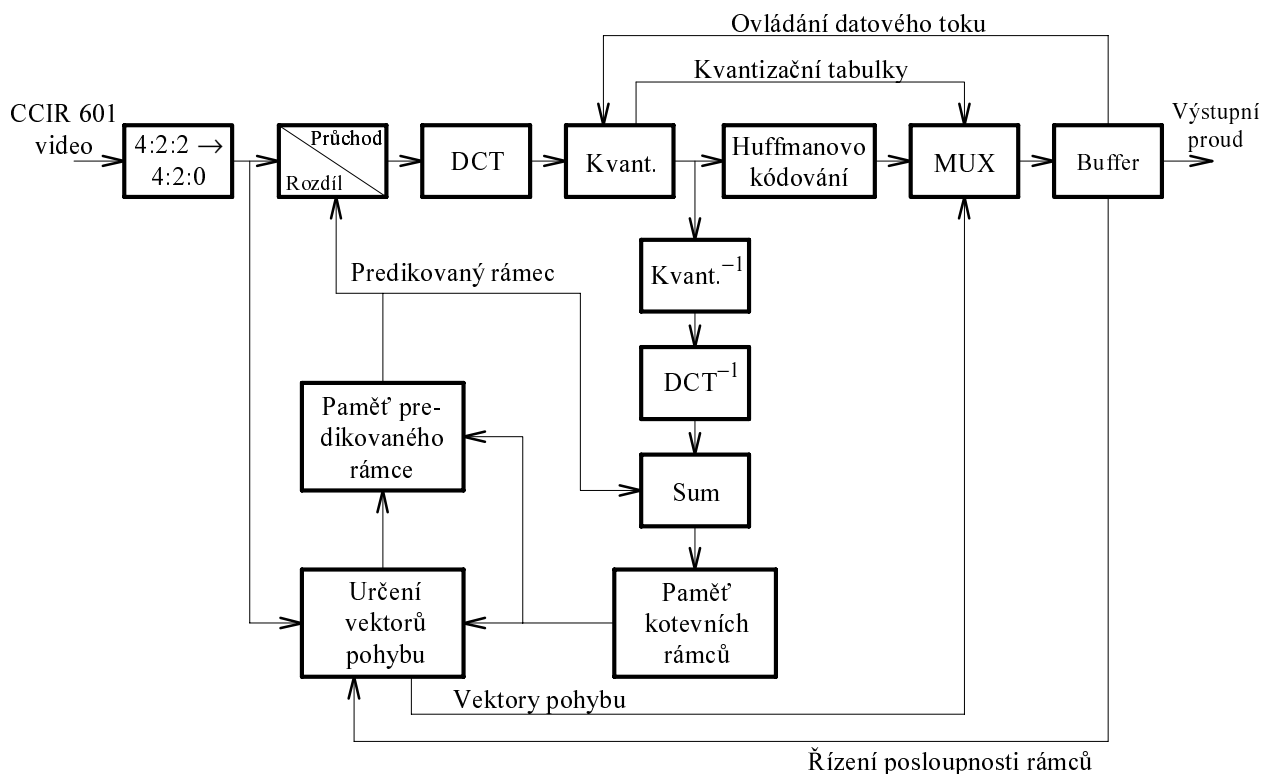
Obr. 7.8. Hledání vektoru pohybu s využitím obrazů s nižší podrobností.

Vector. V poli *CBP* je přenášena bitová maska indikující, které bloky (z šestice $Y_1, Y_2, Y_3, Y_4, C_r, C_b$ bloků možných) jsou v záznamu přenášeny. Pole b_0 až b_5 označují přenášené bloky. Na obr. 7.10 je uvedeno schéma MPEG enkodéru.

makrobloky kódované nezávisle jako makrobloky v rámcích typu I. Poučný může být poněkud podrobnější pohled na způsob kódování makrobloku (obr. 7.9). V poli *Addr* je přenášena informace o poloze bloku v obraze. Pokud se některé části obrazu v čase nemění, nemusí být odpovídající makrobloky kódovány a přenášeny. Přenáší se pouze makrobloky, které se změnily. V poli *Type* se specifikuje typ makrobloku. Pokud např. nebyl nalezen pohybový vektor, může se v tomto poli objevit příznak INTRA indikující, že je makroblok kódován nezávisle, stejně jako makrobloky v rámcích typu I. V poli *Quant* se přenáší měřítko kvantizační matice. Pohybový vektor je přenášen v poli



Obr. 7.9. Kódování makrobloku.

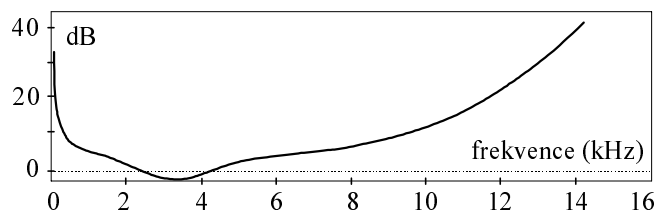


Obr. 7.10. Schéma MPEG enkodéru obrazu.

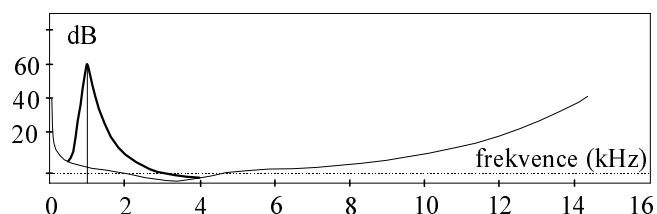
7.4 MPEG komprese zvuku

Kromě kódování obrazových sekvencí řeší norma MPEG také problém komprese a kódování zvukového doprovodu. Norma MPEG-1 předpokládá použití vzorkovacích frekvencí 32, 44.1, 48 kHz a jednoho nebo dvou zvukových kanálů. Norma zavádí tři úrovně komprese: Úroveň I předpokládá cílový datový tok větší než 128 kb/s (typicky např. 192 kb/s), úroveň II předpokládá datový tok okolo 128 kb/s a úroveň III datový tok okolo 64 kb/s. Ve všech případech je zajištěna vysoká věrnost zvuku po jeho dekódování. Vyšším kompresním poměrům dosahovaným v úrovni II a III odpovídá i vyšší složitost komprimačních postupů. Normou MPEG-2 byla zavedena další vylepšení. Jedná se např. o podporu až pěti zvukových kanálů, doplnění podpory nižších cílových datových toků až do 8 kb/s a doplnění podpory nižších vzorkovacích frekvencí 16, 22.05, 24 kHz). V tomto textu se omezíme zejména na postupy použité v nejjednodušší úrovni I normy MPEG-1, a to při komprimaci jediného zvukového kanálu. Všechny úrovně standardu MPEG využívají ke kompresi zvuku nedokonalosti lidského sluchu. Komprese je dosahována tak, že nejsou přenášeny informace, které lidský sluch stejně nemůže rozlišit. Využity jsou zejména následující psychoakustické jevy: 1) proměnná citlivost lidského sluchu v závislosti na frekvenci, 2) frekvenční maskování, 3) temporální maskování.

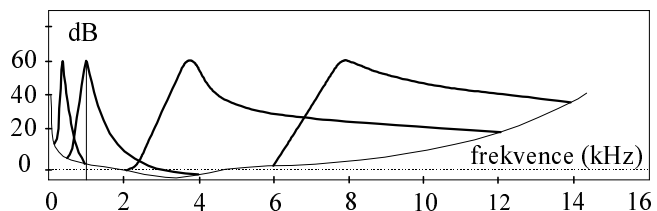
Křivka prahové citlivosti lidského sluchu v tichu v závislosti na frekvenci je na obr. 7.11. Lze pozorovat, že lidský sluch je nejcitlivější v oblasti 2 až 4 kHz. Směrem k vyšším i hlubším tónům se citlivost snižuje. Současně se snižuje také schopnost lidského sluchu rozlišit intenzitu takových tónů. Při kom-



Obr.7.11. Křivka citlivosti lidského ucha v tichu.



Obr.7.12. Maskování tónem 1kHz, 60 dB.



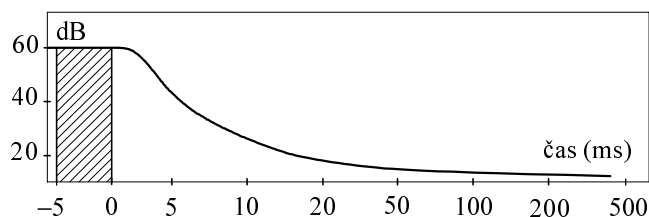
Obr.7.13. Maskování tóny 0.25, 1, 4, 8 kHz, 60 dB.

presi lze tohoto jevu prakticky využít tak, že jsou tóny hlubokých i vysokých kmitočtů kvantovány s větší kvantovací chybou, kterou však lidský sluch nezaznamená. Tóny, jejichž intenzita je nižší než prah citlivosti, nejsou přenášeny vůbec.

Frekvenční maskování je ilustrováno na obr. 7.12. Je zde znázorněna situace, kdy v prostoru zní tón o frekvenci 1 kHz a intenzitě 60 dB. Schopnost lidského sluchu vnímat jiné slabší tóny blízké frekvence a rozlišovat jejich úroveň je podstatně omezena. Obr. 7.12 ukazuje křivku citlivosti lidského sluchu za této situace. Tóny pod touto křivkou lidský sluch nevnímá - říkáme, že jsou maskovány. Praktickým využitím tohoto jevu opět je, že slabší tóny v blízkosti tónů silných nemusí být buď přenášeny vůbec nebo mohou být kódovány se značnou kvantizační chybou. Na obr. 7.13 jsou znázorněny oblasti, které jsou maskovány tóny o kmitočtu 250 Hz, 1, 4 a 8 kHz.

Jev temporálního maskování je ilustrován na obr. 7.14. V čase od -5 do 0 ms zní tón o intenzitě 60 dB. Po doznění tónu v čase 0 nenabude lidský sluch původní citlivosti

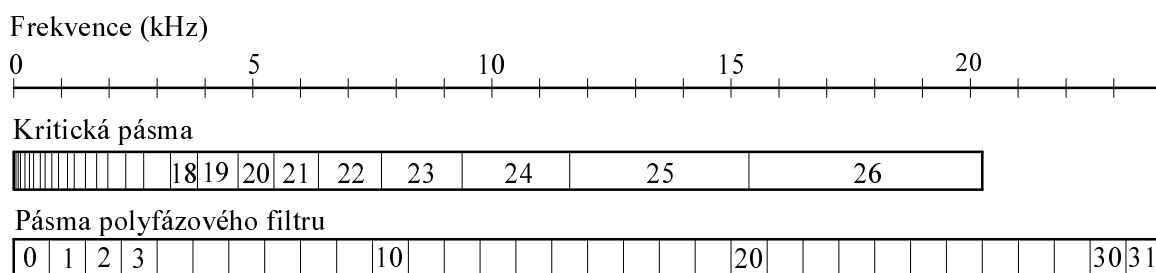
ihned, nýbrž s jistým časovým zpožděním, během něhož citlivost ucha postupně roste. I v této době po doznění tónu může být kvantizační chyba vyšší. Temporálního maskování je využito pouze v úrovni II a III; v úrovni I není využito.



Obr.7.14. Maskování v čase.

Při kompresi zvukového signálu se dále bere v úvahu citlivost sluchu na změny frekvence. Tuto vlastnost lze popsat rozdělením slyšitelného spektra do tzv. kritických pásem. I když je šířka měřená v Hz u jednotlivých kritických pásem různá, vnímá lidský sluch změnu ze spodní na horní hraniční frekvenci v každém kritickém pásmu jako skok stejné velikosti. Situaci ilustruje obr. 7.15. Nahoře je uvedena

stupnice v kHz, pod ní je znázorněno rozdělení do kritických pásem, která lidský sluch vnímá jako pásma stejné šířky. S ohledem na vnímání frekvencí lidským sluchem je zavedena jednotka frekvence 1 bark. Kritická pásma mají šířku právě 1 bark.

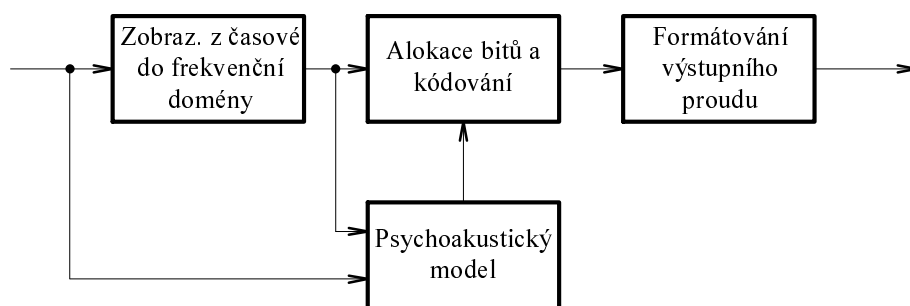


Obr.7.15. Dělení slyšitelného spektra na kritická pásma a pásma polyfázového frekvenčního filtru.

Jednotlivé kroky MPEG komprese zvukového signálu jsou znázorněny na obr. 7.16. V následujícím textu je popíšeme podrobněji. Základním krokem je rozdělení zvukového signálu do 32 frekvenčních pásem. Ačkoli to neodpovídá subjektivnímu vnímání frekvencí lidským sluchem, mají v tomto případě všechna frekvenční pásma stejnou šířku (obr. 7.15). Rozdělení se realizuje digitálním polyfázovým filtrem. Filtr zpracovává vždy 512 vzorků zvukového signálu, které jsou uchovávány v zásobníku FIFO. Vždy, když se do zásobníku nasune 32 nových vzorků, provede se výpočet produkující po jednom vzorku v každém ze 32 frekvenčních pásem. Filtr je normou MPEG předepsán vztahem

$$s(i) = \sum_{k=0}^{63} \sum_{j=0}^7 \cos \left[\frac{\pi(2i+1)(k-16)}{64} \right] \left[C(k+64j)x(k+64j) \right]. \quad (7.9)$$

V uvedeném vztahu označuje $x(k)$ ($k=0,1,\dots, 511$) hodnoty v zásobníku vzorků vstupního signálu, $C(k)$ ($k=0,1,\dots, 511$) jsou váhové koeficienty definované normou a $s(i)$ ($i=0,1,\dots, 31$) je výstupní vzorek v i -tém frekvenčním pásmu. Hodnoty $s(i)$ se počítají při nezměněném obsahu zásobníku vstupních vzorků. Jakmile je všech 32 hodnot $s(i)$ vypočítáno, posune se obsah zásobníku o 32 pozic (nejstarších 32 hodnot je ztraceno) a do prostoru uvolněného na začátku zásobníku (indexy 0 až 31) se nasune 32 nových vzorků. Poznamenejme, že frekvenční charakteristika filtrů není ideální obdélníková, ale lichoběžníková s mírným překrýváním jednotlivých pásem.



Obr.7.16. Schéma MPEG enkodéru zvuku.

Důsledkem této okolnosti je, že i čistě sinusový signál může být rozdělen do dvou sousedních frekvenčních pásem.

Úkolem psychoakustického modelu je analyzovat vstupní signál a stanovit, jaká velikost kvantizačního šumu v jednotlivých frekvenčních pásmech je ještě přijatelná (lidským sluchem nepostřehnutelná). Tato informace je pak využívána enkodérem, který pro každé pásmo určuje, kolik bitů bude pro kódování vzorků pásma použito. Při realizaci psychoakustického modelu je dosti značná volnost. Ve standardu MPEG jsou uvedeny dva příklady (model 1,2). V tomto textu se omezíme pouze na popis modelu 1.

V modelu je nejprve pomocí Fourierovy transformace vstupní zvukový signál transformován do frekvenční domény. Transformace se provádí nad oknem o velikosti 512 (úroveň I) nebo 1024 (úroveň II, III) vzorků. K redukci efektů vyplývajících z použití okna omezené velikosti se využívá Hannovy váhové funkce (ještě před provedením transformace se hodnoty vzorků touto funkcí násobí). Výsledku polyfázového filtru, tj. rozdělení signálu do 32 frekvenčních pásem model nevyužívá, protože je zde zapotřebí jemnějšího rozdělení.

Činnost modelu 1 pokračuje extrakcí tonálních a šumových komponent. Tonální komponenty jsou detekovány jako významné špičky v energetickém spektru. Šumová komponenta je stanovována vždy pouze jedna pro každé kritické pásmo. Její intenzita se určí jako součet intenzit všech netonálních (nešpičkových) hodnot energetického spektra ve vyšetřovaném kritickém pásmu. Se šumovou komponentou se dále zachází podobně jako s komponentou tonální. Za frekvenci šumové komponenty je považován geometrický střed odpovídajícího kritického pásma. Pro zjednodušení dalšího výpočtu provádí model decimaci komponent. Při decimaci jsou vypuštěny všechny šumové komponenty, které mají intenzitu nižší než je práh slyšitelnosti v tichu. Dále jsou vypuštěny všechny tonální komponenty, které se vyskytují v blízkosti (blíže než 0.5 bark) silnější tonální komponenty. Komponenty zbylé po decimaci jsou použity k výpočtu maskovacích úrovní.

Také v psychoakustickém modelu jsou jednotlivé frekvence sdružovány do frekvenčních pásem. Na rozdíl od polyfázového filtru jsou zde ovšem šířky pásem navrženy tak, aby byly lidským sluchem vnímány jako stejné. Jedná se tedy o dělení odvozené od dělení na kritická pásma. Prakticky může být šířka pásma používaná v psychoakustickém modelu např. 1/4 až 1/2 bark. S využitím empiricky získaných křivek maskování (obr. 7.13) a empiricky získaných křivek citlivosti lidského ucha (obr. 7.11) je pro každé pásmo psychoakustického modelu stanovena maskovací úroveň (co je pod touto úrovní, člověk neslyší). Z maskovacích úrovní stanovených pro jednotlivá frekvenční pásma psychoakustického modelu je pak stanovena maskovací úroveň pro každé ze 32 pásem polyfázového filtru. Jistou komplikací je, že se frekvenční pásma psychoakustického modelu a polyfázového filtru neshodují (obr. 7.15). V modelu 1 se pro každé frekvenční pásmo polyfázového filtru maskovací úroveň stanovuje jako minimum z maskovacích úrovní těch pásem psychoakustického modelu, které do vyšetřovaného frekvenčního pásma filtru padnou.

V úrovni I jsou vzorky kódovány v rámcích po 384 vzorcích. Rámec obsahuje 12 vzorků od každého z 32 frekvenčních pásem polyfázového filtru. Kromě samotných vzorků obsahuje rámec ještě další doplňující informace (obr. 7.17). Všechny 12 vzorků jednoho pásma je kódováno na stejném počtu bitů. V poli *bit alloc* je uvedeno celkem 32 hodnot udávajících počet bitů, které jsou použity ke kódování vzorků v jednotlivých frekvenčních pásmech. Hodnota *bit alloc* = 0 znamená, že odpovídající frekvenční pásmo není vůbec přenášeno. Pro frekvenční pásma, v nichž je *bit alloc* ≠ 0, je v poli *scale factors* dále uvedena hodnota, kterou se násobí všechny vzorky pásma. Vzorky jsou přenášeny v poli *samples*. Pole *anc. data* obsahuje případná doplňující data. Počet bitů, které jsou ke kódování vzorků

<i>Header</i>	<i>CRC</i>	<i>Bit alloc</i>	<i>Scale factors</i>	<i>Samples</i>	<i>Anc. data</i>
---------------	------------	------------------	----------------------	----------------	------------------

Obr.7.17. Formát rámce v úrovni I.

v jednotlivých frekvenčních pásmech použity, se stanoví na základě poměru (úroveň signálu v pásmu / maskovací úroveň). Tato hodnota je pro každé pásmo vypočítávána v psychoakustickém modelu.

7.5 Fraktální komprese

Fraktální komprese byla intenzivněji studována zejména v posledním desetiletí. V běžné denní praxi se zatím v masovém měřítku neprosadila, a to pravděpodobně hlavně kvůli větší časové složitosti komprese. Provedení dekomprese je na druhé straně ovšem celkem jednoduché. Za jistých okolností je navíc možné snadno provádět zvětšování detailů obrazu (zoom). Protože se zdá, že vyhlídky fraktální komprese jsou slibné, zmíníme se o ní stručně i v tomto textu. V míře nezbytně nutné nejprve připomeneme matematické základy, o něž se fraktální komprese opírá.

Budeme pracovat s jistými množinami (prostory) prvků. Nechť \mathbf{X} je takový prostor. V \mathbf{X} dále zavedeme metriku. Metrika ke každým dvěma prvkům prostoru \mathbf{X} přiřazuje jejich vzdálenost (reálné číslo). Nechť d je metrika. Dvojici (\mathbf{X}, d) nazveme metrickým prostorem. Posloupnost prvků $\{\varphi_n\}_{n=1}^{\infty}$ nazveme posloupností Cauchyovskou, jestliže pro jakékoli číslo $\varepsilon > 0$ existuje $N > 0$ tak, že $d(\varphi_n, \varphi_m) < \varepsilon$ pro všechny dvojice $m, n > N$. Řekneme, že posloupnost $\{\varphi_n\}_{n=1}^{\infty}$ konverguje k prvku ϕ , jestliže pro jakékoli číslo $\varepsilon > 0$ existuje $N > 0$ tak, že $d(\varphi_n, \phi) < \varepsilon$ pro všechna $n > N$. Prvek ϕ nazveme limitou uvedené posloupnosti. Metrický prostor (\mathbf{X}, d) se nazývá úplným prostorem, jestliže limita každé Cauchyovské posloupnosti v \mathbf{X} je rovněž prvkem \mathbf{X} .

Operátor \mathcal{O} na metrickém prostoru (\mathbf{X}, d) se nazývá kontrahující, jestliže existuje konstanta $0 \leq s < 1$ (součinitel kontrakce) tak, že platí

$$d(\mathcal{O}(\varphi), \mathcal{O}(\psi)) \leq s d(\varphi, \psi) \quad \forall \varphi, \psi \in \mathbf{X}. \quad (7.10)$$

Prvek $\phi \in \mathbf{X}$, pro který platí $\mathcal{O}(\phi) = \phi$, nazveme pevným bodem operátoru \mathcal{O} . Pro akci $\mathcal{O}(\mathcal{O}(\mathcal{O}(\dots \mathcal{O}(\varphi) \dots)))$, kdy je operátor \mathcal{O} aplikován n -krát, zavedeme označení $\mathcal{O}^n(\varphi)$.

Pro fraktální kompresi má zásadní význam následující věta o pevném bodu kontrahujícího operátoru: Nechť \mathcal{O} je kontrahující operátor na úplném metrickém prostoru (\mathbf{X}, d) . Potom \mathcal{O} má právě jediný pevný bod $\phi \in \mathbf{X}$ a pro libovolný prvek $\psi \in \mathbf{X}$ posloupnost $\{\mathcal{O}^n(\psi) | n=0, 1, 2, \dots\}$ konverguje k ϕ . Je tedy

$$\lim_{n \rightarrow \infty} \mathcal{O}^n(\psi) = \phi \quad \forall \psi \in \mathbf{X}. \quad (7.11)$$

Je dosti časté, že se výklad fraktální komprese zahajuje případem binárních obrazů. Za tím účelem se zavádí pojmy Hausdorffův prostor a Hausdorffova dimenze. My zde však učiníme výjimku a začneme ihned s obrazy ve stupních šedi, což je případ, který je prakticky zajímavější (shodně by bylo možné postupovat též pro jednotlivé složky obrazů barevných). Předpokládejme, že obrazy ve stupních šedi jsou popsány pomocí jasové funkce, která je definována nad jistou oblastí - plochou obrazu. Nechť Ω značí tuto plochu a $\varphi(x, y)$ nechť je jasová funkce $\varphi: \Omega \rightarrow I$, kde I je nějaký interval reálných čísel (jasů). Prostor \mathbf{X} je množina všech jasových funkcí, které mohou být nad Ω definovány. Dále zavedeme konkrétní metriku. Nechť φ, ψ jsou prvky \mathbf{X} . Vzdálenost $d(\varphi, \psi)$ definujeme vztahem $d(\varphi, \psi) = \max \{|\varphi(x, y) - \psi(x, y)| \mid x, y \in \Omega\}$ (při použití této metriky lze nejsnáze dokazovat dále uváděná tvrzení). Obecně operátor \mathcal{O} zobrazuje každou jasovou funkci na jinou jasovou funkci a transformuje tak jeden obraz na obraz jiný. V případě kontrahujícího operátoru existuje podle dříve uvedeného teorému právě jediná jasová funkce $\phi(x, y)$, která je pevným bodem uvažovaného operátoru. Fraktální komprese je založena na myšlence, že je možné k zadané jasové funkci nalézt kontrahující operátor tak, aby jasová funkce byla jeho pevným bodem. Jasovou funkci pak není zapotřebí uchovávat, protože může být kdykoli jednoduše vypočítána podle vztahu (7.11). Aby bylo možné metodu označit za kompresi, je samozřejmě nutné předpokládat, že popis odpovídajícího operátoru si vyžádá méně prostoru než popis samotné jasové funkce. Při řešení praktických problémů vznikajících při fraktální kompresi a dekompresi je užitečné následující tvrzení. V případě tohoto tvrzení uvedeme i důkaz, protože se jedná o dobrou ilustraci dosud zavedených pojmů.

Nechť \mathcal{O} je kontrahující operátor na úplném metrickém prostoru (\mathbf{X}, d) , ϕ nechť je jeho pevný bod, s součinitel kontrakce a C nechť je takové reálné číslo, že pro každou dvojici $\varphi, \psi \in \mathbf{X}$ platí $d(\varphi, \psi) \leq C$ (C není menší než největší vzdálenost v \mathbf{X}). Nechť φ je nyní libovolný prvek \mathbf{X} . Pak platí

$$1) d(\mathcal{O}^n(\varphi), \phi) \leq s^n C, \quad 2) d(\varphi, \phi) \leq \frac{d(\varphi, \mathcal{O}(\varphi))}{1-s}. \quad (7.12)$$

Důkaz. Nejprve dokážeme první tvrzení. Prvek ϕ je pevným bodem operátoru. Pro každé i tedy platí $\mathcal{O}^i(\phi) = \phi$. Dosazením do prvního vztahu a uplatněním definice kontrahujícího operátoru dostaneme

$$d(\mathcal{O}^n(\varphi), \mathcal{O}^n(\phi)) \leq s d(\mathcal{O}^{n-1}(\varphi), \mathcal{O}^{n-1}(\phi)) \leq s^2 d(\mathcal{O}^{n-2}(\varphi), \mathcal{O}^{n-2}(\phi)) \leq \dots \leq s^n d(\varphi, \phi) \leq s^n C.$$

Při dokazování druhého vztahu zapíšeme pevný bod operátoru jako limitu $\mathcal{O}^n(\varphi)$ pro $n \rightarrow \infty$. Dále využijeme trojúhelníkové nerovnosti, (v tomto případě zobecněné na n sčítanců). Podobně jako v předchozím případě snadno dokážeme, že platí $d(\mathcal{O}^i(\varphi), \mathcal{O}^{i+1}(\varphi)) \leq s^i d(\varphi, \mathcal{O}(\varphi))$ a konečně využijeme vztahu $(1+s+s^2+s^3+\dots) = 1/(s-1)$. Postupně tak dostaneme

$$d(\varphi, \phi) = \lim_{n \rightarrow \infty} d(\varphi, \mathcal{O}^n(\varphi)) \leq \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} d(\mathcal{O}^i(\varphi), \mathcal{O}^{i+1}(\varphi)) \leq d(\varphi, \mathcal{O}(\varphi)) \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} s^i = \frac{d(\varphi, \mathcal{O}(\varphi))}{1-s}. \quad \bullet$$

Obě tvrzení uvedeného teorému mají praktický význam. Dekomprese obrazu je založena na výpočtu hodnoty $\mathcal{O}^n(\varphi)$. Prakticky však nebude možné volit n nekonečně velké, a hodnota $\mathcal{O}^n(\varphi)$ proto bude pouze jistou aproximací pevného bodu. První tvrzení podává návod pro odhad vzdálenosti mezi skutečným pevným bodem a touto aproximací. Druhého tvrzení lze využít při kompresi. Necht' φ je jaso-
vá funkce obrazu, který má být komprimován. Je zapotřebí nalézt operátor, jehož pevným bodem je funkce φ . Lze očekávat, že ani tento problém nebude v praxi možné řešit zcela přesně. Pravděpodobně postačí nalézt takový operátor, jehož pevný bod je dostatečně blízko funkci φ . Druhé tvrzení poskytuje odhad vzdálenosti pevného bodu ϕ vyšetřovaného operátoru \mathcal{O} od φ .

Dosud nezodpovězenou zůstává otázka, jak k zadané jasové funkci φ nalézt operátor \mathcal{O} , pro který je funkce φ pevným bodem. Rychlé a dostatečně přesné řešení tohoto problému je nejtěžším krokem fraktální komprese. Představme si nejprve, že bychom věděli, že se hledaný operátor vyskytuje v nějaké konečné množině $\{\mathcal{O}_i\}$ operátorů. Alespoň teoreticky by pak nalezení operátoru bylo snadné. Operátory by bylo možné prověřovat jeden po druhém a zjišťovat vzdálenost $d(\varphi, \mathcal{O}_i(\varphi))$. Pro hledaný operátor je vzdálenost $d(\varphi, \mathcal{O}_i(\varphi))$ nulová. V případě komprese obrazů ve stupních šedi hledáme operátor, který jasovou funkci zobrazuje na jinou jasovou funkci. Takových operátorů je zřejmě nespočetně mnoho. Aby byla úloha prakticky zvládnutelná a aby bylo možné uplatnit právě naznačenou metodu prohledávání, je nutné prostor, v němž operátor hledáme, redukovat. Představme si, že redukováný prostor operátorů je opět tvořen konečnou množinou $\{\mathcal{O}_i\}$ operátorů. Z této množiny vybereme ten operátor, jehož pevný bod se nejvíce blíží zadané jasové funkci φ . Podle druhého ze vztahů (7.12) k tomu postačí vyhodnocovat výraz $d(\varphi, \mathcal{O}_i(\varphi))/(1-s)$. Při redukci prověřovaného prostoru vzniká riziko, že hledaný operátor bude vynechán, a že následné řešení proto bude pouze přibližné. Samozřejmě je tak snaha, aby prověřovaný prostor zůstal co nejširší. Na druhé straně je ovšem nutné brát v úvahu také otázku časové složitosti. V dalším odstavci ukážeme jeden z možných postupů redukce.

Předpokládejme, že existuje pokrytí oblasti Ω (obr.7.18). Pokrytí obsahuje konečný počet množin $D_i \subset \Omega$, $i=1,2,\dots, M$ tak, že platí

$$\Omega = \bigcup_{i=1}^M D_i \quad \text{a} \quad D_i \cap D_j = \emptyset \quad \text{pro každé } i \neq j. \quad (7.13)$$

Pro každé i předpokládejme existenci zobrazení $f_i: D_i \rightarrow \Omega$ a zavedeme označení $f_i(D_i) = R_i$. Dále necht' v_i je kontrahující zobrazení na I (I je množina možných hodnot jasu) se součinitelem kontrakce $0 \leq s < 1$. Uvažujme operátor, který je definován předpisem

$$\mathcal{O}(\psi(x,y)) = v_i(\psi(f_i(x,y))) \quad \text{pro } (x,y) \in D_i. \quad (7.14)$$

Nechť \mathbf{X} je prostor jasových funkcí definovaných nad Ω a metrika necht' je definována vztahem $d(\varphi, \psi) = \max\{|\varphi(x,y) - \psi(x,y)| \mid x,y \in \Omega\}$. Pak platí následující tvrzení: 1) Prostor (\mathbf{X}, d) je úplný. 2) Operátor ze vztahu (7.14) je v tomto prostoru kontrahující.

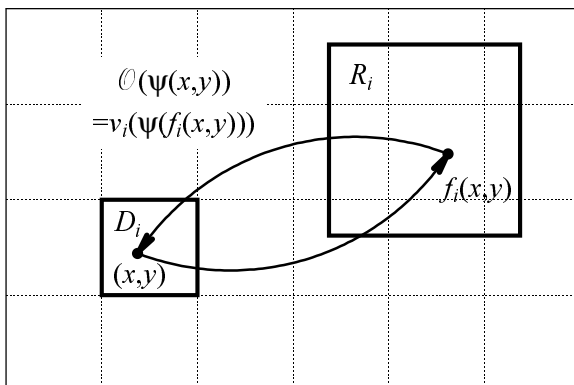
Vztah (7.14) tedy podává návod, jak konstruovat kontrahující operátor, kterého lze použít pro kompresi obrazů ve stupních šedi. Funkce v_i, f_i však stále bohužel nejsou nijak specifikovány, a proto je zatím ve volbě operátoru značná volnost. Při praktické realizaci fraktální komprese lze opět postupovat tak, že se i funkce v_i, f_i hledají na jistém dosti redukovaném prostoru. Následující příklad popisuje konkrétní řešení prezentované v (Barnsley 93).

Oblasti D_i mají tvar čtverce. Předpokládá se, že funkce f_i je afinní transformací. Třída afinních transformací je však dále zúžena, a to na transformace tvaru $\mathbf{x}' = 2\mathbf{A}_i(\mathbf{x} - \mathbf{c}_i) + \mathbf{t}_i$, kde $\mathbf{x} = (x,y)$, $\mathbf{c}_i = (c_i, d_i)$ jsou souřadnice středu oblasti D_i , $\mathbf{t}_i = (t_i, u_i)$ je translační složka transformace a \mathbf{A}_i je matice rozměru 2×2 z následující množiny

$$\mathcal{A} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \right\}. \quad (7.15)$$

Matici \mathbf{A}_i identifikujeme indexem $0 \leq a_i \leq 7$, což je pořadí matice ve výše uvedené množině. Jednotlivé matice realizují identitu, zrcadlení kolem svislé osy oblasti, zrcadlení kolem vodorovné osy, otočení o 180° kolem středu oblasti, zrcadlení kolem hlavní diagonály, otočení o 270° , otočení o 90° a zrcadlení kolem vedlejší diagonály. Je zřejmé, že obrazem R_i čtvercové oblasti D_i je oblast, která je opět čtvercová (obr. 7.18). Oproti čtverci D_i má čtverec R_i dvojnásobné rozměry. Předpokládá se, že $I = \langle 0, b_{\max} \rangle$, kde b_{\max} je maximální hodnota jasu v obraze. Funkce v_i se hledá ve tvaru $v_i(b) = pb + q_i$, kde p je pevně zvolená konstanta $0 \leq p < 1$ (tak je zajištěno, že funkce v_i je kontrahující). Hodnota q_i se stanoví tak, aby střední hodnota jasové funkce $\psi(x,y)$ nad oblastí D_i byla stejná jako střední hodnota jasové funkce $\mathcal{O}(\psi(x,y))$.

Po zavedení výše uvedených redukcí je operátor \mathcal{O} popsán M ticí čtveřic $[(t_1, u_1, a_1, q_1), (t_2, u_2, a_2, q_2), \dots, (t_M, u_M, a_M, q_M)]$ a hodnotou p , která je společná pro všechny oblasti. Je zřejmé, že čtveřice (t_i, u_i, a_i, q_i) popisuje, jak má být operátorem \mathcal{O} vytvořena obrazová funkce nad oblastí D_i . Tento dílčí operátor



Obr. 7.18. Konstrukce kontrahujícího operátoru pro obrazy ve stupních šedi.

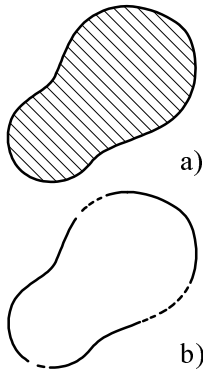
označíme \mathcal{O}_i . Operátor \mathcal{O} lze nalézt postupným probíráním všech oblastí D_i . Pro každou oblast D_i se nalezne operátor \mathcal{O}_i . To lze provést tak, že se postupně probírají všechny možné hodnoty t, u , pro které je splněna podmínka, že oblast $R_i = f_i(D_i)$ padne do Ω . Pro každou dvojici t, u , je stanovena hodnota q (podle pravidla uvedeného v předchozím odstavci je hodnota q stanovena pro každou dvojici t, u jednoznačně) a dále jsou prověřeny všechny možné matice \mathbf{A} . Jako čtveřice (t_i, u_i, a_i, q_i) popisující operátor \mathcal{O}_i se bere ta čtveřice hodnot t, u, a, q , která na oblasti D_i dává minimální hodnotu vzdálenosti $\text{dist}(\psi(x,y), \mathcal{O}(\psi(x,y)))$.

8 Segmentace obrazu

Předpokládejme, že máme obraz, který je popsán obrazovou funkcí, a že máme rozpoznat objekty v tomto obraze. Prvním krokem k tomuto cíli je extrakce objektů, které se v obraze nacházejí. Proces extrakce, v němž jsou objekty separovány od nezajímavého pozadí, nazýváme segmentací obrazu. Segmentace je nejčastěji založena na detekci kontur (hran) ohraničujících jednotlivé objekty nebo na detekci celých oblastí, kterými jsou jednotlivé objekty v obraze reprezentovány. Oba uvedené postupy jsou v praxi široce používány, a proto jsou jejich popisu věnovány poměrně rozsáhlé podkapitoly 8.1, 8.2, 8.3. V podkapitolách 8.4, 8.5 je pak stručně nastíněna problematika detekce rohů objektů a problematika zpracování binárních obrazů.

8.1 Detekce hran

Představujeme si, že každý objekt je v obraze reprezentován souvislou oblastí (obr.8.1a). Každá oblast je obklopena hranicí. Hranice se skládá z hran (případně též z jediné zakřivené hrany). Hrana se skládá z jednotlivých bodů. Významným problémem v analýze obrazů je nalezení hran a celých hranic. Při řešení se často postupuje tak, že se nejprve naleznou jednotlivé body hran. Uvažujme pro jedno-

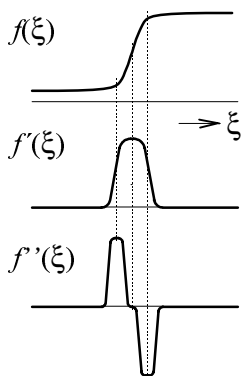


Obr.8.1. Detekce oblastí stanovením hranice.

duchost obrazu ve stupních šedi. Na existenci bodu hrany můžeme usuzovat z průběhu obrazové funkce. Na obr. 8.2 nahoře je znázorněn typický průběh jasu napříč hranou. Za bodu hrany se nejčastěji považuje místo, kde průběh jasu vykazuje náhlou změnu, případně inflexní bod. Jednotlivé nalezené body hran jsou pak různými technikami spojovány do hran a celých hranic. Protože každá plocha by měla být obklopena svojí hranicí (obr. 8.1), zdá se, že detekce hran a detekce celých ploch by teoreticky měly vést ke vzájemně ekvivalentním výsledkům. V praxi se však mohou vyskytnout problémy: Jsou-li např. rozpoznávané objekty v obraze špatně zřetelné nebo je-li obraz poškozen šumem, pak nemusí nalezené hrany tvořit kolem objektů uzavřené smyčky, jak si teoreticky představujeme. Vlivem šumu mohou být na druhé straně hrany detekovány i tam, kde ve skutečnosti nejsou.

8.1.1 Gradientní metody hledání hran

Obr. 8.2 ukazuje typický průběh jasu ve směru napříč hranou a první a druhou derivaci tohoto průběhu. Gradientní metody využívají skutečnosti, že v místě hrany má absolutní hodnota první derivace průběhu jasu vysokou hodnotu. Hodnota derivace popisuje intenzitu kontury v daném bodě - hovoříme o velikosti hrany. Operátory, které velikost hrany stanovují (obvykle se aplikují v každém bodě obrazu), bývají nazývány hranovými operátory. Z uvedeného vyplývá, že nejjednoduššími hranovými



Obr.8.2. Průběh jasu a jeho první a druhé derivace v místě hrany.

operátory jsou zřejmě derivace $\frac{\partial f}{\partial x}$ a $\frac{\partial f}{\partial y}$, které popisují změnu úrovně jasu ve směru os x a y . Těchto operátorů by bylo možné použít k hledání hran rovnoběžných se souřadnými osami. Při hledání hran obecného směru je zapotřebí vyšetřovat průběh jasu ve směru kolmém na směr potenciální hrany. Necht' vektor $\mathbf{n} = (\cos \theta, \sin \theta)$ popisuje směr kolmý ke směru eventuální hrany a ξ necht' je souřadnice měřená v tomto směru. K ověření skutečnosti, zda v daném bodě existuje hrana známého směru, lze použít derivaci ve směru. Pro derivaci ve směru kolmo k hraně máme

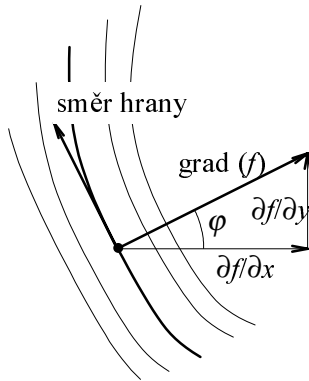
$$\frac{\partial f}{\partial \xi} = \text{grad}(f) \cdot \mathbf{n} = \frac{\partial f}{\partial x} \cos \theta + \frac{\partial f}{\partial y} \sin \theta. \quad (8.1)$$

Za velikost hrany v daném bodě můžeme považovat hodnotu $|\frac{\partial f}{\partial \xi}|$. Při hledání hrany její směr předem obvykle neznáme. Rovni-

ce (8.1) a úvahy, které k ní vedly, však podávají návod, jak situaci řešit. Pro rozhodnutí, zda ve vyšetřovaném bodě je či není hrana, bude rozhodující směr, v němž je změna jasu největší. Tímto směrem je gradient obrazové funkce. Směr hrany je pak kolmý ke směru gradientu. Za velikost hrany lze vzít velikost gradientu. Označme $e(x,y)$ velikost hrany, $\varphi(x,y)$ směr gradientu a $\psi(x,y)$ směr hrany v bodě (x,y) . Dále pro stručnost zavedme značení $f_x(x,y) = \partial f(x,y)/\partial x$, $f_y(x,y) = \partial f(x,y)/\partial y$. Pak máme:

$$e(x,y) = \sqrt{f_x^2(x,y) + f_y^2(x,y)}, \quad (8.2)$$

$$\varphi(x,y) = \arctan\left[\frac{f_y(x,y)}{f_x(x,y)}\right], \quad \psi(x,y) = \varphi(x,y) + \frac{\pi}{2}. \quad (8.3)$$



Obr.8.3. Stanovení velikosti a směru hrany na základě gradientu.

Cílem praktického výpočtu zpravidla bývá rozhodnout, zda vyšetřovaný bod obrazu je či není bodem ležícím na hranici nějakého objektu. V nejjednodušším případě lze za body hranice považovat ta místa, kde hodnota $e(x,y)$ je větší než nějaká pevná, předem zvolená prahová hodnota. Tento jednoduchý postup má ovšem své nedostatky: 1) Hranice objektu obvykle vyjde tlustší než jeden bod (intuitivně si možná představujeme, že tloušťka hranice je právě jeden bod). 2) Uvedený postup neřeší problém doplnění chybějících a odstranění nadbytečných částí hranice diskutovaný v úvodu podkapitoly 8.1. Důmyslnější postupy popíšeme postupně později (např. v souvislosti s Cannyho detektorem hran).

Až dosud jsme předpokládali, že obrazová funkce je spojitá. Při praktické implementaci pracujeme ovšem nejčastěji s funkcí diskrétní. Jestliže nahradíme derivace diferencemi, můžeme vztahů (8.2), (8.3) snadno použít i pro diskrétní případ. Pro difference můžeme psát

$$f_x(x,y) = f(x+1,y) - f(x,y), \quad (8.4)$$

$$f_y(x,y) = f(x,y+1) - f(x,y). \quad (8.5)$$

Při praktické realizaci výpočtu je hodnota $e(x,y)$ stanovována a porovnávána s hodnotou prahu ve všech bodech (pixelech) obrazu. Poznamenejme, že při stanovení délky gradientu lze použít i jiných metrik, než je metrika uvedená ve vztahu (8.2) (kořeny těchto snah sahají ovšem poněkud do minulosti, kdy byl výpočet odmocniny časově velmi náročnou, a proto obávanou operací). Můžeme použít např. následujících vztahů:

$$e(x,y) = |f_x(x,y)| + |f_y(x,y)|, \quad e(x,y) = \max\{|f_x(x,y)|, |f_y(x,y)|\}. \quad (8.6)$$

Skutečnost, že při praktickém výpočtu na použité metrice příliš nezáleží, ukazuje následující příklad, který ponecháváme čtenáři k samostatnému řešení. Příklad: Dokažte, že pro každou dvojici nezáporných čísel a,b platí $(a+b)/\sqrt{2} \leq \sqrt{a^2+b^2} \leq \sqrt{2} \max\{a,b\}$.

V minulosti bylo realizováno mnoho praktických metod pro detekci hran, jejichž společným teoretickým základem je výpočet velikosti gradientu. Některé z nich popíšeme dále podrobněji.

Robertsův operátor: Tento hranový operátor je jedním z nejstarších, ale stále ještě používaným operátorem pro stanovení velikosti hrany. Operátor stanovuje difference ve dvou na sebe kolmých diagonálních směrech. Velikost hrany v bodě (x,y) obrazu se počítá podle předpisu

$$e(x,y) = \sqrt{(f(x,y) - f(x+1,y+1))^2 + (f(x+1,y) - f(x,y+1))^2}. \quad (8.7)$$

Někdy také bývá uváděn vztah, v němž jsou místo funkčních hodnot použity jejich odmocniny. Tato varianta údajně více odpovídá vnímání hran lidským zrakem.

$$e(x, y) = \sqrt{(\sqrt{f(x, y)} - \sqrt{f(x+1, y+1)})^2 + (\sqrt{f(x+1, y)} - \sqrt{f(x, y+1)})^2}. \quad (8.8)$$

Operátor Prewittové: Tento operátor využívá k výpočtu velikosti hrany v bodě o souřadnicích x, y hodnot obrazové funkce ve všech sousedních pixelech. Tyto hodnoty pro stručnost označíme A, B, C, D, F, G, H, I (obr. 8.4). Operátor počítá derivace ve směrech os x, y průměrem

A	B	C
D	$f(x, y)$	F
G	H	I

Obr.8.4. Značení hodnot obrazové funkce pro operátor Prewittové.

$$f_x(x, y) = \frac{1}{3}[(C - A) + (F - D) + (I - G)],$$

$$f_y(x, y) = \frac{1}{3}[(A - G) + (B - H) + (C - I)]. \quad (8.9)$$

Uvedený postup stanovení derivací průměrem přispívá k redukci šumu. V případech, kdy nezáleží na absolutní velikosti hrany, lze koeficienty $1/3$ vypustit. Jestliže po výpočtu velikosti hrany následuje prahování, v němž se rozhoduje, zda v daném bodě hrana existuje či nikoli, lze vypuštění koeficientů kompenzovat volbou velikosti prahu.

Sobelův operátor: Tento operátor je podobný operátoru Prewittové. Rozdíl spočívá v tom, že se pro výpočet derivací v jednotlivých směrech používá vážený průměr dle vztahů

$$f_x(x, y) = \frac{1}{4}[(C - A) + 2(F - D) + (I - G)], \quad f_y(x, y) = \frac{1}{4}[(A - G) + 2(B - H) + (C - I)]. \quad (8.10)$$

I v tomto případě lze za stejných okolností jako u operátoru Prewittové koeficienty $1/4$ vypustit.

Kirschův operátor: Také tento operátor lze zařadit mezi operátory založené na výpočtu gradientu. Veličina mající povahu derivace je počítána celkem v osmi směrech a za velikost hrany je vzata její maximální hodnota. Pro stručnost přeznačíme funkční hodnoty obrazové funkce tak, jak je uvedeno na obr. 8.5. Velikost hrany v bodě (x, y) se vypočítá dle vzorce

A_0	A_1	A_2
A_7	$f(x, y)$	A_3
A_6	A_5	A_4

$$e(x, y) = \max_{i=0}^7 [5S_i - 3T_i], \quad (8.11)$$

$$\text{kde} \quad S_i = A_i + A_{i+1} + A_{i+2}, \quad (8.12)$$

$$T_i = A_{i+3} + A_{i+4} + A_{i+5} + A_{i+6} + A_{i+7}. \quad (8.13)$$

Obr.8.5. Značení hodnot obrazové funkce pro Kirschův operátor.

Index i , pro který je ve výrazu (8.11) dosaženo maxima, určuje směr hrany. Jestliže ve výrazech (8.12), (8.13) vyjdou hodnoty indexů větší než 7, upraví se operací modulo 8.

8.1.2 Detekce hran hledáním průchodu druhé derivace nulou

Z obr. 8.2 je zřejmé, že druhá derivace průběhu jasu ve směru napříč hranou nabývá dvou extrémů opačného znaménka a v místě hrany znaménko mění. Protože směr hrany není zpravidla předem znám, je nutné druhou derivaci počítat přinejmenším ve dvou směrech, např. ve směru souřadných os x, y . K výpočtu lze použít Laplaceova operátoru. Zaveďme označení $f_{xx}(x, y) = \partial^2 f(x, y) / \partial x^2$, $f_{yy}(x, y) = \partial^2 f(x, y) / \partial y^2$. Laplaceův operátor je definován vztahem

$$\nabla^2 f(x, y) = f_{xx}(x, y) + f_{yy}(x, y). \quad (8.14)$$

Je-li obrazová funkce diskretní, použijeme místo derivací diference. Pro body uvnitř obrazu (nikoli v krajních řádcích a sloupcích) můžeme použít symetrických diferenčních vzorců

$$f_{xx}(x, y) = f(x-1, y) - 2f(x, y) + f(x+1, y), \quad (8.15)$$

$$f_{yy}(x, y) = f(x, y-1) - 2f(x, y) + f(x, y+1). \quad (8.16)$$

Dosazením vztahů (8.15), (8.16) do (8.14) dostaneme pro realizaci Laplaceova operátoru předpis

$$\nabla^2 f(x, y) = f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) - 4f(x, y). \quad (8.17)$$

Vztah (8.17) lze vypočítat pomocí konvoluce obrazové funkce $f(x, y)$ s konvoluční maskou uvedenou na obr. 8.6a. V souvislosti s určováním hran pomocí druhé derivace bývá uváděna také maska dle obr.8.6b, která realizuje výpočet součtu druhých derivací ve čtyřech směrech $0^\circ, 45^\circ, 90^\circ, 135^\circ$. Je zřejmé, že v krajních řádcích a sloupcích obrazu nelze masky z obr. 8.6 použít. Pro uvedené případy lze v případě potřeby snadno odvodit speciální masky vyžadující pouze hodnoty uvnitř obrazu. Např. pro výpočet hodnoty Laplaceova operátoru v dolním řádku obrazu snadno odvodíme masku dle obr.

0	1	0
1	-4	1
0	1	0

a)

Obr.8.6. Konvoluční masky pro výpočet Laplaceova operátoru.

1	1	1
1	-8	1
1	1	1

b)

0	1	0
1	-3	1
0	0	0

a)

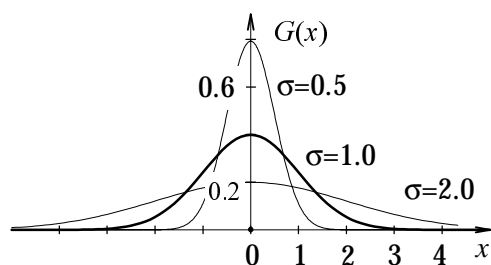
Obr.8.7. Příklady masek pro body na krajích a v rozích obrazu.

0	1	0
0	-2	1
0	0	0

b)

8.7a, pro výpočet hodnoty v levém dolním rohu pak masku uvedenou na obr. 8.7b (ostatní potřebné masky získáme rotací masek zde uvedených). Hrany v obraze nalezneme analýzou výsledku poskytnutého Laplaceovým operátorem. Za hranu považujeme místo, kde mezi dvěma dostatečně velikými a dostatečně blízko položenými extrémy opačného znaménka výsledek poskytnutý operátorem mění znaménko.

Právě popsaná metoda hledání hran založená na výpočtu druhé derivace a na použití Laplaceova operátoru je dosti citlivá na šum. I při malém zašumění obrazu je detekováno značné množství falešných hran. Problém lze řešit filtrací šumu, která se provede ještě před tím, než jsou hledány hrany. Marr a Hildreth (Maar 80) popsali metodu, v níž pro redukci šumu používají konvoluci s gaussianem. Protože se tato metoda používá dosti často, popíšeme ji podrobněji. Dvourozměrný gaussian je definován vztahem (obr.8.8)

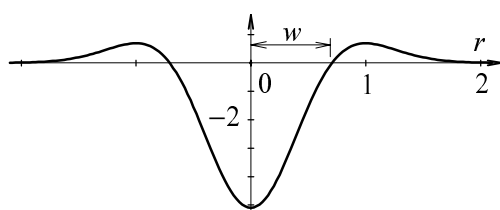
**Obr.8.8.** Průběh Gaussiánu $G(x)$ pro $\sigma = 0.5, \sigma = 1, \sigma = 2$.

$$G(x, y) = G(x)G(y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right). \quad (8.18)$$

Hodnotou σ lze „regulovat šířku“ gaussianu a tudíž i míru filtrace. Běžně používané hodnoty σ jsou 0.5 až 3. Realizaci konvoluce obrazové funkce s gaussianem a následnou aplikaci Laplaceova operátoru popisuje vztah

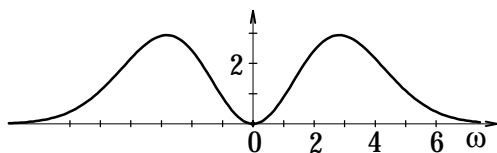
$$\begin{aligned} \nabla^2 [G(x, y) * f(x, y)] &= \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(x-\xi, y-\eta) f(\xi, \eta) d\xi d\eta \\ &= \left[\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) G(x, y) \right] * f(x, y) = [\nabla^2 G(x, y)] * f(x, y). \end{aligned} \quad (8.19)$$

Vztah (8.19) ukazuje, že popisovanou posloupnost akcí lze realizovat konvolucí obrazové funkce s funkcí $\nabla^2 G$. Položíme-li $r = \sqrt{x^2 + y^2}$, můžeme funkci $\nabla^2 G$ zapsat ve tvaru

**Obr.8.9.** Průběh funkce $\nabla^2 G(r)$, $\sigma = 0.5$.

$$\nabla^2 G(r) = \left(\frac{r^2 - 2\sigma^2}{2\pi\sigma^6} \right) \exp\left(-\frac{r^2}{2\sigma^2}\right). \quad (8.20)$$

Funkce $\nabla^2 G$ je rotačně symetrická. Její průběh (průřez průběhu) je znázorněn na obr. 8.9. Označme w poloměr kružnice, kde funkce přechází ze záporných do kladných hodnot (obr. 8.9). Z rovnice (8.20) vychází $w = \sigma\sqrt{2}$. K důkladnějšímu osvětlení principu činnosti



Obr.8.10. Frekvenční charakteristika $\mathcal{F}\{\nabla^2 G\}$ pro $\sigma = 0.5$.

této metody vyšetříme ještě frekvenční charakteristiku operátoru $\nabla^2 G$. Provedením Fourierovy transformace a položením $u^2+v^2 = \omega^2$ dostaneme

$$\mathcal{F}\{\nabla^2 G(x, y)\} = \omega^2 \exp\left(-\sigma^2 \frac{\omega^2}{2}\right). \quad (8.21)$$

Průběh frekvenční charakteristiky $\mathcal{F}\{\nabla^2 G\}$ je znázorněn na obr.8.10. Je zřejmé, že se jedná o pásmový filtr.

Vrchol frekvenční charakteristiky vychází v místě $\omega = \sqrt{2}/\sigma$. Po provedení konvoluce $[\nabla^2 G(x,y)] * f(x,y)$ pokračuje metoda hledáním průchodů nulou stejně, jak již bylo popsáno dříve.

8.1.3 Parametrické modely hrany

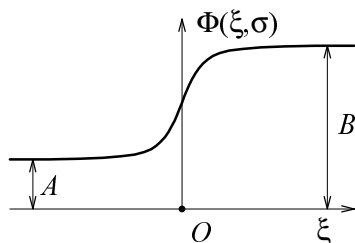
Řekněme, že máme rozhodnout, zda je v obraze v místě o souřadnicích (x,y) hrana. V pozitivním případě máme stanovit také její velikost a směr. Můžeme to provést tak, že jistým počtem funkčních hodnot obrazové funkce $f(x,y)$ z okolí bodu (x,y) proložíme vhodnou plochu, která průběh obrazové funkce aproximuje. O existenci hrany pak rozhodneme a její parametry stanovíme na základě vyšetřování průběhu aproximující funkce. Aby se potlačil vliv šumu, prokládá se aproximační funkce zpravidla větším počtem funkčních hodnot obrazové funkce, než je nejnižší teoreticky nutný počet. Proložení lze realizovat minimalizací součtu čtverců diferencí funkčních hodnot aproximační funkce od funkčních hodnot obrazové funkce ve zvolených bodech. Nejjednodušší plochou, kterou lze jako modelu hrany použít, je rovina $z = ax + by + c$. Nalezneme-li rovinu aproximující průběh obrazové funkce v daném místě, pak snadno získáme velikost hrany $e(x,y) = |\text{grad}(ax + by + c)| = \sqrt{(a^2 + b^2)}$ a také její směr $\varphi(x,y) = \text{arctg}(b/a) + \pi/2$. Rovinu můžeme nalézt např. tak, aby co nejlépe aproximovala průběh obrazové funkce ve čtyřech bodech $f(x,y)$, $f(x+1,y)$, $f(x,y+1)$ a $f(x+1,y+1)$. Neznámé hodnoty a,b,c nalezneme minimalizací chyby $\varepsilon = [ax + by + c - f(x,y)]^2 + [a(x+1) + by + c - f(x+1,y)]^2 + [ax + b(y+1) + c - f(x,y+1)]^2 + [a(x+1) + b(y+1) + c - f(x+1,y+1)]^2$ uplatněním podmínek $\partial\varepsilon/\partial a = 0$, $\partial\varepsilon/\partial b = 0$, $\partial\varepsilon/\partial c = 0$.

Dále ukážeme případ, kdy je jako model hrany použita důmyslnější plocha. Předpokládejme např., že v řezu napříč hranou má jas průběh, který je dobře vystižen funkcí

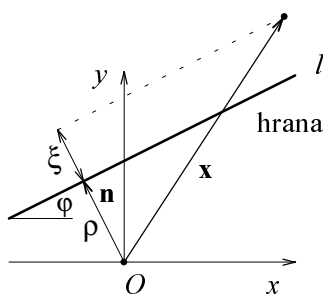
$$B + (A - B)\Phi(\xi, \sigma). \quad (8.22)$$

Význam parametrů A,B objasňuje obr. 8.11. Funkce $\Phi(\xi, \sigma)$ je integrálem gausiánu. Je tedy

$$\Phi(\xi, \sigma) = \int_{t=-\infty}^{\xi} G(t, \sigma) dt, \quad \text{kde } G(t, \sigma) = \exp\left(-\frac{t^2}{2\sigma^2}\right). \quad (8.23)$$



Obr. 8.11. Parametrický model hrany.



Obr. 8.12. Parametrický model hrany (pohled shora).

Abychom přešli k dvojrozměrnému případu, uvažme, že $\mathbf{n}\mathbf{x} = \rho$, kde $\mathbf{n} = (-\sin \varphi, \cos \varphi)$, $\mathbf{x} = (x,y)$, je rovnice přímky l (obr. 8.12). Jestliže bod reprezentovaný vektorem \mathbf{x} na přímce l neleží, pak vzdálenost bodu od přímky je $\xi = \mathbf{n}\mathbf{x} - \rho = -x \sin \varphi + y \cos \varphi - \rho$. Předpokládáme-li, že hrana leží na přímce l , pak lze průběh jasu v okolí bodu hrany dobře aproximovat funkcí

$$B + (A - B)\Phi(-x \sin \varphi + y \cos \varphi - \rho, \sigma). \quad (8.24)$$

Celkem má tedy popsaný model pět parametrů A,B,σ,φ,ρ . Hodnota $(A-B)$ určuje velikost (výšku) hrany, parametr σ její strmost a parametr φ směr hrany v bodě \mathbf{x} . Parametry lze určit na základě pod-

mínky, aby se v jistém počtu bodů (pixelů) ležících v okolí bodu x hodnoty funkce (8.24) co nejvíce přimykaly skutečným hodnotám obrazové funkce. Nejmenší počet bodů umožňující úlohu řešit je pět. S ohledem na redukci šumu je však výhodnější volit počet bodů vyšší. Pro hledané parametry lze pak sestavit předeterminovaný systém, který je možné řešit minimalizací chyby. Jistou komplikací ovšem je, že se jedná o problém nelineární. K jeho řešení je proto zapotřebí použít numerických metod.

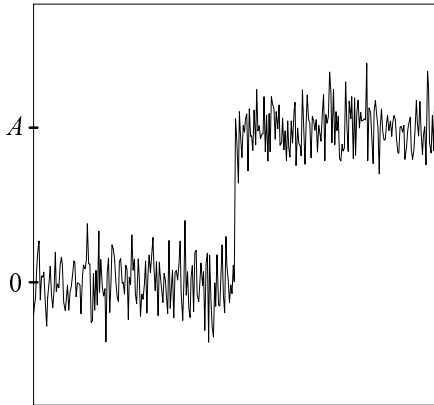
8.1.4 Cannyho detektor hran

Popisovaný detektor hran byl publikován Canny (Canny 83, 86) a později detailně studován mnoha dalšími autory. Při návrhu detektoru vycházel Canny z toho, že starší detektory hran byly konstruovány více či méně intuitivně, a formuloval proto požadavky, které by měl detektor splňovat. K nalezení detektoru pak přistoupil jako k optimalizační úloze (tj. hledal detektor, který vytyčené požadavky splňuje co nejlépe). Canny navržené požadavky byly následující: 1) minimalizovat pravděpodobnost chybné detekce, 2) najít polohu hrany v obraze co nejpřesněji, 3) bod hrany identifikovat jednoznačně. Postup, který vede k návrhu odpovídajícího filtru, popíšeme podrobněji.

Předpokládejme nejprve, že vstupní signál je jednorozměrný (získané výsledky později zobecníme i na signály dvojrozměrné). Matematickým modelem hrany ve vstupním signálu (ve shodě s původním pramenem označme vstupní signál $I(x)$) je skok výšky A umístěný do počátku. Doplněn je dále aditivní šum $n(x)$ (obr.8.13). Předpokládáme, že šum je bílý s nulovou střední hodnotou. Vstupní signál má tedy tvar

$$I(x) = Au_{-1}(x) + n(x), \text{ kde } u_{-1}(x) = \begin{cases} 0 & \text{pro } x < 0 \\ 1 & \text{pro } x \geq 0 \end{cases}. \quad (8.25)$$

Předpokládáme, že hranový operátor bude založen na výpočtu konvoluce vstupního signálu s funkcí $f(x)$. Tuto funkci, kterou zatím neznáme, stanovíme tak, aby byla co nejlépe splněna kriteria formulovaná v úvodu této podkapitoly. Označme $O(x_0)$ signál, který je výsledkem konvoluce. Podle definice konvoluce máme



Obr. 8.13. Průběh signálu ve směru napříč hranou.

$$O(x_0) = \int_{-\infty}^{+\infty} I(x_0 - x)f(x)dx. \quad (8.26)$$

Nejprve vyhodnotíme, jaký výstup v bodě $x_0 = 0$ odpovídá skoku velikosti A (bez šumu). Dosazením rovnice (8.25) do rovnice (8.26) a položením $n(x) = 0$, $x_0 = 0$ dostaneme

$$\int_{-\infty}^{+\infty} Au_{-1}(-x)f(x)dx = A \int_{-\infty}^0 f(x)dx. \quad (8.27)$$

Dále vypočítáme střední hodnotu čtverce signálu, který na svém výstupu poskytuje operátor v bodě $x_0 = 0$, jestliže je na vstup přiložen pouze šum $n(x)$. Dostaneme

$$E \left\{ \left[\int_{-\infty}^{\infty} n(-x)f(x)dx \right]^2 \right\} = \int_{-\infty}^{\infty} E \{ n^2(-x) \} f^2(x)dx = n_0^2 \int_{-\infty}^{\infty} f^2(x)dx. \quad (8.28)$$

Při úpravě výše uvedeného vztahu jsme se opírali o předpoklad, že $n(x)$ je bílý šum s nulovou střední hodnotou a že tedy pro všechna $x_1 \neq x_2$ platí $E \{ n(x_1)n(x_2) \} = 0$. Dále jsme předpokládali, že hodnota $E \{ n^2(x) \}$ je pro všechna x konstantní, a položili jsme proto $E \{ n^2(x) \} = n_0^2$ (jedná se o varianci šumové složky vstupního signálu). Definujme poměr signál/šum (SNR) jako podíl odezvy na skok velikosti A ku odmocnině střední hodnoty čtverce odezvy na šum. Pro hodnocení pravděpodobnosti chybné detekce dále zavedme míru $\Sigma = n_0 SNR / A$, která je závislá pouze na hledané funkci $f(x)$, nikoli na velikosti vstupního signálu. Máme tedy

$$SNR = \frac{A \int_{-\infty}^0 f(x) dx}{n_0 \sqrt{\int_{-\infty}^{\infty} f^2(x) dx}}, \quad \Sigma = \frac{n_0}{A} SNR = \frac{\int_{-\infty}^0 f(x) dx}{\sqrt{\int_{-\infty}^{\infty} f^2(x) dx}}. \quad (8.29)$$

Při detekci hrany právě konstruovaným hranovým operátorem budeme za hranu považovat místo, kde operátor (konvoluce) dává největší hodnotu. V případě hrany umístěné do počátku by to teoreticky mělo být místo $x_0 = 0$. Šum však způsobí, že tomu tak nebude zcela přesně. Nalezneme proto místo, kde výstup operátoru nabývá extrému. K tomu vypočítáme první derivaci výstupního signálu a položíme ji rovnu nule

$$\frac{d}{dx_0} O(x_0) = O'(x_0) = \frac{d}{dx_0} \int_{-\infty}^{\infty} I(x_0 - x) f(x) dx = 0. \quad (8.30)$$

Pro derivování konvoluce máme

$$\frac{d}{dx_0} \int_{-\infty}^{\infty} I(x_0 - x) f(x) dx = \frac{d}{dx_0} \int_{-\infty}^{\infty} I(x) f(x_0 - x) dx = \int_{-\infty}^{\infty} I(x) f'(x_0 - x) dx = \int_{-\infty}^{\infty} I(x_0 - x) f'(x) dx. \quad (8.31)$$

Výstup rozdělíme na složku odpovídající skoku umístěnému do počátku a na složku odpovídající šumu. Tyto složky označíme postupně O_s, O_n . Derivace odezvy na skok má v bodě x_0 velikost

$$O'_s(x_0) = \int_{-\infty}^{\infty} A u_{-1}(x_0 - x) f'(x) dx = A \int_{-\infty}^{x_0} f'(x) dx = A f(x_0). \quad (8.32)$$

Derivaci $O'_s(x_0)$ odezvy na skok rozvineme kolem počátku v Taylorovu řadu. Předpokládáme při tom, že hledaná funkce $f(x)$ bude antisymetrická, a že proto bude platit $f(0) = 0$. Dostaneme

$$O'_s(x_0) = A f(x_0) \approx A f(0) + x_0 A f'(0) = x_0 A f'(0). \quad (8.33)$$

Derivaci odezvy hranového operátoru na šum bude náhodná proměnná s nulovou střední hodnotou a s variancí

$$E\{O_n'^2(x_0)\} = E\left\{\left[\int_{-\infty}^{\infty} n(x_0 - x) f'(x) dx\right]^2\right\} = n_0^2 \int_{-\infty}^{\infty} f'^2(x) dx. \quad (8.34)$$

Pro podmínku extrému pak máme

$$O'(x_0) = O'_s(x_0) + O'_n(x_0) = 0 \quad \text{a tedy} \quad O'_s(x_0) = -O'_n(x_0). \quad (8.35)$$

Z předchozí rovnice dále plyne $E\{O_s'^2(x_0)\} = E\{O_n'^2(x_0)\}$. (8.36)

Dosažením vztahů (8.33) a (8.34) do rovnice (8.36) dostaneme:

$$E\{x_0^2\} = \frac{n_0^2 \int_{-\infty}^{\infty} f'^2(x) dx}{A^2 f'^2(0)} = \sigma_{x_0}^2. \quad (8.37)$$

V předchozím výrazu jsme označení σ_{x_0} použili pro aproximaci směrodatné odchylky polohy maxima na výstupu operátoru od skutečné polohy skoku, který leží v počátku. Schopnost operátoru správně lokalizovat hranu hodnotíme podle převrácené hodnoty této odchylky. Zavedeme míru $\Lambda = n_0/(\sigma_{x_0} A)$, která závisí pouze na vlastnostech hranového operátoru a nikoli na signálu. S využitím vztahu (8.37)

dostaneme
$$\Lambda = \frac{|f'(0)|}{\sqrt{\int_{-\infty}^{\infty} f'^2(x) dx}}. \quad (8.38)$$

Výkonnost hranového operátoru pak hodnotíme součinem $\Sigma \Lambda$. S využitím vztahů (8.29), (8.38) máme

$$\Sigma(f) \Lambda(f) = \frac{\int_{-\infty}^0 f(x) dx}{\sqrt{\int_{-\infty}^{\infty} f^2(x) dx}} \frac{|f'(0)|}{\sqrt{\int_{-\infty}^{\infty} f'^2(x) dx}}. \quad (8.39)$$

Poslední podmínkou je, aby jedné hraně odpovídalo pokud možno jediné maximum na výstupu operátoru. Tento požadavek vyjádříme jako požadavek na funkci $f(x)$. Využijeme pomocného tvrzení, které uvedeme bez důkazu a které říká, že střední vzdálenost x_{zc} mezi dvěma průchody nulou odezvy funkce g na gaussovský šum je dána výrazem (odezvou se rozumí konvoluce funkce g s šumem)

$$x_{zc} = \pi \left(\frac{-R(0)}{R''(0)} \right)^{\frac{1}{2}}, \quad (8.40)$$

kde $R(\tau)$ je autokorelační funkce funkce g . Snadno ověříme, že platí (předpokládáme $g(-\infty)=g(\infty)=0$)

$$R(0) = \int_{-\infty}^{\infty} g^2(x) dx, \quad R''(0) = \int_{-\infty}^{\infty} g(x)g''(x) dx = - \int_{-\infty}^{\infty} g'^2(x) dx. \quad (8.41)$$

V našem případě hledáme střední vzdálenost x_{zc} mezi dvěma průchody nulou odezvy funkce f' . Ze vztahů (8.40), (8.41) dostaneme

$$x_{zc} = \pi \left(\frac{\int_{-\infty}^{\infty} f'^2(x) dx}{\int_{-\infty}^{\infty} f''^2(x) dx} \right)^{\frac{1}{2}}. \quad (8.42)$$

Vzdálenost x_{\max} mezi dvěma přilehlými maximy odezvy funkce na šum je dvojnásobkem x_{zc} . Tuto hodnotu vyjádříme relativně vzhledem k šířce operátoru. Šířkou operátoru rozumíme interval $\langle -W, W \rangle$. Na tomto intervalu předpokládáme funkční hodnoty f nenulové (mimo tento interval předpokládáme funkční hodnoty nulové; dále připomeňme, že funkci f předpokládáme antisymetrickou). Je tedy

$$x_{\max} = 2x_{zc} = kW. \quad (8.43)$$

Schopnost detektoru identifikovat hranu jednoznačně bude tím lepší, čím vyšší bude hodnota k . Dosud neznámá funkce f by tedy měla maximalizovat výraz (8.39) a při tom dávat pokud možno co nejvyšší hodnotu k . V (Canny 83) je popsáno řešení založené na maximalizaci funkcionalu (8.39). Jsou postupně voleny různé hodnoty k (0.075 až 0.7) a rovnice (8.43) je použita jako doplňující podmínka. Pro každou hodnotu k je numerickou optimalizací stanovena funkce f . Na základě takto získaných výsledků je ukázáno, že pro vyšší hodnoty k (což je prakticky zajímavý případ) lze hledanou funkci f dosti dobře aproximovat pomocí první derivace gaussiánu. Na tuto možnost se dále zaměříme. Připomeňme, že gausián je definován předpisem

$$G(x) = \exp\left(-\frac{x^2}{2\sigma^2}\right). \quad (8.44)$$

Funkce f má tedy tvar
$$f(x) \approx G'(x) = -\frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right). \quad (8.45)$$

Pro výše uvedenou aproximaci funkce f vypočteme následující hodnoty (Canny 83)

$$|f'(0)| = \frac{1}{\sigma^2}, \quad \int_{-\infty}^0 f(x) dx = 1, \quad \int_{-\infty}^{\infty} f^2(x) dx = \frac{\sqrt{\pi}}{2\sigma}, \quad \int_{-\infty}^{+\infty} f'^2(x) dx = \frac{3\sqrt{\pi}}{4\sigma^3}, \quad \int_{-\infty}^{+\infty} f''^2(x) dx = \frac{15\sqrt{\pi}}{8\sigma^5}. \quad (8.46)$$

Odtud pak index $\Sigma\Lambda$ výkonnosti operátoru a hodnota k pro operátor $f(x) = G'(x)$ vychází

$$\Sigma\Lambda = \sqrt{\frac{8}{3\pi}} \approx 0.92 \quad \text{a} \quad k = \sqrt{\frac{4}{15}} \approx 0.51. \quad (8.47)$$

Zjištěné hodnoty jsou jen o málo menší než hodnoty pro operátor optimální. Podstatnou výhodou použití gausiánu je ale snadný výpočet.

Až doposud jsme se soustředili na jednorozměrný případ. Zbývá ukázat, jak pracuje Cannyho detektor v případě dvojrozměrném. Předpokládejme nejprve, že máme hledat hrany známého směru. Ve směru napříč hranou použijeme konvoluci s právě nalezenou funkcí f . Ve směru podél hrany uplatníme projekci. Jestliže je f aproximována derivací gausiánu, pak je výhodné, aby funkcí realizující projekci byl rovněž gausián (projekce bude počítána jako konvoluce s gausiánem). Jako obvykle není zapo-

třebí výpočet provádět pro všechny možné směry hran, ale postačí jej provést pro dva směry (např. x,y). Necht' $I(x,y)$ je obrazová funkce. Pro hrany rovnoběžné s osou y a hrany rovnoběžné s osou x pak dává hranový operátor následující hodnoty

$$[I(x,y)*G(y)]*G_x(x) = \frac{\partial}{\partial x} \{I(x,y)*[G(x)G(y)]\} = E_x(x,y), \quad (8.48)$$

$$[I(x,y)*G(x)]*G_y(y) = \frac{\partial}{\partial y} \{I(x,y)*[G(x)G(y)]\} = E_y(x,y). \quad (8.49)$$

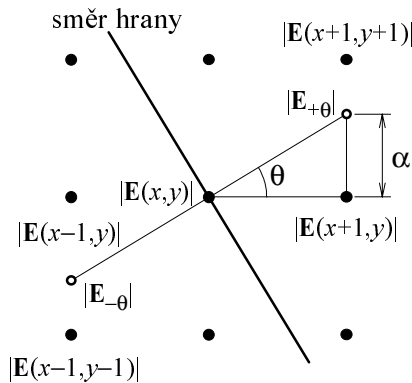
V uvedených výrazech realizuje první člen v hranatých závorkách projekci. Výstupem dvojrozměrného detektoru hran je vektorové pole $\mathbf{E}(x,y) = (E_x(x,y), E_y(x,y))$. Pro velikost a směr hrany máme

$$|\mathbf{E}(x,y)| = \sqrt{E_x^2(x,y) + E_y^2(x,y)}, \quad \varphi(x,y) = \arctg(E_y(x,y)/E_x(x,y)) + \pi/2. \quad (8.50)$$

Řekneme, že v nějakém bodě obrazu je hrana, jestliže $|\mathbf{E}(x,y)|$ v tom bodě nabývá lokálního extrému (připomeňme, že jsme se snažili, aby jedné hraně příslušelo pouze jediné maximum) a jestliže je hodnota $|\mathbf{E}(x,y)|$ v tom bodě dostatečně veliká. Předpokládejme nyní, že je obraz popsán diskrétní obrazovou funkcí. Zaměříme se nejdříve na první podmínku. Bod (x,y) může být bodem hrany jen tehdy, jestliže platí (obr. 8.14)

$$|\mathbf{E}_{-\theta}| < |\mathbf{E}(x,y)| > |\mathbf{E}_{+\theta}| \quad (8.51)$$

Hodnoty $|\mathbf{E}_{+\theta}|, |\mathbf{E}_{-\theta}|$ stanovíme interpolací (obr. 8.14). Např. leží-li směr gradientu v prvním oktantu, pak máme



$$|\mathbf{E}_{+\theta}| = \alpha |\mathbf{E}(x+1, y+1)| + (1-\alpha) |\mathbf{E}(x+1, y)|, \quad (8.52)$$

$$|\mathbf{E}_{-\theta}| = \alpha |\mathbf{E}(x-1, y-1)| + (1-\alpha) |\mathbf{E}(x-1, y)|. \quad (8.53)$$

Význam hodnoty α je patrný z obr. 8.14. Podobné vztahy lze snadno odvodit i pro zbývající oktanty. Podmínku, že hodnota $|\mathbf{E}(x,y)|$ má být v bodě hrany dostatečně velká lze realizovat prahováním. Účinné je prahování s hysterezí, při němž se používá dvou hodnot prahu t_{low}, t_{high} . Za body hrany jsou nejprve označeny všechny pixely, kde je velikost hrany větší než t_{high} . Dále jsou pak za body hrany označeny také ty pixely, kde je velikost hrany větší než t_{low} (ale menší než t_{high}) a které sousedí s pixelem, který byl již dříve označen za bod hrany. Tento krok lze opakovat několikrát.

Obr. 8.14. K rozhodnutí maxima.

8.1.5 Stanovení hran na základě textury

V tomto odstavci se zaměříme na detekci hran tvořených náhlou změnou textury ve vyšetřovaném obraze. Jednoduché je řešení, které předpokládá, že v důsledku rozdílnosti textur po obou stranách hrany je rozdílná také střední hodnota jasu v oblastech přiléhajících k hraně. Necht' $f^{(r)}(x,y)$ označuje střední hodnotu jasu v kruhovém okolí bodu (x,y) . Poloměr okolí je r . Předpokládejme, že v obraze máme hrana, jejíž směr je $\varphi + \pi/2$. Umístěme dvě taková okolí $f^{(r)}(x_L, y_L), f^{(r)}(x_R, y_R)$ nalevo a napravo od hrany (obr.8.15) a určíme diferenci

$$\begin{aligned} f_{\varphi}^{(r)}(x,y) &= |f^{(r)}(x_R, y_R) - f^{(r)}(x_L, y_L)| \\ &= |f^{(r)}(x + r \cos\varphi, y + r \sin\varphi) - f^{(r)}(x - r \cos\varphi, y - r \sin\varphi)|. \end{aligned} \quad (8.54)$$

Existence hrany v bodě (x,y) se projeví vysokými hodnotami $f_{\varphi}^{(r)}(x,y)$. Výraz (8.54) předpokládá znalost směru hrany. Není-li směr znám, lze vypočítat diference ve směru os x a y , kterých pak použijeme stejně, jak bylo popsáno u metody gradientní. Pro diference a velikost hrany máme

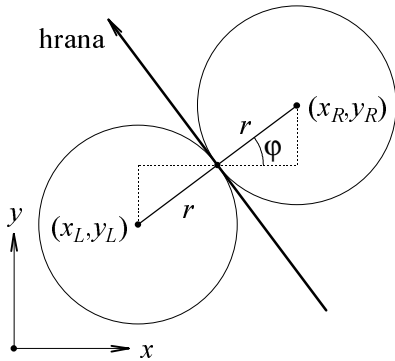
$$f_x^{(r)}(x, y) = f^{(r)}(x+r, y) - f^{(r)}(x-r, y), \quad f_y^{(r)}(x, y) = f^{(r)}(x, y+r) - f^{(r)}(x, y-r), \quad (8.55)$$

$$e^{(r)}(x, y) = \sqrt{\left[f_x^{(r)}(x, y)\right]^2 + \left[f_y^{(r)}(x, y)\right]^2}. \quad (8.56)$$

Při praktické realizaci je možné nahradit kruhové okolí bodu (x, y) praktičtějším okolím čtvercovým, pro které lze střední hodnotu obrazové funkce vyjádřit jednoduše vztahem

$$f^{(r)}(x, y) = \left(\sum_{u=x-r}^{x+r} \sum_{v=y-r}^{y+r} f(u, v) \right) / (2r+1)^2. \quad (8.57)$$

Volba poloměru okolí r zasluží pozornosti. Jestliže je poloměr příliš malý, pak není v procesu průměrování dostatečně vyhlazen vzorek textury, což způsobuje detekci falešných hran. Příliš velký poloměr naopak způsobí, že nejsou detekovány krátké hrany. Je proto výhodné stanovit poloměr okolí na základě velikosti nejmenšího objektu, který má být detekován. Volbu je zpravidla nutné ověřit experimentem. Naznačme na závěr ještě jeden poněkud komplikovanější postup pro detekci hran z textury. Uvažujme bod (x, y) a jeho dvě okolí tak, jak bylo naznačeno na obr. 8.15. Pro obě okolí sestrojíme normalizované histogramy $p_L(z)$, $p_R(z)$ jasů. Usuzujeme, že rozdílné histogramy implikují existenci hrany v bodě (x, y) . Míru rozdílnosti obou histogramů lze určit např. ze vztahu



Obr.8.15. Ke stanovení hrany na základě textury.

$$\int \left[p_L(z) - p_R(z) \right]^2 dz.$$

8.1.6 Hledání hran v barevných obrazech

Až doposud jsme předpokládali, že zpracovávaný obraz je v každém místě charakterizován jedinou skalární hodnotou - úrovní jasu. Máme-li hledat hrany v barevných obrazech, můžeme je před zpracováním převést na černobílý a dále je zpracovávat dříve popsány postupy. Nevýhodou tohoto přístupu je ovšem ztráta jisté části informace. Proto může být někdy výhodné zpracovávat barevné obrazy přímo, bez jejich převodu na obrazy černobílé. Předpokládejme, že barva každého bodu obrazu je popsána trojicí r, g, b intenzit červené, zelené a modré barevné složky. Barevný obraz je tedy popsán vektorovou obrazovou funkcí

$$\mathbf{f}(x, y) = (r(x, y), g(x, y), b(x, y)). \quad (8.58)$$

Pro stanovení hran v barevných obrazech můžeme upravit např. již dříve popsanou gradientní metodu. Jako velikost hrany můžeme vzít např. hodnotu

$$e(x, y) = \sqrt{\left| \text{grad}(r(x, y)) \right|^2 + \left| \text{grad}(g(x, y)) \right|^2 + \left| \text{grad}(b(x, y)) \right|^2}. \quad (8.59)$$

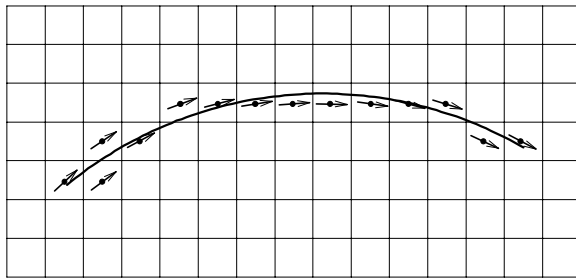
Za směr hrany můžeme považovat směr kolmý k vektoru

$$\mathbf{g}(x, y) = \text{grad}(r(x, y)) + \text{grad}(g(x, y)) + \text{grad}(b(x, y)). \quad (8.60)$$

Při praktické realizaci uvedených vztahů opět nahradíme derivace diferencemi. Na závěr poznamenejme, že podle našich zkušeností jsou výsledky získané analýzou barevných obrazů sice jednoznačně lepší než v případě převodu na obraz černobílý, avšak toto zlepšení není vždy tak významné, aby vyvážilo komplikace při výpočtu.

8.2 Spojování hran

Hranové operátory z předchozí podkapitoly určovaly velikost a případně i směr hrany jednotlivě v každém bodě obrazu, tedy lokálně. To je ovšem poněkud v rozporu s intuitivní představou hran jako



Obr.8.16. Proložení globální hrany.

delších souvislých úseček případně křivek v obrazech. Hrany chápané tímto druhým způsobem nazýváme v této podkapitole globálními hranami. K nalezení globálních hran lze využít výsledků poskytovaných lokálními operátory, které jsou však dále zpracovány v procesu tzv. spojování hran. Proces spojování je ilustrován na obr. 8.16. Ve vyobrazeném případě umožňují lokálně stanovené velikosti a směry hran snadno proložit hranu globální.

8.2.1 Heuristické sledování hrany

V heuristických metodách sledování hrany jsou globální hrany předpokládány ve tvaru úsečky nebo křivky (často se např. jedná o kruhový oblouk). Vstupní informací pro tyto metody je obraz, v němž je pro každý pixel stanovena velikost a případně i směr hrany. Takový obraz lze získat pomocí některého z lokálních hranových operátorů. Někdy se také může jednat o naskenovaný obraz technického výkresu, v němž jsou čáry po naskenování upraveny ztenčováním. Heuristické sledování globální hrany zahrnuje zpravidla následující kroky: 1) nalezení počátečního bodu globální hrany, 2) predikce polohy následujícího bodu, 3) hledání bodu hrany v okolí predikované polohy, 4) určení následujícího kroku algoritmu. Jednotlivé kroky popíšeme dále podrobněji.

Nalezení počátečního bodu globální hrany: V tomto kroku nalezneme postupným prohledáváním vstupního obrazu některý bod globální hrany. Je nutné, aby tento bod byl detekován spolehlivě, což lze zajistit volbou vysoké hodnoty prahu. Jestliže je k dispozici také směr hrany, lze této informace použít pro predikci polohy druhého bodu hrany. Není-li směr k dispozici, hledáme druhý bod hrany mezi všemi sousedy bodu prvního.

Predikce polohy následujícího bodu: Jestliže již bylo nalezeno několik předchozích bodů hrany, pak lze na základě této znalosti a na základě očekávaného tvaru hrany (úsečka, křivka) odhadnout polohu následujícího bodu. Podle očekávaného tvaru globální hrany proložíme doposud nalezenými body přímkou nebo křivku. K proložení se často používá metody nejmenších čtverců. Proložení snadno získáme predikci polohy následujícího bodu. Pro stanovení okolí, ve kterém má být následující bod hledán, se používají různé heuristické vztahy, které se opírají o úvahu, že čím více bodů hrany bylo již nalezeno a čím kvalitněji lze předpokládaný tvar nalezenými body proložit, tím je odhad polohy následujícího bodu jistější, a prohledávané okolí může být proto menší. Predikovat lze rovněž očekávanou velikost lokální hrany, což umožňuje stanovit vhodnou velikost prahu, který bude v následujícím bodě pro rozhodnutí o existenci hrany použit.

Hledání bodu v okolí predikované polohy: Bod hrany hledáme v oblasti, jejíž poloha a velikost byly stanoveny predikcí. Jestliže je ve stanoveném okolí nalezeno více vyhovujících bodů, lze vybrat nejlepší z nich (tj. bod, který je blízko predikované polohy a v němž je vysoká velikost hrany).

Určení následujícího kroku: Je nutné rozhodnout, zda lze ve sledování globální hrany pokračovat nebo zda má být sledování ukončeno. Jestliže je ve stanoveném okolí nalezen bod, který dostatečně přesně odpovídá predikci a jestliže lze předpokládaný tvar hrany proložit dosud nalezenou množinou bodů dostatečně kvalitně, lze ve sledování pokračovat. Jinak je nutné sledování ukončit. Kvalitu proložení lze hodnotit např. pomocí maxima vzdálenosti nalezených bodů od prokládané křivky.

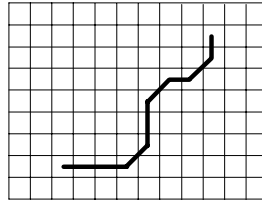
Po ukončení sledování jedné hrany přechází algoritmus na sledování hrany následující. Hrany, které již byly nalezeny, jsou ze vstupního obrazu odstraněny, aby nebyly detekovány opakovaně. Metody

heuristického sledování hrany jsou účinné. Jejich implementace je však bohužel obtížnější a vyžaduje pracné řešení řady drobných detailů.

Na závěr tohoto odstavce poznamenejme, že jednou z možností, jak popsat extrahované křivky, je použít tzv. Freemanova kódu. Ve Freemanově kódu je křivka popsána od jednoho ze svých konců. Postupně jsou zapisovány kódy směrů tak, jak křivka postupuje od jednoho pixelu k pixelu následujícímu.

3	2	1
4	x	0
5	6	7

Obr.8.17. Kódování směrů ve Freemanově kódu.



Obr.8.18. Kódování tvaru křivky pomocí Freemanova kódu (0001221012).

Kódování směrů ukazuje obr. 8.17. Příklad kódování křivky je uveden na obr. 8.18. Extrahované křivky lze ovšem také popsat pomocí úseček a oblouků (kruhových nebo parabolických), které jimi lze proložit. Někdy je popis pomocí Freemanova kódu meziřadně k získání uvedeného popisu.

8.2.2 Proložení přímkou a křivky

V tomto odstavci ukážeme, jak lze metodou nejmenších čtverců proložit danou množinou bodů přímkou, kružnicí nebo kruhový oblouk. Zaměříme se nejprve na přímku. Rovnici přímkou uvažujeme ve tvaru $\alpha x + \beta y + \gamma = 0$. Předpokládejme, že $\beta \neq 0$, a děleme rovnici touto hodnotou. Dostane rovnici $ax + y + c = 0$. Příмка má být proložena n body. Souřadnice i -tého bodu označme x_i, y_i . Hodnota $\varepsilon_i = ax_i + y_i + c$ je úměrná vzdálenosti i -tého bodu od přímkou. Hodnoty a, c lze stanovit z podmínky, aby součet čtverců hodnot ε_i byl minimální. Hledáme tedy minimum výrazu

$$\varepsilon = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (ax_i + y_i + c)^2. \quad (8.61)$$

Obvyklým postupem, kdy položíme $\partial\varepsilon/\partial a = 0$ a $\partial\varepsilon/\partial c = 0$, obdržíme následující výsledek

$$a = \frac{\sum_{i=1}^n x_i \sum_{i=1}^n y_i - n \sum_{i=1}^n x_i y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2}, \quad c = \frac{\sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2}. \quad (8.62)$$

Dosažená hodnota ε určuje kvalitu proložení. Pokud by pro prokládanou přímkou platilo $\beta = 0$, pak nelze provést dělení rovnice přímkou hodnotou β . V takovém případě lze ovšem dělit hodnotou α , což dává rovnici $x + by + c = 0$. Tento případ lze řešit analogicky jako případ předchozí. Výsledné vztahy pro b, c vycházejí podobně - postačí zaměnit a za b a prohodit x_i a y_i . V praxi zpravidla dopředu nevíme, který z obou případů nastane. Můžeme proto vypočítat hodnoty jmenovatele ve vzorcích pro obě varianty a použít tu, pro níž je absolutní hodnota jmenovatele větší.

Při proložení kružnicí vyjdeme z rovnice kružnice $(x - a)^2 + (y - b)^2 - r^2 = 0$, kde a, b jsou hledané souřadnice středu kružnice a r je její poloměr. Metodou nejmenších čtverců lze snadno odvodit řešení, kdy minimalizujeme součet čtvrtých mocnin vzdáleností zadaných bodů od prokládané kružnice. Minimalizujeme tedy hodnotu

$$\varepsilon = \sum_{i=1}^n \left[(x_i - a)^2 + (y_i - b)^2 - r^2 \right]^2. \quad (8.63)$$

Položíme-li $\partial\varepsilon/\partial a = 0$, $\partial\varepsilon/\partial b = 0$, $\partial\varepsilon/\partial r = 0$, pak po zdlouhavém výpočtu obdržíme výsledek

$$r = \sqrt{\frac{1}{n} \left(na^2 - 2a \sum_{i=1}^n x_i + \sum_{i=1}^n x_i^2 + nb^2 - 2b \sum_{i=1}^n y_i + \sum_{i=1}^n y_i^2 \right)}, \quad (8.64)$$

$$a = \frac{\sum_{i=1}^n x_i^3 + \sum_{i=1}^n xy_i^2 - \frac{1}{n} \sum_{i=1}^n x_i^2 \sum_{i=1}^n x_i - \frac{1}{n} \sum_{i=1}^n y_i^2 \sum_{i=1}^n x_i - 2b(\sum_{i=1}^n x_i y_i - \frac{1}{n} \sum_{i=1}^n x_i \sum_{i=1}^n y_i)}{2[\sum_{i=1}^n x_i^2 - \frac{1}{n} (\sum_{i=1}^n x_i)^2]}, \quad (8.65)$$

$$b = \frac{A + B + C}{D}, \quad (8.66)$$

kde

$$A = \left(\frac{1}{n} \sum_{i=1}^n x_i \sum_{i=1}^n y_i - \sum_{i=1}^n x_i y_i\right) \left(\sum_{i=1}^n x_i^3 + \sum_{i=1}^n x_i y_i^2\right),$$

$$B = \frac{1}{n} \left(\sum_{i=1}^n x_i y_i \sum_{i=1}^n x_i - \sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i\right) \left(\sum_{i=1}^n x_i^2 + \sum_{i=1}^n y_i^2\right),$$

$$C = \left(\sum_{i=1}^n x_i^2 - \frac{1}{n} [\sum_{i=1}^n x_i]^2\right) \left(\sum_{i=1}^n y_i^3 + \sum_{i=1}^n y_i x_i^2\right),$$

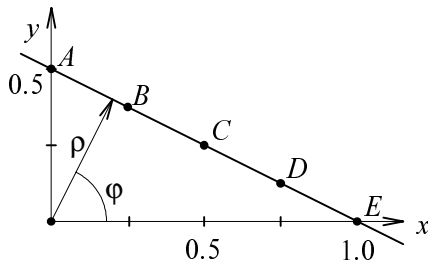
$$D = 2\left[\left(\sum_{i=1}^n y_i^2 - \frac{1}{n} [\sum_{i=1}^n y_i]^2\right) \left(\sum_{i=1}^n x_i^2 - \frac{1}{n} [\sum_{i=1}^n x_i]^2\right) - \left(\sum_{i=1}^n x_i y_i - \frac{1}{n} \sum_{i=1}^n x_i \sum_{i=1}^n y_i\right)^2\right].$$

Proložení kružnice nebo kruhového oblouku může být užitečné i tehdy, jestliže křivky, které z obrazu extrahujeme, nemají tvar kružnice. Proložení kružnice několika po sobě jdoucími body křivky můžeme získat aproximaci oskulační kružnice. Převrácená hodnota poloměru oskulační kružnice je rovna křivosti křivky. Průběh křivosti podél křivky je významnou informací, která, jak ukážeme v kapitole 9, může být využita i při rozpoznání.

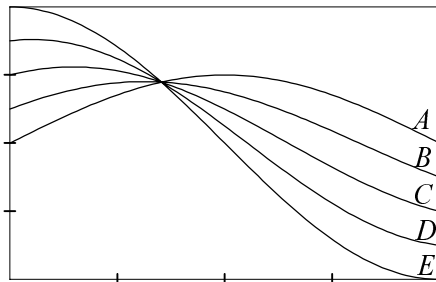
8.2.3 Houghova transformace

Předpokládejme, že jsme již pomocí některé z metod z podkapitoly 8.1 rozhodli, které body obrazu považujeme za body hran. Nyní chceme ověřit, zda jsou body hran v obraze položeny tak, že naznačují existenci přímek (úseček) v obraze. Metoda, kterou zde popíšeme, byla navržena Houghem a později zdokonalena Dudou a Hartem (Duda 72). Metoda je známa pod názvem Houghova transformace. Uvažujme rovnici přímky ve tvaru (obr.8.19)

$$\rho = x \cos \varphi + y \sin \varphi. \quad (8.67)$$



Obr.8.19. Houghova transformace. Body ležící na přímce.

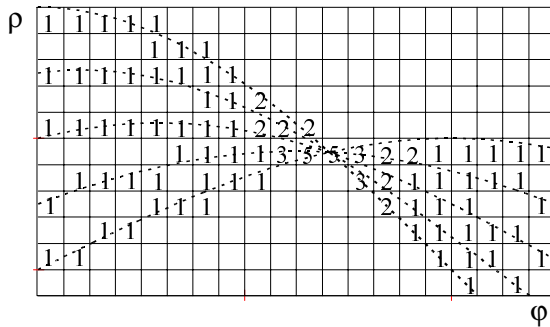


Obr.8.20. Bodům ležícím na přímce odpovídá v prostoru φ, ρ průsečík sinusoid.

Pro přímku procházející bodem Q reprezentovaným vektorem (x_Q, y_Q) pak platí $\rho = x_Q \cos \varphi + y_Q \sin \varphi$. Uvažujme nyní prostor φ, ρ . V tomto prostoru odpovídá každé přímce procházející bodem Q jediný bod. Svazku přímek procházejících bodem Q pak v tomto prostoru odpovídá sinusoida $\rho = x_Q \cos \varphi + y_Q \sin \varphi$. Mějme nyní v obraze např. body A, B, C, D, E , které leží na přímce (obr.8.19). Svazkům možných přímek v těchto bodech odpovídají v prostoru φ, ρ sinusoidy A, B, C, D, E (obr.8.20). Obrazem přímky, na které leží body A, B, C, D, E (tedy přímky, která je společná všem svazkům), je v prostoru φ, ρ bod, v němž se sinusoidy A, B, C, D, E protínají.

Dosud provedené úvahy vedou k následujícímu postupu výpočtu. Rozdělme prostor φ, ρ na oblasti. Nechť např. dvojice (i, j) označuje v prostoru φ, ρ obdélníkovou oblast, pro niž platí $\varphi_{i-1} \leq \varphi < \varphi_i$, $\rho_{j-1} \leq \rho < \rho_j$. Zaveďme nad takto diskretizovaným prostorem

φ, ρ dvourozměrný histogram $h(i, j)$. Počáteční hodnota $h(i, j)$ je pro všechna (i, j) nula. Nyní probíráme všechny pixely obrazu. Pro každý pixel $Q = (x_Q, y_Q)$, kde v obraze nalezneme hranu (předpokládáme,



Obr.8.21. Bodům ležícím na přímce odpovídá vysoká hodnota v histogramu.

že neznáme její směr), sestrojíme v prostoru φ, ρ sinusoidu (obr. 8.21). Při tom v histogramu inkrementujeme hodnoty $h(i, j)$ pro ty oblasti (i, j) , kterými sinusoida prochází. Po skončení tohoto procesu $h(i, j) = n$ popisuje skutečnost, že na přímce charakterizované hodnotami $\varphi \in \langle \varphi_{i-1}, \varphi_i \rangle$, $\rho \in \langle \rho_{j-1}, \rho_j \rangle$ leží v obraze n bodů, které byly považovány za hrany. Pomocí Houghovy transformace tak můžeme zjistit, že v obraze leží v jedné přímce vysoký počet bodů hran. V případě potřeby je toto zjištění možné dále doplnit analýzou, jak jsou detekované body rozloženy podél nalezené přímky.

Jestliže v jednotlivých bodech obrazu dokážeme hranu nejen detekovat, ale také dostatečně spolehlivě určit její směr, pak lze popsanou metodu zjednodušit. Necht' (x, y) je bod hrany a ψ směr hrany v tomto bodě. Potom podle vztahu $\rho = x \cos(\psi - \pi/2) + y \sin(\psi - \pi/2)$ můžeme ihned vypočítat odpovídající hodnotu ρ . V tomto případě tedy místo generování sinusoidy generujeme pouze jednotlivé body a histogram inkrementujeme pouze v jediné oblasti. Na závěr poznamenejme, že Houghovu transformaci lze zobecnit také na křivky. Protože počet hledaných parametrů křivky udává potřebný rozměr histogramu, jedná nejčastěji o křivky se dvěma, nejvýše třemi parametry.

8.3 Detekce oblastí

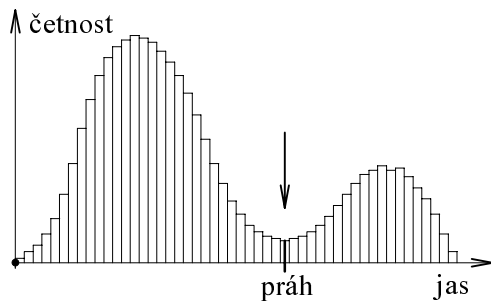
V podkapitole 8.1 jsme popsali metody segmentace obrazu, které byly založeny na detekci hranice oblastí. Jiná skupina metod se zaměřuje přímo na vyhledávání celých oblastí v obraze. Oblasti jsou vyhledávány na základě jistým způsobem zvoleného kritéria homogenity oblastí. Takovým kritériem mohou být např. konstantní nebo dosti blízké hodnoty jasu, barvy, pokrytí stejnou texturou atd. Protože ze znalosti hranice dokážeme určit také vnitřek oblastí a naopak, zdá se, že by metody založené na detekci hranic i metody založené na detekci celých oblastí měly být vzájemně rovnocenné. Už na obr. 8.1 jsme však naznačili, že tomu tak v praxi nemusí vždy být. Metody založené na detekci oblastí jsou často upřednostňovány u dosti zašumělých obrazů, kde by detekce hranic byla nespolehlivá. Metody detekce oblastí lze rozdělit do tří základních skupin: 1) detekce prahováním, 2) detekce narůstáním oblastí, 3) detekce dělením oblastí. V následujících odstavcích tyto metody popíšeme podrobněji.

8.3.1 Prahování

Prahování je nejjednodušší metodou detekce celých oblastí v obraze. V nekomplikovaných případech se při tom však současně jedná o metodu rychlou a spolehlivou. Výsledkem prahování je binární obraz, v němž je bodům (pixelům) nalezených oblastí obvykle přiřazena hodnota 1, zatímco zbývajícím bodům (bodům pozadí) je přiřazena hodnota 0. Při prahování se vychází z předpokladu, že body hledaných oblastí mají stejný nebo dosti podobný jas. Jedna z možných realizací prahování spočívá v tom, že pixel je označen jako bod oblasti, jestliže jeho jas padne do intervalu $\langle a, b \rangle$. Velmi často je v rozhodovacím kritériu použita pouze jediná hodnota t - tzv. práh. Body, jejichž jas je vyšší než hodnota prahu, jsou detekovány jako body hledaných oblastí a body, jejichž jas je nižší, jsou detekovány jako body pozadí (případně naopak). Prahování se tedy děje na základě předpisu

$$g(x, y) = \begin{cases} 1, & f(x, y) \in \langle a, b \rangle \\ 0, & \text{jinak} \end{cases} \quad \text{nebo} \quad g(x, y) = \begin{cases} 1, & f(x, y) \geq t \\ 0, & \text{jinak} \end{cases} \quad (8.68)$$

Úspěšnost prahování závisí na znalosti správné hodnoty prahu. Jestliže tuto hodnotu neznáme, je možné pokusit se ji stanovit na základě informací získaných z obrazu, který má být segmentován. Pro obrazy s bimodálním histogramem (histogramem se dvěma vrcholy) jasu se např. často doporučuje

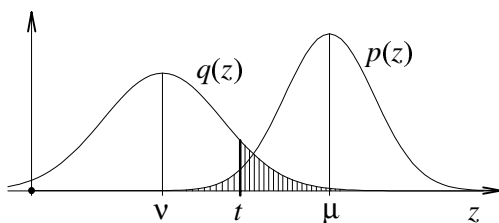


Obr.8.22. Bimodální histogram jasu.

volit jako hodnotu t prahu hodnotu, v níž histogram dosahuje mezi oběma vrcholy minima (obr. 8.22). Lze snadno domyslet, že tato heuristická metoda má racionální základ. Předpokládá totiž, že v obraze existují dva druhy pixelů: pixely náležící hledaným oblastem a pixely náležící pozadí. Oba druhy pixelů jsou při tom relativně četné a mají dosti odlišný jas, což v histogramu dává vzniknout zmíněným dvěma vrcholům. Nejsou-li na druhé straně uvedené předpoklady splněny, nemusí být výsledky metody uspokojivé. Ne vždy je také histogram bimodální.

V případech, kdy je obraz sice kontrastní, avšak v různých svých částech má nerovnoměrnou úroveň jasu, nelze často najít jedinou hodnotu prahu tak, aby vyhovovala pro všechny části obrazu. V takovém případě lze použít prahování s proměnnou hodnotou prahu, kdy se pro různá místa obrazu používá různých hodnot prahu. Metodu lze realizovat tak, že se obraz rozdělí na několik částí a pro každou se najde individuální hodnotu prahu. Je-li histogram úrovní jasů vyšetřované části obrazu bimodální, lze hodnotu prahu stanovit např. postupem popsáním v předchozím odstavci. Jestliže ve vyšetřované části obrazu velmi převažují buď body oblasti nebo body pozadí, pak hodnotu prahu z histogramu uvažované části obrazu spolehlivě stanovit nelze. V takovém případě lze hodnotu prahu stanovit např. jako průměr hodnot prahů z přilehlých oblastí.

Ukážeme dále důmyslnější metodu stanovení prahu, která je založena na minimalizaci rizika a na využití počtu pravděpodobnosti. Předpokládejme, že jasy pixelů oblastí, které mají být detekovány, mají normální rozložení hustoty pravděpodobnosti $p(z)$ se střední hodnotou μ a směrodatnou odchylkou σ



Obr.8.23. Stanovení prahu minimalizací chyby.

(obr.8.23). Podobně předpokládáme, že také jasy pixelů pozadí mají normální rozložení hustoty pravděpodobnosti $q(z)$ se střední hodnotou v a směrodatnou odchylkou τ (přísně teoreticky vzato nemohou být ovšem rozložení normální, protože prakticky jsou jasové úrovně nenulové pouze v jistém intervalu; může se však jednat o dobré přiblížení). Dále předpokládejme, že podíl pixelů hledaných oblastí na celkovém počtu pixelů obrazu je θ a že platí $\mu > v$. Předpokládejme nyní, že zvolíme práh t .

Označme $P(t)$ pravděpodobnost jevu, že bod hledané oblasti bude nesprávně vyhodnocen jako bod pozadí. Dále označme $Q(t)$ pravděpodobnost jevu, že bod pozadí bude správně vyhodnocen jako bod pozadí. Pak $1 - Q(t)$ je pravděpodobnost, že bod pozadí bude nesprávně vyhodnocen jako bod oblasti. Pro pravděpodobnosti $P(t)$ a $Q(t)$ máme (obr.8.23)

$$P(t) = \int_{-\infty}^t p(z) dz, \quad Q(t) = \int_{-\infty}^t q(z) dz. \quad (8.69)$$

Dále musíme uvážit, že poměrné zastoupení bodů objektu v obraze je θ a poměrné zastoupení bodů pozadí v obraze je $(1-\theta)$. Výsledná pravděpodobnost chybné klasifikace je pak

$$\varepsilon = \theta P(t) + (1-\theta)[1 - Q(t)]. \quad (8.70)$$

K nalezení minima výraz (8.70) derivujeme a derivaci položíme rovnu nule. Dostaneme

$$\frac{\partial \varepsilon}{\partial t} = \theta \frac{\partial P(t)}{\partial t} - (1-\theta) \frac{\partial Q(t)}{\partial t} = \theta p(t) - (1-\theta)q(t) = 0. \quad (8.71)$$

$$\text{Odtud máme} \quad (1-\theta)q(t) = \theta p(t). \quad (8.72)$$

Podle předpokladu mají hustoty pravděpodobnosti $p(z)$ i $q(z)$ normální rozložení. Je tedy

$$p(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(t-\mu)^2}{2\sigma^2}\right], \quad q(t) = \frac{1}{\tau\sqrt{2\pi}} \exp\left[-\frac{(t-\nu)^2}{2\tau^2}\right]. \quad (8.73)$$

Dosažením vztahů (8.73) do (8.72) a logaritmováním máme

$$\ln(1-\theta) - \ln \tau - \frac{(t-\nu)^2}{2\tau^2} = \ln \theta - \ln \sigma - \frac{(t-\mu)^2}{2\sigma^2}. \quad (8.74)$$

$$\text{Po úpravě dostaneme} \quad \tau^2(t-\mu)^2 - \sigma^2(t-\nu)^2 = 2\sigma^2\tau^2 \ln \frac{\tau\theta}{\sigma(1-\theta)}. \quad (8.75)$$

Výraz (8.75) je kvadratickou rovnicí pro t . K řešení je zapotřebí znát hodnoty $\mu, \sigma, \nu, \tau, \theta$. Povšimněme si blíže ještě případu, kdy $\theta = 1/2$ a $\sigma = \tau$. V tomto případě se rovnice 8.75 zjednoduší na tvar

$$(t-\mu)^2 = (t-\nu)^2. \quad (8.76)$$

Řešením rovnice (8.76) je intuitivně očekávaná hodnota $t = (\mu + \nu) / 2$.

8.3.2 Metoda spojování oblastí

Metody založené na spojování začínají pracovat s malými oblastmi (dokonce např. i s jednotlivými pixely), které postupně spojují do větších oblastí. Proces se iterativně opakuje tak dlouho, dokud v obraze existují oblasti, které lze spojovat. Obvyklými kritérii, která jsou uplatňována při rozhodování, zda lze dvě sousedící oblasti spojit do jediné větší oblasti, jsou následující: Obě oblasti by měly mít stejný nebo podobný jas, barvu, texturu atd. Mezi oběma oblastmi by neměla být zřetelná hrana (existence slabě zřetelné nebo přerušované hrany mezi oběma oblastmi se někdy ale připouští). Spojení oblastí je tím více žádoucí, čím větší je délka, podél níž se oblasti vzájemně dotýkají (délka se měří relativně vzhledem k rozměrům oblastí). Spojení může být naopak nežádoucí v případě, že by spojením vznikla oblast nevhodného nebo neočekávaného tvaru. Metodám založeným na spojování oblastí byla v minulosti věnována značná pozornost. Testování výše uvedených kritérií bylo realizováno různými heuristickými postupy. Metodu spojování oblastí je možné považovat za metodu účinnou. Je však náročnější na implementaci a zpravidla i na výpočetní výkon počítače.

8.3.3 Metoda dělení oblastí

Tato metoda zpravidla začíná s jedinou oblastí tvořenou celým obrazem a dále opakovaně provádí její dělení na dílčí oblasti. Proces dělení pokračuje tak dlouho, dokud není ve všech vzniklých oblastech dosaženo splnění kritéria homogenity. Dále uvedeme dva příklady kritéria homogenity.

Dělení oblasti modální metodou: Často používaným prostředkem k posouzení homogenity oblastí je histogram některé její vlastnosti (např. histogram úrovně jasů). Vychází se z představy, že jestliže histogram obsahuje dva nebo více vrcholů (je bimodální nebo multimodální), pak je vyšetřovaná oblast nehomogenní a je nutné ji rozdělit. K rozdělení lze opět využít znalosti histogramu. Např. u bimodálního histogramu lze za dělicí prahovou hodnotu zvolit hodnotu, která odpovídá „údolí“ mezi oběma vrcholy (obr. 8.22). Provedeme-li prahování oblasti s touto hodnotou prahu, je oblast rozdělena na dvě nebo více nových oblastí. U multimodálních histogramů lze provést dělení buď na základě většího počtu prahů nebo lze vybrat jediný práh odpovídající nejvýznamnějšímu „údolí“ v histogramu a dělení provádět několikrát. Druhý postup bývá považován za spolehlivější. Proces dělení končí, když histogram každé oblasti obsahuje nejvýše jeden vrchol.

Dělení oblasti metodou diskriminantu: Existence nehomogenit v oblasti se ne vždy projeví tím, že je histogram (např. jasových úrovní) bimodální. V takovém případě je nutné hledat jiná kritéria. V (Otsu 78) byla např. prezentována metoda, která používá diskriminačního kritéria. Předpokládejme, že pracujeme s úrovněmi jasů $1, 2, \dots, L$. Uvažujme oblast, která má n pixelů. Označme n_i počet pixelů oblas-

ti, které mají jas i . Dále zavedme $p_i = n_i/n$. Střední hodnota a rozptyl jasů v uvažované oblasti jsou popsány výrazy

$$\mu = \sum_{i=1}^L ip_i, \quad \sigma^2 = \sum_{i=1}^L (i-\mu)^2 p_i. \quad (8.77)$$

Vysoká hodnota σ naznačuje, že je nutné oblast rozdělit. Dělení oblasti provedeme prahováním s prahem t . Stanovení prahu je cílem následující úvahy. Počet pixelů s jasnem menším nebo rovným t označme m_1 , počet pixelů s jasnem větším než t označme m_2 ($m_1 + m_2 = n$). Zavedme $w_1 = m_1/n$, $w_2 = m_2/n$. Střední hodnota a rozptyl jasů pro oblasti s jasnem nižším (vyšším) než t je

$$\mu_1 = \sum_{i=1}^t \frac{ip_i}{w_1}, \quad \sigma_1^2 = \sum_{i=1}^t \frac{(i-\mu_1)^2 p_i}{w_1}, \quad \mu_2 = \sum_{i=t+1}^L \frac{ip_i}{w_2}, \quad \sigma_2^2 = \sum_{i=t+1}^L \frac{(i-\mu_2)^2 p_i}{w_2}. \quad (8.78)$$

Pro rozptyl σ_w^2 uvnitř tříd a rozptyl σ_b^2 mezi třídami platí (Otsu 78)

$$\sigma_w^2 = w_1\sigma_1^2 + w_2\sigma_2^2, \quad \sigma_b^2 = w_1(\mu_1 - \mu)^2 + w_2(\mu_2 - \mu)^2 = w_1w_2(\mu_1 - \mu_2)^2. \quad (8.79)$$

Neznámou hodnotu prahu určíme tak, aby minimalizovala σ_w^2 nebo maximalizovala σ_b^2 . Protože platí $\sigma_w^2 + \sigma_b^2 = \sigma^2$, jsou obě kriteria ekvivalentní. Nevýhodou této jinak pěkné metody je, že se při stanovení hodnoty prahu nelze vyhnout iteračnímu procesu.

8.4 Detekce rohů

Někdy se další zpracování obrazu může opírat o nalezení rohů (vrcholů) oblastí, což může být výhodné zejména z následujících důvodů: 1) Rohy bývají bohatě zastoupeny v obrazech scén vytvořených člověkem i v obrazech scén „přírodních“. 2) Polohu rohů lze v obrazech zpravidla jednoznačně určit. 3) Rohy v obrazech často korespondují s rohy (vrcholy) skutečných trojdimenzionálních objektů, čehož lze využít např. při rekonstrukci scény. Uvažujme obraz ve stupních šedi a vyšetřujme průběh křivosti plochy popsané obrazovou funkcí. Je zřejmé, že v místě rohu není křivost v žádném směru nulová (to je rozdíl ve srovnání s hranou). Detektory rohů jsou často založeny na uvedeném pozorování. Dále uvedeme dva příklady detektorů.

8.4.1 Kitchen-Rosenfeldův detektor rohů

Tento detektor byl publikován v (Kitchen 82). Necht' $I(x,y)$ označuje jas obrazu v bodě x,y a necht' $I_x, I_y, I_{xx}, I_{yy}, I_{xy}$ jsou derivace funkce I . Autoři navrhují v každém bodě obrazu počítat hodnotu

$$k = \frac{I_{xx}I_y^2 + I_{yy}I_x^2 - 2I_{xy}I_xI_y}{I_x^2 + I_y^2} \quad (8.80)$$

a ukazují, že uvedená hodnota je úměrná velikosti křivosti a velikosti gradientu obrazové funkce v uvažovaném bodě. Bod (x,y) lze považovat za roh, jestliže je v něm hodnota k dostatečně vysoká.

8.4.2 Harrisův detektor rohů

K novějším detektorům rohů patří detektor navržený Harrisem a Stephensem (Harris 88). V případě tohoto detektoru konstruujeme v každém bodě obrazu matici

$$\mathbf{M}(x,y) = \begin{bmatrix} \langle (I_x)^2 \rangle & \langle I_x I_y \rangle \\ \langle I_x I_y \rangle & \langle (I_y)^2 \rangle \end{bmatrix}. \quad (8.81)$$

Lomené závorky označují konvoluci s Gaussovou funkcí:

$$\langle \psi(x,y) \rangle = \psi(x,y) * [G(x)G(y)], \quad (8.82)$$

kde znak $*$ označuje konvoluci a $G(u)$ je Gaussova funkce

$$G(u) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{u^2}{2\sigma^2}\right). \quad (8.83)$$

Volba hodnoty σ ovlivňuje citlivost detektoru. Derivace ve vztahu (8.81) jsou v praxi aproximovány diferencemi. K rozhodnutí, zda daný bod (x,y) lze považovat za roh, lze využít funkce (Harris 88)

$$\text{cor}(x, y) = \det(\mathbf{M}(x, y)) - 0.04 \cdot \text{trace}^2(\mathbf{M}(x, y)). \quad (8.84)$$

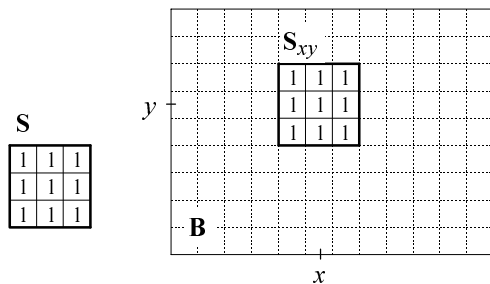
Bod (x,y) považujeme za roh, jestliže jsou splněny následující dvě podmínky: 1) $\text{cor}(x,y)$ je větší než nějaká předem stanovená hodnota; 2) hodnota $\text{cor}(x,y)$ je největší v okně o předem definovaných rozměrech (okno je svým prostředním pixelem umístěno do bodu (x,y)).

8.5 Zpracování binárních obrazů

Binárním obrazem nazýváme takový obraz, v němž obrazová funkce v každém bodě (v každém pixelu) nabývá jedné ze dvou možných hodnot. Binární obrazy jsou zpravidla výsledkem metod provádějících segmentaci obrazu (v pixelech náležících objektům např. obrazová funkce nabývá hodnoty 1, v pixelech pozadí nabývá hodnoty 0). Před tím, než jsou binární obrazy analyzovány, lze je zpracovat některým ze speciálních postupů, které dále ve stručném přehledu popíšeme.

8.5.1 Matematická morfologie

Tvůrcem matematické morfologie je Serra (Serra 82). Z poměrně nedávných publikací zabývajících se touto problematikou uveďme alespoň práci (Dougherty 92). Teorie matematické morfologie je dosti obsáhlá, a proto můžeme v tomto textu uvést pouze základní informace. Označme \mathbf{B} vstupní binární obraz. Dále budeme pracovat s pomocným binárním obrazem \mathbf{S} . Význam tohoto pomocného obrazu



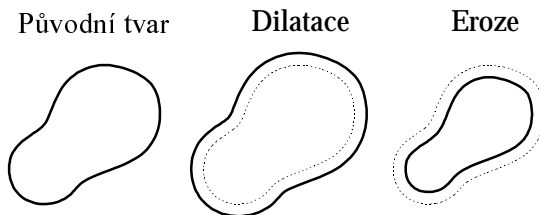
Obr.8.24. K významu masky u morfologických operací.

bude obdobný jako význam masky u konvoluce - budeme jej postupně přikládat na různá místa obrazu \mathbf{B} . Označení \mathbf{S}_{xy} budeme používat pro obraz, který vznikne translací pomocného obrazu \mathbf{S} tak, aby počátek obrazu \mathbf{S} padl do bodu o souřadnicích (x,y) (obr.8.24). Základními operacemi matematické morfologie jsou *eroze* a *dilatace*. Tyto operace jsou definovány následujícími vztahy (obr.8.25)

$$\mathbf{E} = \mathbf{B} \otimes \mathbf{S} = \{x, y \mid \mathbf{S}_{x,y} \subseteq \mathbf{B}\}, \quad (8.85)$$

$$\mathbf{D} = \mathbf{B} \oplus \mathbf{S} = \{x, y \mid \mathbf{S}_{x,y} \cap \mathbf{B} \neq \emptyset\}. \quad (8.86)$$

Vysvětlíme podrobněji význam vztahu (8.85). Erozí binárního obrazu \mathbf{B} za použití masky \mathbf{S} vznikne obraz \mathbf{E} , který je opět binární. Předpokládejme, že jednotlivé body binárního obrazu nesou hodnotu 0 nebo 1. V bodě o souřadnicích (x,y) je v obraze \mathbf{E} hodnota 1, jestliže je v obraze \mathbf{B} hodnota 1 alespoň na těch místech, kde je hodnota 1 v masce \mathbf{S}_{xy} . Jinak je v obraze \mathbf{E} v bodě o souřadnicích (x,y) hodnota 0. Ve vztahu (8.85) jsou obrazy formálně reprezentovány jako množiny - jedná se o množiny pixelů nesoucích hodnotu 1. Analogicky by bylo možné interpretovat také vztah (8.86).



Obr.8.25. Dilatace a eroze.

Dalšími často používanými operacemi matematické morfologie jsou otevření a uzavření. Otevření je definováno jako eroze následovaná dilatací (vztah 8.87). Uzavření je naopak dilatace následovaná erozí (vztah 8.88). Máme tedy

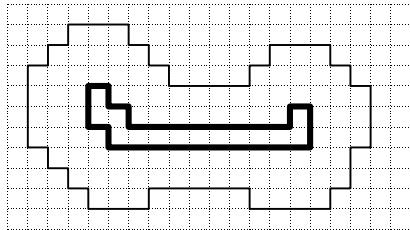
$$\mathbf{B} \circ \mathbf{S} = (\mathbf{B} \otimes \mathbf{S}) \oplus \mathbf{S}, \quad (8.87)$$

$$\mathbf{B} \bullet \mathbf{S} = (\mathbf{B} \oplus \mathbf{S}) \otimes \mathbf{S}. \quad (8.88)$$

Z definice je zřejmé, že otevření eliminuje malé a tenké objekty a rozděljuje objekty v místech, kde jsou tenké. Uzavření naopak vyplňuje malé a tenké díry v objektech a spojuje objekty, které leží blízko sebe. Poznamenejme, že někdy může být výhodné použít většího počtu otevření nebo uzavření po sobě. Alternativně může také být otevření a uzavření realizováno tak, že je použito většího počtu operací eroze a stejného počtu operací dilatace.

8.5.2 Ztenčování

Cílem ztenčování je reprezentovat objekty v obraze jako lineární útvary (obr.8.26). Ztenčování může být realizováno pomocí opakovaně prováděné eroze (tj. opakovaným odstraňováním krajních pixelů

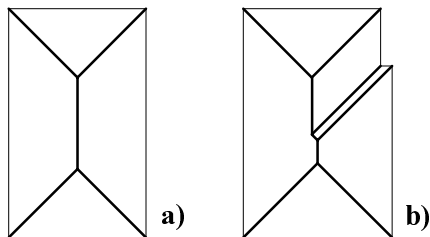


Obr.8.26. Objekt před a po ztenčení.

z objektu). Postup provádění každého erozního kroku je při tom modifikován tak, aby nedošlo k porušení souvislosti objektu. Při praktickém výpočtu jsou dle vztahu (8.85) nejprve detekovány pixely objektu kandidující na to, že do nich bude zapsána hodnota 0 (tj. budou z objektu odstraněny). Tato hodnota je pak ale skutečně zapsána jen do některých z kandidujících pixelů. Pixely jsou vybrány tak, aby zápisem hodnoty 0 nedošlo k rozdělení zbývajících částí objektu na více částí. Proces opakovaného provádění eroze pokračuje tak dlouho, dokud lze nějaké pixely odstraňovat.

8.5.3 Kostra

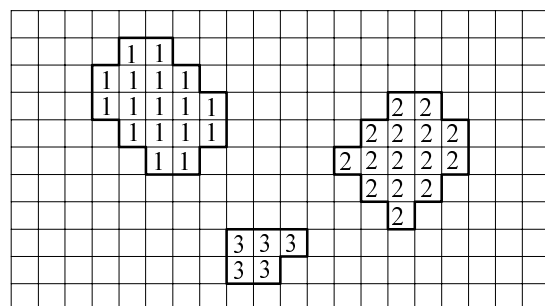
Také v případě kostry je cílem popsat tvar plošného objektu pomocí lineárních útvarů. Často se uvádí (Fu 86, Castleman 96) následující neformální definice kostry: Představme si oblast, jejíž kostru hledáme, jako trávník. Jestliže trávník po celém jeho okraji zapálíme a bude-li se oheň šířit všemi směry stejnou rychlostí, pak množinu bodů, kde se setkají „různé ohně“ nazveme kostrou. Kostrou oblasti ve tvaru kruhu je jeho střed. Příklad na obr. 8.27a ukazuje kostru obdélníka. Na obr. 8.27b je uveden příklad naznačující obtíže, které vzniknou při doslovné aplikaci dříve uvedené definice kostry. Ačkoli malý výstupek na pravé straně obdélníka může být pouhou nepřesností způsobenou digitalizací, výsledná kostra se dosti liší od kostry ideálního obdélníka. Při praktické implementaci algoritmů pro stanovení kostry je zpravidla zapotřebí tyto jevy potlačit.



Obr.8.27. a) Kostra obdélníka. b) Vliv chyby způsobené digitalizací.

8.5.4 Indexování

Pro další zpracování binárního obrazu může být někdy výhodné provést tzv. indexování oblastí. Indexováním rozumíme, že každé oblasti (každému rozpoznávanému objektu) je přiřazeno přirozené číslo, které je pak zapsáno do každého pixelu oblasti (obr.8.28). Za jednu oblast při tom považujeme množinu vzájemně sousedících pixelů, v nichž byla v původním binárním obraze zapsána hodnota 1. Při rozhodování, zda pixely sousedí, rozlišujeme tzv. čtyřsousednost a osmisousednost. Představme si pixely jako čtverečky podobně, jak je vyobrazeno na obr. 8.28. Při čtyřsousednosti považujeme dva pixely za sousedící, jestliže mají společnou stranu. Každý pixel tak má čtyři sousedy. V případě osmisousednosti považujeme za sousedy pixelu všech osm pixelů, které s daným pixellem sousedí hranou nebo pouze rohem. Indexovaného obrazu lze využít např. při praktické implementaci výpočtu atributů objektů při příznakovém rozpoznávání.



Obr.8.28. Indexování oblastí.

9 Měření objektů pro příznakové rozpoznání

V této kapitole předpokládáme, že obraz byl segmentován a že objekty, které mají být rozpoznány, byly úspěšně separovány. Příznakové metody rozpoznání jsou založeny na tom, že každý objekt, který má být rozpoznán, je popsán pomocí vhodných číselných hodnot, tzv. příznaků. Zpravidla je zapotřebí použít více příznaků, které vytvářejí vektor příznaků. Vektor příznaků vypočítaný pro rozpoznávaný objekt nese všechny podstatné informace o objektu a je jedinou informací pro jeho následné rozpoznání (při rozpoznání pak již nepracujeme s obrazy objektů, ale pouze s jejich vektory příznaků). Rozhodnout o tom, jaké veličiny by měly být použity jako příznaky vedoucí k rozpoznání, je obtížné. Neexistuje jednoznačný návod, jak tento problém řešit. Nejobecnějším pravidlem je, že počet příznaků a příznaky samotné by měly být voleny tak, aby od sebe spolehlivě dokázaly odlišit objekty jednotlivých tříd (kapitola 10). Na druhé straně by počet příznaků měl být rozumný, tedy ne příliš velký (za prakticky přijatelnou hodnotu lze považovat počet nepřekračující 10-20 příznaků). Je pravidlem, že při řešení jednotlivých problémů je zapotřebí provést teoretický návrh i experimentální ověření zvolené množiny příznaků. Obecně lze říci, že příznak je pro rozpoznání užitečný, jestliže splňuje následující vlastnosti: Hodnoty příznaku by měly vycházet podobně pro objekty jedné třídy a naopak dosti odlišné pro objekty různých tříd. Hodnoty příznaku by měly být pokud možno nezávislé na hodnotách jiných příznaků, které jsou pro rozpoznání také použity. V následujících podkapitolách uvedeme několik typických příkladů, jakých příznaků lze pro rozpoznání použít.

9.1 Momenty

Momenty různého stupně patří k příznakům, které jsou používány dosti často. Důvodem je, že jejich schopnost rozlišit od sebe objekty různých tříd často vyhovuje a jejich výpočet je snadný. Moment vztahený k souřadné soustavě obrazu je definován vztahem

$$m_{p,q} = \iint_{\Omega} x^p y^q f(x,y) dx dy, \quad (9.1)$$

kde $f(x,y)$ je obrazová funkce, p,q udávají stupeň momentu a Ω je ta část obrazu, kterou považujeme za rozpoznávaný objekt. (Poznamenejme, že zde uvádíme vztahy platné pro spojitou obrazovou funkci. Zápis vztahů platných pro funkci diskrétní jistě nebude čtenáři činit potíže - integraci postačí nahradit sumací.) Hodnoty p,q lze volit $p \geq 0, q \geq 0$. Moment $m_{0,0}$ je plocha objektu vážená hodnotou jasu. V úlohách, kde je plocha objektu významná, může být i moment $m_{0,0}$ použit jako jeden z příznaků pro rozpoznání. Protože zbývající momenty $m_{p,q}$ jsou závislé na umístění rozpoznávaného objektu v obraze, dává se zpravidla přednost výpočtu momentů vzhledem k osám procházejícím těžištěm objektu. K výpočtu polohy (x_t, y_t) těžiště objektu lze použít momentů $m_{0,0}, m_{1,0}, m_{0,1}$. Máme

$$x_t = \frac{m_{1,0}}{m_{0,0}}, \quad y_t = \frac{m_{0,1}}{m_{0,0}}. \quad (9.2)$$

Také poloha objektu (reprezentovaná polohou těžiště) může být v některých úlohách významná, a proto i hodnoty x_t, y_t někdy mohou být použity jako příznaky pro rozpoznání. Moment $\mu_{p,q}$ vzhledem k osám procházejícím těžištěm objektu je definován vztahem

$$\mu_{p,q} = \iint_{\Omega} (x - x_t)^p (y - y_t)^q f(x,y) dx dy. \quad (9.3)$$

Momenty $\mu_{p,q}$ k těžištním osám sice nezávisí na poloze objektu v obraze, avšak závisí na velikosti objektu i na jeho orientaci (natočení). Tam, kde jsou velikost objektu a jeho natočení významné, lze k rozpoznání použít přímo momentů $\mu_{p,q}$. Pokud není pro rozpoznání rozhodující velikost objektu, je možné použít normalizovaných momentů $\nu_{p,q}$

$$\nu_{p,q} = \frac{\mu_{p,q}}{(m_{0,0})^\gamma}, \quad \text{kde } \gamma = \left(\frac{p+q}{2} \right) + 1. \quad (9.4)$$

Jestliže při rozpoznání objektu nemá záležet ani na jeho orientaci (natočení), pak je nutné momenty počítat k hlavním osám objektu. Hlavní osy vytvářejí souřadnou soustavu, která má počátek v těžišti rozpoznávaného objektu a vzhledem k souřadné soustavě obrazu je pootočená o úhel θ . Pootočení θ je dáno podmínkou, aby momenty $\mu'_{2,0}$, $\mu'_{0,2}$ vypočtené k hlavním osám nabyly extrémních hodnot. Ukážeme, jak lze úhel θ stanovit. Rovnici přímky, která je první hlavní osou, hledáme ve tvaru

$$x \sin \theta - y \cos \theta + \rho = 0. \quad (9.5)$$

Hodnotami x, y při tom rozumíme souřadnice měřené v souřadné soustavě obrazu. Vzdálenost r bodu o souřadnicích x, y od první hlavní osy je

$$r = x \sin \theta - y \cos \theta + \rho. \quad (9.6)$$

Pro moment setrvačnosti k této ose pak vychází

$$\mu'_{0,2} = \iint_{\Omega} r^2 f(x, y) dx dy. \quad (9.7)$$

Po úpravě máme
$$\mu'_{0,2} = \iint_{\Omega} (x \sin \theta - y \cos \theta + \rho)^2 f(x, y) dx dy. \quad (9.8)$$

Zavedme následující substituci (souřadnice s vlnovkou jsou měřeny vzhledem k soustavě, jejíž počátek je posunut do těžiště objektu):

$$\tilde{x} = x - x_t, \quad \tilde{y} = y - y_t, \quad \text{a tedy} \quad x = \tilde{x} + x_t, \quad y = \tilde{y} + y_t. \quad (9.9)$$

Z rovnice (9.8) pak máme
$$\mu'_{0,2} = \iint_{\Omega} (\tilde{x} \sin \theta - \tilde{y} \cos \theta)^2 f(x, y) dx dy. \quad (9.10)$$

Položením $\partial \mu'_{0,2} / \partial \theta = 0$ obdržíme
$$\operatorname{tg} 2\theta = \frac{2\mu_{1,1}}{\mu_{2,0} - \mu_{0,2}}, \quad (9.11)$$

kde
$$\mu_{2,0} = \iint_{\Omega} \tilde{x}^2 f(x, y) dx dy, \quad \mu_{1,1} = \iint_{\Omega} \tilde{x} \tilde{y} f(x, y) dx dy, \quad \mu_{0,2} = \iint_{\Omega} \tilde{y}^2 f(x, y) dx dy. \quad (9.12)$$

Jak je patrné z rovnice (9.11), lze alternativně definovat hlavní osy objektu jako osy, ke kterým je moment $\mu_{1,1}$ (tzv. deviační moment) roven nule. Vztah (9.11) je užitečný i tehdy, když je natočení objektu pro rozpoznání významné. Úhel θ lze totiž chápat jako úhel popisující natočení objektu v obraze a lze jej použít jako jeden z příznaků. Na závěr tohoto odstavce o momentech poznamenejme, že kromě přímého použití momentů různého stupně lze jako příznaků pro rozpoznání použít též různých hodnot z momentů odvozených. Jako příklad uveďme podíl $\mu'_{2,0} / \mu'_{0,2}$, který charakterizuje podlouhlost.

9.2 Pravoúhlost a podlouhlost

Hranici vyšetřovaného objektu postupně rotujeme v rozsahu 0-90°. Krok úhlu rotace volíme v jednotkách stupňů (např. 5°). Po provedení rotace opíšeme kolem hranice objektu pravoúhelník, jehož strany jsou rovnoběžné se stranami obrazu (to je jednoduché - po provedení rotace stačí nalézt extrémní souřadnice hranice). Ze všech pravoúhelníků vybereme ten, který má nejmenší plochu. Plochu objektu označme A_O , plochu nejmenšího pravoúhelníka a délky jeho stran označme postupně A_R, a, b . Pravoúhlost (R) a podlouhlost (S) jsou pak definovány vztahy

$$R = \frac{A_O}{A_R}, \quad S = \frac{a}{b}. \quad (9.13)$$

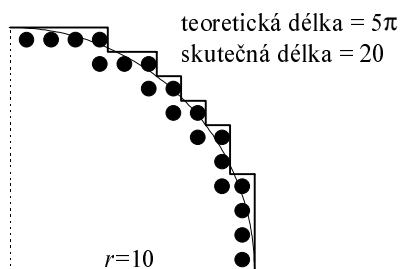
9.3 Kruhovitost

Nechť P, A označují délku hranice objektu a jeho plochu. Kruhovitost je pak definována vztahem

$$C = \frac{P^2}{A}. \quad (9.14)$$

Pro kruh vychází hodnota kruhovosti $C=4\pi$, pro čtverec je $C=16$, pro objekty nepravidelného tvaru jsou hodnoty vyšší. Poznamenejme, že hodnota 4π pro kruh vychází z teoretické délky hranice. V závislosti na výpočetním postupu se může skutečná délka hranice zjištěná v digitalizovaném obraze

od délky teoretické lišit. Např. na obr. 9.1 je teoretická délka čtveroblouku 5π . Prakticky zjištěná délka jeho digitalizovaného obrazu je však 20. Pokud by popsany jev byl při rozpoznávání na závalu, lze mu čelit vyhlazením tvaru hranice. Je například možné jistým počtem pixelů hranice proložit křivku. Délku této křivky lze pak chápat jako délku odpovídající části hranice.



Obr. 9.1. K rozdílu mezi teoretickou a skutečnou délkou hranice.

9.4 Energie hranice

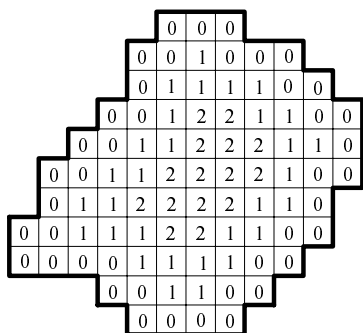
Předpokládejme, že známe průběh křivosti $k(s)$ hranice objektu. Délka obvodu objektu nechť je P . Příznak „energie hranice“ koncentruje informaci o průběhu funkce $k(s)$ do jediné hodnoty

$$E = \frac{1}{P} \int_0^P [k(s)]^2 ds. \quad (9.15)$$

Uvažujeme-li různé objekty se stejnou plochou, pak nejmenší energie hranice vychází pro kruh, a to $E = (1/R)^2$. Pro složitější tvary hranice jsou hodnoty energie vyšší.

9.5 Průměrná vzdálenost pixelu od hranice

Nechť d_i označuje vzdálenost i -tého pixelu objektu od hranice objektu, jehož plocha je A . Protože by bylo zdlouhavé pro každý pixel objektu vypočítávat vzdálenost euklidovskou, rozumíme zde vzdáleností počet řad pixelů, které leží mezi hranicí a uvažovaným pixelem (obr. 9.2). Průměrnou vzdáleností μ_d pixelu od hranice pak rozumíme hodnotu



Obr. 9.2. Vzdálenost pixelu od hranice.

$$\mu_d = \frac{1}{N} \sum_{i=1}^N d_i, \quad (9.16)$$

kde N je celkový počet pixelů v objektu. Koncentrovanou informaci o tvaru objektu nese pak např. příznak

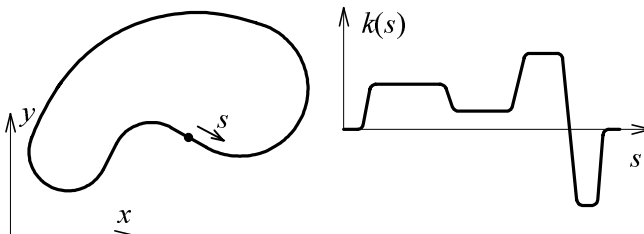
$$S = \frac{A}{\mu_d^2}. \quad (9.17)$$

9.6 Popis tvaru objektu pomocí průběhu křivosti jeho hranice

Z elementární teorie křivek je známo, že průběh $k(s)$ křivosti křivku jednoznačně určuje až na její polohu (uvažujeme zde křivky rovinné a pod pojmem křivost máme na mysli první křivost). Nabízí se možnost využít této skutečnosti tak, že pro rozpoznávání objektů využijeme příznaků odvozených z průběhu křivosti jejich hranice. Obr. 9.3 ukazuje křivku (hranici objektu) a odpovídající průběh $k(s)$ křivosti. Je zřejmé, že funkce $k(s)$ je periodická s periodou P , což je délka křivky. Funkci $k(s)$ lze proto rozvinout ve Fourierovu řadu

$$k(s) = \sum_{-\infty}^{\infty} c_n \exp\left[j2\pi n \frac{s}{P}\right], \quad (9.18)$$

$$\text{kde } c_n = \frac{1}{P} \int_0^P k(s) \exp\left[-j2\pi n \frac{s}{P}\right] ds. \quad (9.19)$$



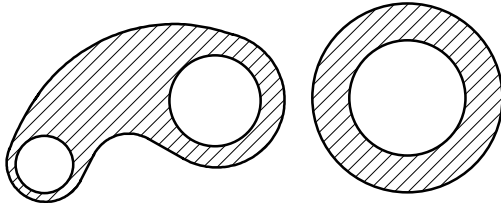
Obr. 9.3. Objekt a průběh křivosti po obvodě.

Zpravidla lze hlavní tvary objektu popsat pomocí amplitud jistého počtu nejnižších frekvenčních složek. Vyšší složky popisují pro rozpoznání zpravidla méně významné tvarové detaily; může se též jednat o šumy. Prakticky lze průběh křivosti hranice získat např. tak, že skupinami po sobě jdoucích pixelů tvořících hranici prokládáme kruhové oblouky nebo jinou vhodnou křivku. V případě kruhového oblouku platí mezi jeho

poloměrem a křivostí jednoduchý vztah $k=1/r$. V případě aproximace hranice objektu množinou kruhových oblouků má pak funkce $k(s)$ po částech konstantní průběh. Prokládáme-li obecnější křivku s parametrickou rovnicí $\mathbf{x}=\mathbf{x}(s)$, pak pro křivost máme $k(s)=|\mathbf{x}''(s)|$ (z diferenciální geometrie křivek je známo, že uvedený vztah platí pro takovou parametrizaci křivky, kde parametr s měří délku oblouku). Poznamenejme, že kromě prokládání kruhových oblouků nebo křivek, byla navržena také řešení, kde je informace o průběhu křivosti získávána analýzou řetězce popisujícího tvar hranice ve Freemanově kódu (kap. 8.2.1, obr. 8.18).

9.7 Eulerovo číslo

Necht' C je počet navzájem nesouvislých částí rozpoznávaného objektu a H necht' je počet děr v objektu (obr.9.4). Eulerovo číslo je rozdíl $C-H$. Obvyklé metody segmentace zpravidla neumožňují detekovat objekt s více než jednou částí jako objekt jediný, ale každou část detekují jako objekt samostatný (na rozdíl od obr. 9.4). Eulerovo číslo je proto v tomto případě $1-H$. Je samozřejmé, že užitečným příznakem pro rozpoznání může být také jednoduše pouze počet děr v rozpoznávaném objektu, který s Eulerovým číslem úzce souvisí.



Obr. 9.4. Objekt se dvěma nesouvislými částmi a třemi děrami ($C=2, H=3$).

detekovat objekt s více než jednou částí jako objekt jediný, ale každou část detekují jako objekt samostatný (na rozdíl od obr. 9.4). Eulerovo číslo je proto v tomto případě $1-H$. Je samozřejmé, že užitečným příznakem pro rozpoznání může být také jednoduše pouze počet děr v rozpoznávaném objektu, který s Eulerovým číslem úzce souvisí.

9.8 Atributy odvozené z histogramu jasu

Jestliže máme rozpoznávat objekty, pro něž je charakteristické jisté rozložení jasu po ploše objektu, nebo objekty, které jsou pokryty texturou, pak může být užitečné zvolit jako příznaky sloužící k rozpoznání také hodnoty odvozené z jasu rozpoznávaného objektu. Připomeňme, že rozložení jasu po ploše objektu bylo jistým způsobem vzato v úvahu již při výpočtu momentů. V tomto odstavci ukážeme další možnosti. Označme $b \in \{0, 1, 2, \dots, m\}$ jas pixelu (m je maximální možný jas). Pro rozpoznávaný objekt vypočítejme histogram zastoupení jednotlivých hodnot jasů v objektu. Označme N celkový počet pixelů rozpoznávaného objektu a $N(b)$ počet pixelů v objektu, které mají jas právě b . Provedme normalizaci histogramu takto: $p(b)=N(b)/N$. Hodnotu $p(b)$ lze interpretovat jako pravděpodobnost jevu, že pixel objektu má jas právě b . Funkce $p(b)$ podrobně charakterizuje zastoupení jednotlivých jasů v objektu. Informaci obsaženou ve funkci $p(b)$ je žádoucí koncentrovat do pokud možno malého množství příznaků, které mohou být použity k rozpoznání. Interpretace funkce $p(b)$ jako hustoty pravděpodobnosti umožňuje využít postupů známých z teorie pravděpodobnosti: Jako příznaků vedoucích k rozpoznání lze použít zejména střední hodnoty a momentů různého stupně. Nejběžněji používané příznaky uvádíme v následujícím přehledu.

$$\text{Střední hodnota jasu:} \quad \mu_b = \sum_{b=0}^m b p(b). \quad (9.20)$$

$$\text{Variance:} \quad \sigma_b^2 = \sum_{b=0}^m (b - \mu_b)^2 p(b). \quad (9.21) \quad \text{Šikmost:} \quad s_b = \frac{1}{\sigma_b^3} \sum_{b=0}^m (b - \mu_b)^3 p(b). \quad (9.22)$$

$$\text{Energie:} \quad E_b = \sum_{b=0}^m p^2(b). \quad (9.23) \quad \text{Entropie:} \quad T_b = \sum_{b=0}^m p(b) \log(p(b)). \quad (9.24)$$

Poznamenejme, že z provedených výzkumů vyplývá, že lidské oko jen s obtížemi rozlišuje textury, které se shodují v momentech do druhého řádu včetně a liší se pouze v momentech vyšších řádů. To však neznamená, že by uvedené charakteristiky nemohly být využity pro rozpoznání počítačové, pokud jsou difference v momentech vyšších řádů skutečně významné.

9.9 Atributy odvozené z frekvenčního spektra jasu

Pro rozpoznání objektů s texturou lze též využít příznaků odvozených z frekvenčního spektra jasu rozpoznávaného objektu. Provedeme Fourierovu transformaci obrazové funkce nad oblastí, kterou

rozpoznávaný objekt zaujímá. Jistou komplikací je, že rozpoznávaný objekt může mít libovolný tvar. Jednoduchý postup, který umožňuje aplikovat běžné algoritmy Fourierovy transformace (zejména rychlé), spočívá v tom, že do rozpoznávaného objektu vepíšeme obdélník nebo čtverec vhodných rozměrů a provedeme Fourierovu transformaci obrazové funkce nad tímto obdélníkem (alternativně by bylo možné vepsat obdélníků více tak, aby pokryly pokud možno celý rozpoznávaný objekt a jednotlivá získaná frekvenční spektra jistým způsobem průměrovat). Vepsáním obdélníku jsme získali vzorek textury, kterou je objekt pokryt. Víme, že Fourierův obraz nese všechny informace obsažené ve vzoru. Pro účely rozpoznání je žádoucí informaci obsaženou ve Fourierově obrazu opět nějakým způsobem koncentrovat do pokud možno malého počtu příznaků. Lze to provést např. takto: Nechť $f(x,y)$ je obrazová funkce nad obdélníkem vepsaným do rozpoznávaného objektu a $F(u,v)$ nechť je její Fourierův obraz. Jako příznaky lze použít např. následující veličiny

$$U_{p,q} = \int_{-\infty}^{\infty} \int_p^q F(u,v) du dv, \quad V_{p,q} = \int_p^q \int_{-\infty}^{\infty} F(u,v) du dv, \quad (9.25)$$

kde p,q jsou zvolené reálné hodnoty. Je zřejmé, že hodnoty $U_{p,q}, V_{p,q}$ vyjadřují, jak jsou zastoupeny frekvence v pásech širokých $|p-q|$ rovnoběžných s osami v,u . Pro rozpoznání lze samozřejmě zvolit libovolný počet takových pásů. Alternativně lze hodnotu $F(u,v)$ samozřejmě také integrovat nad jinou vhodnou oblastí - např. nad obdélníkovou oblastí konečných rozměrů. Vyjádřeme dále u,v v polárních souřadnicích takto: $u = \rho \cos \varphi, v = \rho \sin \varphi$. Nechť $F(\rho, \varphi)$ je Fourierův obraz popsáný v těchto polárních souřadnicích. Jako příznaky lze v tomto případě použít následujících hodnot

$$A_{p,q} = \int_p^q \int_0^{2\pi} F(\rho, \varphi) d\varphi d\rho, \quad B_{s,t} = \int_0^t \int_s^{\infty} F(\rho, \varphi) d\varphi d\rho. \quad (9.26)$$

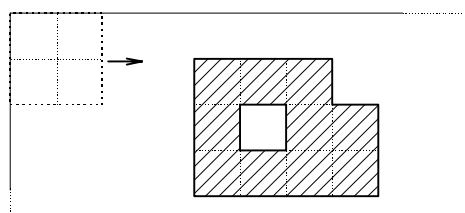
Hodnota $A_{p,q}$ vyjadřuje zastoupení frekvencí v mezikruží o šířce $|p-q|$. Příznaku $A_{p,q}$ je na místě použít tehdy, jestliže chceme rozpoznávat objekty pokryté texturou, která může být vzhledem ke stranám obrazu libovolně natočena. Je-li naopak natočení textury významné a měla-li by velikost natočení sloužit k rozpoznání, pak je na místě použít příznaku $B_{s,t}$, který vyjadřuje zastoupení frekvencí v klínu o šířce $|t-s|$ úhlových jednotek.

9.10 Algoritmus výpočtu plochy, obvodu a Eulerova čísla

Ačkoli podrobná implementace výpočtu dosud uvedených příznaků není příliš obtížná, popíšeme v tomto odstavci jednoduchý postup, kterého lze použít při výpočtu plochy, obvodu a Eulerova čísla. Předností postupu je, že nevyžaduje explicitní určení hranice, a je proto výhodný tehdy, jestliže byly segmentací určeny přímo plochy náležící jednotlivým objektům (např. bylo-li použito prahování). Postup spočívá v tom, že zjišťujeme, kolikrát je možné do různých míst objektu umístit matice obsažené v dále uvedených množinách Q_1, Q_2, Q_3, Q_4, Q_D :

$$Q_1 = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right\}, \quad Q_2 = \left\{ \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \right\}, \quad (9.27, 28)$$

$$Q_3 = \left\{ \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \right\}, \quad Q_4 = \left\{ \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right\}, \quad Q_D = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right\}. \quad (9.29, 30, 31)$$



Obr. 9.5. Výpočet plochy, obvodu a Eulerova čísla ($n(Q_1)=5, n(Q_2)=8, n(Q_3)=5, n(Q_4)=1, n(Q_D)=0, A=10, P=18, E=0$).

Prakticky můžeme výpočet provést tak, že začneme v některém rohu obrazu a dále systematicky prověřujeme po řádcích a po sloupcích všechny možné pozice v obraze tak, že do nich umísťujeme matice z množin Q_1, Q_2, Q_3, Q_4, Q_D . Za výskyt matice v objektu považujeme situaci, kdy se pixely objektu vyskytují právě na místech jedniček v uvažované matici. Nechť $n(Q_1), n(Q_2), n(Q_3), n(Q_4), n(Q_D)$ označují počty výskytů matic z množin $Q_1, Q_2, Q_3,$

Q_4 , Q_D v objektu. Pro plochu, obvod a Eulerovo číslo objektu pak platí následující vztahy:

$$A = \frac{1}{4} [n(Q_1) + 2n(Q_2) + 3n(Q_3) + 4n(Q_4) + 2n(Q_D)], \quad (9.32)$$

$$P = n(Q_1) + n(Q_2) + n(Q_3) + 2n(Q_D), \quad (9.33)$$

$$E = \frac{1}{4} [n(Q_1) - n(Q_3) + 2n(Q_D)] \quad (\text{pro čtyřsousednost}), \quad (9.34)$$

$$E = \frac{1}{4} [n(Q_1) - n(Q_3) - 2n(Q_D)] \quad (\text{pro osmisousednost}). \quad (9.35)$$

Popsaný postup je ilustrován na obr. 9.5. Čárkovanou čarou je znázorněna počáteční poloha matice 2×2 v levém horním rohu obrazu. Šipka znázorňuje posouvání polohy po sloupcích a po řádcích. Poznamenáváme, že délka obvodu zahrnuje také délku obvodu případných děr.

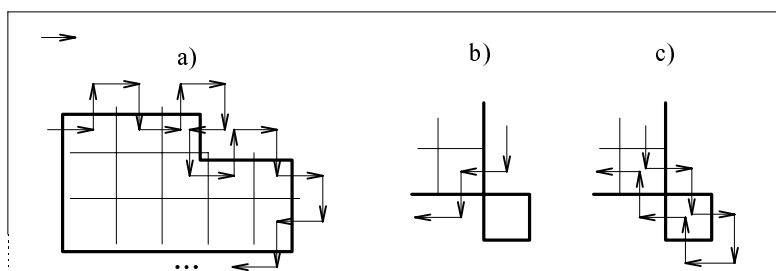
Jestliže objekty, které mají být rozpoznány, byly původně hladké, pak zkreslení velikosti plochy a obvodu, které bylo způsobeno digitalizací, lze do jisté míry kompenzovat použitím modifikovaných vztahů pro plochu a obvod

$$A = \frac{1}{4} n(Q_1) + \frac{1}{2} n(Q_2) + \frac{7}{8} n(Q_3) + n(Q_4) + \frac{3}{4} n(Q_D), \quad (9.36)$$

$$P = n(Q_2) + \frac{1}{\sqrt{2}} [n(Q_1) + n(Q_3) + 2n(Q_D)]. \quad (9.37)$$

9.11 Algoritmus určení hranice

Jestliže byly segmentací určeny přímo plochy náležící jednotlivým objektům (např. bylo-li použito prahování), pak může být pro výpočet příznaků někdy zapotřebí explicitně určit také hranice objektů. Ne vždy totiž vystačíme pouze s délkou hranice, kterou lze postupem podle předchozího odstavce stanovit i bez jejího explicitního určení. Určení hranice lze provést sledováním pomocí následujícího algoritmu: 1) Procházej pixely obrazu po řádcích a po sloupcích tak dlouho, až narazíš na první pixel objektu (tento pixel je nutně pixelem ležícím na hranici objektu). 2) Sleduj dále hranici objektu podle následujícího pravidla: 2a) Nacházíš-li se uvnitř objektu, otoč se doleva a posuň se o jeden pixel. 2b)



Obr. 9.6. Stanovení hranice jejím sledováním.

Nacházíš-li se vně objektu, pak se otoč doprava a posuň se o jeden pixel. Počáteční směr v prvním nalezeném pixelu objektu je zleva doprava. Během činnosti algoritmu jsou procházeny body ležící na hranici objektu (obr. 9.6a). Algoritmus končí nalezením uzavřené hranice. Poznamenejme, že právě popsany algoritmus nalezení hranice je

sice jednoduchý, avšak ne vždy je získaný výsledek zcela uspokojivý. Jisté problémy může způsobit skutečnost, že algoritmus prochází některé pixely hranice dvakrát. Obsahuje-li hranice pixely sousedící pouze rohem (což může nastat, jestliže uvažujeme osmisousednost), pak navíc může dojít k chybnému procházení a ztrátě jisté části hranice. Situaci ilustruje obr. 9.6. Na obr. 9.6c je správně nalezena celá hranice, zatímco na obr. 9.6b byla při procházení část hranice vynechána. Uvedeným potížím lze sice čelit pečlivým ošetřením speciálních případů, ale algoritmus se tak ovšem bohužel komplikuje.

9.12 Hodnocení vhodnosti zvolené množiny příznaků

Na závěr této kapitoly se pokusíme o diskusi, jak lze hodnotit vhodnost volby příznaků pro rozpoznání. Pro jednoduchost předpokládejme, že máme pouze dva příznaky, které zde označíme x , y . Předpokládejme dále, že máme k dispozici vhodnou trénovací množinu. Můžeme proto vypočítat střední

hodnoty a variance obou příznaků v jednotlivých třídách (v následujících výrazech je i indexem třídy, stříškou vyznačujeme, že se jedná o odhad vypočítaný na základě trénovací množiny, N_i je počet hodnot získaných v jednotlivých třídách; podrobnější výklad pojmu třída lze nalézt v následující kapitole)

$$\hat{\mu}_{x,i} = \frac{1}{N_i} \sum_{j=1}^{N_i} x_{i,j}, \quad \hat{\mu}_{y,i} = \frac{1}{N_i} \sum_{j=1}^{N_i} y_{i,j}, \quad (9.38)$$

$$\hat{\sigma}_{x,i}^2 = \frac{1}{N_i} \sum_{j=1}^{N_i} (x_{i,j} - \hat{\mu}_{x,i})^2, \quad \hat{\sigma}_{y,i}^2 = \frac{1}{N_i} \sum_{j=1}^{N_i} (y_{i,j} - \hat{\mu}_{y,i})^2. \quad (9.39)$$

Jak jsme již uvedli v úvodu této kapitoly, měly by být hodnoty příznaků pro objekty patřící stejným třídám podobné. Hodnoty variancí ve vztazích (9.39) by proto měly být co nejmenší. Zvolené příznaky by dále měly být pokud možno nezávislé. Závislost či nezávislost příznaků lze popsat pomocí kovariance. Normalizovaná kovariance je definována vztahem

$$\hat{\sigma}_{xy,i}^2 = \frac{1}{N_i \hat{\sigma}_{x,i} \hat{\sigma}_{y,i}} \sum_{j=1}^{N_i} (x_{i,j} - \hat{\mu}_{x,i})(y_{i,j} - \hat{\mu}_{y,i}). \quad (9.40)$$

Hodnoty normalizované kovariance vycházejí z intervalu $\langle -1, 1 \rangle$. Hodnota 0 vyjadřuje, že příznaky jsou nezávislé. Jiné hodnoty vyjadřují závislost. Čím větší je absolutní hodnota kovariance, tím větší je závislost. Hodnota 1 např. vyjadřuje fakt, že se jedná o příznaky zcela závislé. V takovém případě je ovšem použití obou příznaků nadbytečné - bez újmy na kvalitě rozpoznání by bylo možné jeden z příznaků vypustit. (Také v případě hodnoty -1 se jedná o naprostou závislost příznaků. Odchytky hodnot příznaků od jejich střední hodnoty však mají opačné znaménko.) Schopnost příznaků rozlišovat (separovat) mezi jednotlivými třídami lze hodnotit pomocí normalizované vzdálenosti mezi třídami, kterou lze vyjádřit následujícími vztahy:

$$\hat{D}_{x,i,j} = \frac{|\hat{\mu}_{x,i} - \hat{\mu}_{x,j}|}{\sqrt{\hat{\sigma}_{x,i}^2 + \hat{\sigma}_{x,j}^2}}, \quad \hat{D}_{y,i,j} = \frac{|\hat{\mu}_{y,i} - \hat{\mu}_{y,j}|}{\sqrt{\hat{\sigma}_{y,i}^2 + \hat{\sigma}_{y,j}^2}}. \quad (9.41)$$

K úspěšnému rozpoznání je důležité, aby uvedené hodnoty separačních vzdáleností mezi třídami byly pokud možno co největší.

9.13 Karhunen-Loéveho transformace

Předpokládejme, že \mathbf{x} je náhodný vektor dimenze n (tedy sloupcový vektor obsahující n náhodných proměnných). Nechť $\boldsymbol{\mu}_x$ je vektor středních hodnot a \mathbf{C}_x matice kovariance pro tento náhodný vektor.

Je tedy
$$\boldsymbol{\mu}_x = E\{\mathbf{x}\}, \quad \mathbf{C}_x = E\{(\mathbf{x} - \boldsymbol{\mu}_x)(\mathbf{x} - \boldsymbol{\mu}_x)^T\}. \quad (9.42)$$

Nyní vytvořme náhodný vektor \mathbf{y} lineární transformací

$$\mathbf{y} = \mathbf{A}(\mathbf{x} - \boldsymbol{\mu}_x), \quad (9.43)$$

kde \mathbf{A} je matice, která ve svých řádcích obsahuje vlastní vektory \mathbf{v}_i matice \mathbf{C}_x . Pro další výklad bude užitečné, budeme-li předpokládat, že vlastní vektory jsou do řádků matice zapsány v pořadí podle klesající hodnoty odpovídajícího vlastního čísla λ_i . Pro vektor \mathbf{y} nyní vypočítáme vektor středních hodnot $\boldsymbol{\mu}_y$ a matici kovariance \mathbf{C}_y . Dostaneme

$$\boldsymbol{\mu}_y = E\{\mathbf{A}(\mathbf{x} - \boldsymbol{\mu}_x)\} = \mathbf{0}, \quad (9.44)$$

$$\mathbf{C}_y = E\{(\mathbf{y} - \boldsymbol{\mu}_y)(\mathbf{y} - \boldsymbol{\mu}_y)^T\} = \mathbf{A}\mathbf{C}_x\mathbf{A}^T. \quad (9.45)$$

Protože jednotlivé řádky matice \mathbf{A} jsou vlastními vektory \mathbf{v}_i matice \mathbf{C}_x , protože platí $\mathbf{C}_x \mathbf{v}_i = \lambda_i \mathbf{v}_i$ a konečně protože vlastní vektory tvoří ortonormální bázi (to platí v případě, že vlastní čísla jsou různá, v případě vícenásobných vlastních čísel můžeme však provést ortonormalizaci), vychází

$$\mathbf{C}_y = \mathbf{A} \mathbf{C}_x \mathbf{A}^T = \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \dots \\ \mathbf{v}_n^T \end{bmatrix} \mathbf{C}_x [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \dots \quad \mathbf{v}_n] = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}. \quad (9.46)$$

Zopakujme, že \mathbf{C}_y je kovarianční matice náhodného vektoru \mathbf{y} . Připomeňme, že diagonální prvky kovarianční matice obsahují variance jednotlivých složek vektoru a nediagonální prvky pak kovariance dvou různých složek. Dále připomeňme, že nulová hodnota kovariance dvou náhodných proměnných signalizuje jejich vzájemnou nekorelovanost. Výsledek, ke kterému jsme dospěli ve vztahu (9.46), tedy ukazuje, že lineární transformace (9.43) má za následek, že jednotlivé složky vektoru \mathbf{y} jsou nekorelované. Transformace tedy odstraňuje korelaci mezi jednotlivými složkami vektoru \mathbf{x} , vlastní čísla λ_i matice \mathbf{C}_x navíc udávají varianci jednotlivých složek transformovaného náhodného vektoru \mathbf{y} . Poznamenejme, že k transformaci (9.43) je též možné realizovat inverzní transformaci (platnost druhé rovnosti v následujícím vztahu vyplývá z faktu, že \mathbf{A} je ortonormální)

$$\mathbf{x} = \mathbf{A}^{-1} \mathbf{y} + \boldsymbol{\mu}_x = \mathbf{A}^T \mathbf{y} + \boldsymbol{\mu}_x. \quad (9.47)$$

Až doposud jsme se zabývali případem, kdy se při transformaci neměnil počet složek transformovaného vektoru. Vytvořme nyní matici \mathbf{B} rozměru $m \times n$ ($m < n$) tak, že jejích m řádků bude obsahovat vlastní vektory matice \mathbf{C}_x . Vybereme při tom ty vektory, které odpovídají největším vlastním číslům. Předpokládejme nyní pro jednoduchost, že $\boldsymbol{\mu}_x = \mathbf{0}$. Opět můžeme zřejmě provést lineární transformaci

$$\mathbf{y} = \mathbf{B} \mathbf{x}. \quad (9.48)$$

Analogicky ke vztahu (9.47) můžeme vypočítat

$$\hat{\mathbf{x}} = \mathbf{B}^T \mathbf{y}. \quad (9.49)$$

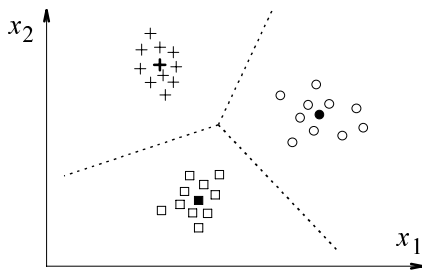
Vztah (9.48) ukazuje, že je možné vektor \mathbf{x} reprezentovat pomocí vektoru o menším počtu složek. Vztah (9.49) naznačuje, že na základě této reprezentace bude možné původní vektor opět zrekonstruovat. Samozřejmou podmínkou je, aby původní (\mathbf{x}) a zrekonstruovaná ($\hat{\mathbf{x}}$) hodnota si byly co nejbližší. Lze ukázat, že pro střední kvadratickou chybu (MSE) uvedené aproximace platí

$$MSE = \sum_{i=m+1}^n \lambda_i. \quad (9.50)$$

Podle uvedeného vztahu je chyba reprezentace rovna součtu vlastních čísel odpovídajících těm vlastním vektorům, které nebyly zahrnuty do matice \mathbf{B} . Při praktické realizaci proto do matice \mathbf{B} zahrneme pouze vlastní vektory odpovídající dostatečně velkým vlastním číslům. Vlastní vektory odpovídající malým vlastním číslům lze vynechat, protože mají na přesnost reprezentace jen malý vliv.

10 Příznakové metody analýzy obrazu

Příznakové metody analýzy obrazu předpokládají, že máme k dispozici popis rozpoznávaného objektu ve formě vektoru příznaků $\mathbf{x}=(x_1,x_2,\dots, x_m)^T$. Jednotlivými složkami tohoto vektoru jsou hodnoty vhodně zvolených příznaků (kapitola 9). Podle oboru hodnot, kterých mohou příznaky nabývat, je možné rozlišovat spojité, diskrétní a binární (logické, dichotomické) příznaky. Každý konkrétní vektor příznaků reprezentuje nějaký bod v m -rozměrném prostoru X^m , který zde budeme nazývat příznakovým prostorem. Příznakový prostor X^m je kartézským součinem oborů hodnot všech uvažovaných příznaků a je tedy tvořen všemi možnými hodnotami vektoru $\mathbf{x}=(x_1,x_2,\dots, x_m)^T$. Jsou-li např. hodnoty všech zvolených příznaků reálné, je X^m prostorem reálných m -tic. V konkrétních úlohách nemusí ovšem hodnoty příznaků nabývat všech teoreticky možných hodnot z X^m . Obvykle se stává, že je hodnota příznaků v konkrétní úloze nějak omezena. Často jsou např. hodnotami příznaků pouze čísla z jistého intervalu. Zavedeme proto obor χ hodnot příznaků, které se v uvažované úloze mohou vyskytovat. Samozřejmě platí $\chi \subseteq X^m$.



Obr.10.1. Obrazy objektů jednotlivých tříd tvoří shluky v prostoru příznaků.

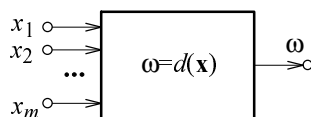
Mějme objekt O_1 , kterému odpovídá vektor příznaků \mathbf{x}_1 , a objekt O_2 , kterému odpovídá vektor příznaků \mathbf{x}_2 . Vektory $\mathbf{x}_1, \mathbf{x}_2$ reprezentují body v prostoru příznaků X^m . Příznakové metody analýzy obrazu jsou založeny na myšlence, že podobnost nebo shoda objektů O_1, O_2 se projeví tím, že délka vektoru $\mathbf{x}_1 - \mathbf{x}_2$ v prostoru X^m je malá nebo nulová. Množinu objektů stejného druhu nazýváme třídou. Pokud bychom konstruovali např. systém pro rozpoznávání dvojrozměrných geometrických objektů, pak by třídami byly jednotlivé obrazce, které chceme rozpoznávat, např. kružnice, čtverec, trojúhelník atd. Cílem rozpoznání je stanovení třídy objektu. Označíme n počet tříd, které mají být rozpoznávány. Jednotlivé třídy dále označme $\omega_i, 1 \leq i \leq n$. Množina všech tříd je pak $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$. Na obr.

10.1 je znázorněna situace, kdy rozpoznávání spočívá v zařazení rozpoznávaného objektu do jedné ze tří tříd. K rozpoznání je použito dvou příznaků x_1, x_2 . Značky znázorňují obrazy objektů jednotlivých tříd v prostoru příznaků. Je typické, že obrazy objektů jednotlivých tříd tvoří v prostoru příznaků shluky. Pro potřeby pozdějšího výkladu jsou dále silněji vyznačeny tzv. etalony a čárkovaně jsou vyznačeny hranice mezi třídami.

10.1 Klasifikátory

Klasifikátorem nazveme zobrazení $d: \chi \rightarrow \Omega$. Jedná se tedy zobrazení, které ke každému možnému vektoru příznaků přiřadí identifikátor třídy, do které rozpoznávaný objekt náleží. Můžeme tedy psát

$$\omega = d(\mathbf{x}). \quad (10.1)$$



Obr.10.2. Ilustrace činnosti klasifikátoru.

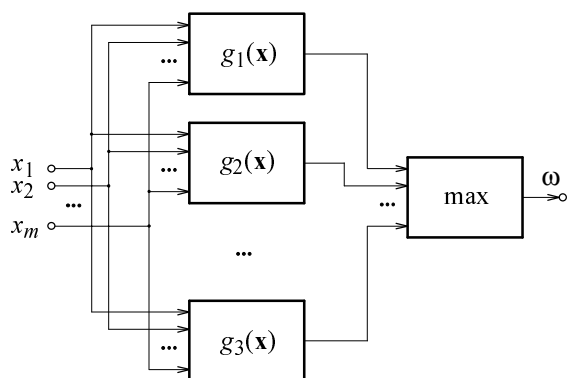
Schématicky je činnost klasifikátoru ilustrována na obr.10.2. Rozhodovací pravidlo, podle kterého se provádí klasifikace, indukuje rozklad oblasti χ na podoblasti. Podoblast χ_i ($\chi_i \subseteq \chi$) je tvořena těmi hodnotami vektoru \mathbf{x} , pro které platí $d(\mathbf{x})=\omega_i$. Základním problémem při konstrukci klasifikátoru je návrh rozhodovacího pravidla. Této otázce se dále budeme věnovat podrobněji. Podle realizace rozhodovacího kritéria můžeme klasifikátory charakterizovat jako klasifi-

kátory deterministické nebo nedeterministické, s pevným nebo s proměnným počtem příznaků, klasifikátory bez učení nebo s učením. V tomto textu se však omezíme pouze na nejjednodušší případy.

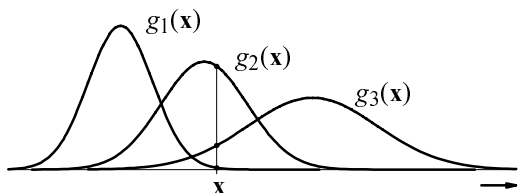
10.1.1 Klasifikace pomocí diskriminačních funkcí

Při klasifikaci pomocí diskriminačních funkcí předpokládáme, že existuje n reálných diskriminačních funkcí $g_1(\mathbf{x}), g_2(\mathbf{x}), \dots, g_n(\mathbf{x})$ definovaných nad \mathcal{X} . Rozpoznání je v tomto případě založeno na stanovení funkce dávající pro zadaný vektor příznaků maximální hodnotu. Hledaná třída je stanovena podle pravidla:

$$\text{jestliže pro zadané } \mathbf{x} \text{ platí } g_r(\mathbf{x}) > g_s(\mathbf{x}) \text{ pro každé } s \in (\{1, 2, \dots, n\} \setminus \{r\}), \text{ pak } d(\mathbf{x}) = \omega_r. \quad (10.2)$$



Obr.10.3. Schéma klasifikace pomocí diskriminačních funkcí.



Obr.10.4. Klasifikace diskriminačními funkcemi $g_r(\mathbf{x}) = p(\mathbf{x}|\omega_r)$.

Funkci klasifikátoru ilustruje obr.10.3. Známe-li funkce $g_1(\mathbf{x}), g_2(\mathbf{x}), \dots, g_n(\mathbf{x})$, pak jsou tím také určeny oblasti jednotlivých tříd. Máme

$$\chi_r = \{\mathbf{x} | d(\mathbf{x}) = \omega_r\}. \quad (10.3)$$

Hranici mezi dvěma sousedními oblastmi χ_r, χ_s lze stanovit řešením rovnice

$$g_r(\mathbf{x}) = g_s(\mathbf{x}). \quad (10.4)$$

Z předchozího výkladu je zřejmé, že klíčovým problémem je nalezení vhodné množiny diskriminačních funkcí. Tomuto problému budou věnovány následující odstavce. Zatím pro ilustraci pouze poznamenáváme, že jednoduchou volbou je např. $g_r(\mathbf{x}) = p(\mathbf{x}|\omega_r)$, kde $p(\mathbf{x}|\omega_r)$ je podmíněná hustota pravděpodobnosti pro hodnoty příznaků za předpokladu, že se jedná o objekt třídy ω_r . Pro jednorozměrný prostor příznaků je tato volba ilustrována na obr 10.4. V dalším textu tento přístup dále zdokonalíme. V případě klasifikace do dvou tříd (tzv. dichotomie) pracuje klasifikátor pouze se dvěma diskriminačními funkcemi. Pro nalezení maxima tak stačí zjistit znaménko rozdílu $g_1(\mathbf{x}) - g_2(\mathbf{x})$. Jestliže přidělíme třídám identifikátory 1, -1, lze klasifikátor jednoduše zkonstruovat podle předpisu

$$\omega = \text{sgn}(g_1(\mathbf{x}) - g_2(\mathbf{x})). \quad (10.5)$$

Proveďme ještě několik obecných úvah o tvaru diskriminačních funkcí. Nejjednodušším tvarem diskriminační funkce je funkce lineární, která má tvar

$$g_r(\mathbf{x}) = a_{r,0} + a_{r,1}x_1 + a_{r,2}x_2 + \dots + a_{r,m}x_m. \quad (10.6)$$

Na koeficient $a_{r,i}$ můžeme pohlížet jako na váhu příznaku x_i pro třídu ω_r , $a_{r,0}$ je práh, kterým je hodnota diskriminační funkce posouvána o konstatní velikost. Rovnici plochy v X^m , na níž leží hranice mezi oblastmi χ_r, χ_s , snadno určíme ze vztahu 10.4. Vyjde

$$(a_{r,0} - a_{s,0}) + (a_{r,1} - a_{s,1})x_1 + (a_{r,2} - a_{s,2})x_2 + \dots + (a_{r,m} - a_{s,m})x_m = 0. \quad (10.7)$$

Vidíme, že v případě lineárních diskriminačních funkcí je plocha tvořící hranici mezi třídami nadrovinou v prostoru X^m . Je zřejmé, že při n diskriminačních funkcích je nejvyšší možný počet takových nadrovin dán výrazem $n(n-1)/2$. V případě dichotomického klasifikátoru jsou k dispozici pouze dvě diskriminační funkce, a v prostoru příznaků existuje tudíž pouze jediná nadrovina, která je rozděljuje na dvě části. Pokud lze danou úlohu řešit za použití lineárních diskriminačních funkcí, tedy pokud je možné najít hranice oddělující oblasti odpovídající různým třídám tak, že jsou tyto hranice částmi

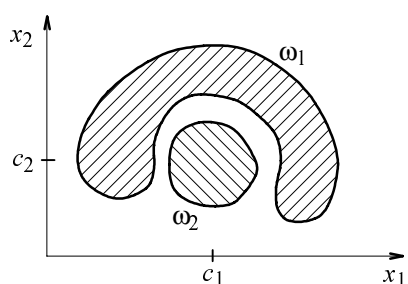
nadrovin, pak úlohu nazveme úlohou s lineárně separabilními třídami. Lineární separabilita je z hlediska praktické realizace samozřejmě výhodná. Nejsou-li třídy lineárně separabilní, pak lze někdy původní prostor X^m příznaků zobrazit do nového prostoru Y^q . Prostor Y^q a zobrazení $X^m \rightarrow Y^q$ vyhledáme tak, aby v Y^q bylo dosaženo lineární separability. Poznamenejme ovšem, že nalézt takový prostor a zobrazení nemusí být jednoduché a někdy to může být i nemožné. Odhlédněme od této komplikace a předpokládejme, že prostor a zobrazení požadovaných vlastností jsou známy a že platí (ξ je vektor příznaků v prostoru Y^q)

$$\xi = \Phi(\mathbf{x}). \quad (10.8a)$$

Po složkách tedy $\xi_1 = \Phi_1(\mathbf{x}), \xi_2 = \Phi_2(\mathbf{x}), \dots, \xi_q = \Phi_q(\mathbf{x}). \quad (10.8b)$

Pro diskriminační funkci $g_r(\mathbf{x})$ pak s využitím vztahů (10.6) a (10.8) máme

$$g_r(\xi) = g_r(\Phi(\mathbf{x})) = a_{r,0} + a_{r,1}\Phi_1(\mathbf{x}) + a_{r,2}\Phi_2(\mathbf{x}) + \dots + a_{r,q}\Phi_q(\mathbf{x}). \quad (10.9)$$



Obr.10.5. Vhodnou transformací lze někdy úlohu převést na lineárně separabilní problém.

Obr.10.5 ukazuje příklad, kdy by mohl být popsán postup užitečný. V příznakovém prostoru x_1, x_2 znázorněném na obrázku jistě nejsou třídy ω_1, ω_2 lineárně separabilní. Úlohu však snadno převedeme na lineárně separabilní problém transformací $\xi = \sqrt{[(x_1 - c_1)^2 + (x_2 - c_2)^2]}$. Klasifikátor bude pracovat pouze s jediným příznakem. V případech, které nelze vyřešit podobným způsobem a jestliže lineární diskriminační funkce nepostačuje, lze někdy využít následujícího postupu. Funkci $g_r(\mathbf{x})$ reprezentujeme množinou $\{g_{r,k}(\mathbf{x})\}$ lineárních funkcí. Hodnotu diskriminační funkce $g_r(\mathbf{x})$ pak stanovujeme podle vztahu

$$g_r(\mathbf{x}) = \max_k \{g_{r,k}(\mathbf{x})\}. \quad (10.10)$$

10.1.2 Klasifikace pomocí minimální vzdálenosti

Při klasifikaci pomocí minimální vzdálenosti jsou jednotlivé třídy objektů reprezentovány tzv. etalony. Etalon \mathbf{e}_r je vektor obsahující hodnoty příznaků, které jsou v jistém smyslu typické pro třídu ω_r a které tak třídu ω_r reprezentují (na obr.10.1 jsou etalony vyznačeny silně). Je-li k dispozici vhodná trénovací množina, pak stanovení etalonů nečiní potíže. Nejjednodušší možností je použít střední hodnoty

$$\mathbf{e}_r = \frac{1}{N_r} \sum_{i=1}^{N_r} \mathbf{x}_{r,i}, \quad (10.11)$$

kde N_r je počet trénovacích vzorků pro třídu ω_r a $\mathbf{x}_{r,i}$ jsou hodnoty vektoru příznaků získané pro jednotlivé vzorky náležící do této třídy. Etalony je zapotřebí stanovit pro všechny třídy. Rozpoznání pak probíhá takto: Necht' \mathbf{x} reprezentuje rozpoznávaný objekt v prostoru příznaků. Klasifikátor zařadí rozpoznávaný objekt do té třídy, jejíž etalon má od bodu \mathbf{x} nejmenší vzdálenost. Pro zatřídění tedy hledáme minimum

$$\min_r (\text{dist}(\mathbf{e}_r, \mathbf{x})). \quad (10.12)$$

Jako vzdálenost mezi bodem reprezentujícím rozpoznávaný objekt v prostoru příznaků a etalonem \mathbf{e}_r můžeme vzít vzdálenost euklidovskou. Protože vzdálenost je vždy kladná, můžeme místo minima hledat minimum jejího čtverce, což je pohodlnější. Hledáme tedy

$$\min_r (\text{dist}^2(\mathbf{e}_r, \mathbf{x})) = \min_r \left(\sum_{i=1}^m (e_{r,i} - x_i)^2 \right). \quad (10.13)$$

Ukážeme, že klasifikace založená na hledání minima vzdálenosti patří do skupiny metod založených na použití diskriminačních funkcí. K tomu postačí převést úlohu o hledání minima na úlohu o hledání maxima. Je zřejmé, že

$$\min_r(\text{dist}^2(\mathbf{e}_r, \mathbf{x})) = -\max_r(-\text{dist}^2(\mathbf{e}_r, \mathbf{x})). \quad (10.14)$$

Vidíme tedy, že diskriminační funkce $g_r(\mathbf{x})$ má v případě klasifikace pomocí minimální vzdálenosti tvar

$$g_r(\mathbf{x}) = -\text{dist}^2(\mathbf{e}_r, \mathbf{x}). \quad (10.15)$$

Poznamenejme, že ačkoli jsou funkce $g_r(\mathbf{x})$ v tomto případě nelineární, jsou hranice mezi dvěma sousedními oblastmi tvořeny částmi nadrovin. Uvažujme etalony $\mathbf{e}_r, \mathbf{e}_s$, jim odpovídající oblasti χ_r, χ_s a předpokládejme, že se jedná o oblasti sousední. Hranice mezi oblastmi χ_r, χ_s je kolmá ke spojnici $\mathbf{e}_r, \mathbf{e}_s$ a půlí vzdálenost mezi $\mathbf{e}_r, \mathbf{e}_s$.

10.1.3 Určení diskriminačních funkcí minimalizací rizika

Až doposud jsme předpokládali, že každý bod v prostoru příznaků může být obrazem objektů nejvýše jediné třídy. Uvažme nyní situaci, kdy se na jediný bod v prostoru příznaků mohou zobrazit objekty více než jedné třídy. Protože klasifikátor ke každému bodu v prostoru příznaků přiřazuje třídu jedinou, je zřejmé, že se v takovém případě nelze vyhnout chybné klasifikaci. Popsaná situace může nastat i při pečlivé volbě množiny příznaků jednoduše tehdy, jestliže jsou si objekty dvou různých tříd příliš podobné, než aby bylo možné ve všech případech předpokládat jejich zcela bezchybné rozpoznání. V tomto případě, kdy se nutně musíme smířit s tím, že klasifikátor občas odpovídá chybně, lze návrh klasifikátoru založit na minimalizaci ztráty. Předpokládejme, že známe funkci $\lambda(\omega_r|\omega_s)$ popisující ztrátu, kterou utrpíme, jestliže rozpoznávaný objekt, který je ve skutečnosti ze třídy ω_s , klasifikátor chybně zařadí do třídy ω_r (samozřejmým požadavkem je, aby funkční hodnoty $\lambda(\omega_r|\omega_s)$ byly nezáporné). Hodnoty, kterých může funkce $\lambda(\omega_r|\omega_s)$ nabývat, lze uspořádat do matice

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda(\omega_1|\omega_1) & \lambda(\omega_1|\omega_2) & \dots & \lambda(\omega_1|\omega_n) \\ \lambda(\omega_2|\omega_1) & \lambda(\omega_2|\omega_2) & \dots & \lambda(\omega_2|\omega_n) \\ \dots & \dots & \dots & \dots \\ \lambda(\omega_n|\omega_1) & \lambda(\omega_n|\omega_2) & \dots & \lambda(\omega_n|\omega_n) \end{bmatrix}. \quad (10.16)$$

Vhodným měřítkem pro hodnocení kvality klasifikátoru je střední hodnota ztráty, kterou při práci klasifikátoru jeho chybnými rozhodnutími utrpíme. Na minimalizaci této střední ztráty lze založit návrh i učení klasifikátoru. Základní myšlenka je tato: Předpokládejme, že jsme se již rozhodli pro jistý typ klasifikátoru a že pouze zbývá nastavit jeho parametry (v tuto chvíli si např. můžeme zjednodušeně představit, že jsme se rozhodli pro klasifikaci pomocí lineárních diskriminačních funkcí a že zbývá určit hodnoty konstant $a_{r,i}$). Označme \mathbf{a} vektor všech parametrů, které mají být v klasifikátoru nastaveny. Střední ztrátu J vyjádříme jako funkci \mathbf{a} . Je tedy $J=J(\mathbf{a})$. Zatím neznámý vektor \mathbf{a} určíme z podmínky, aby střední ztráta $J(\mathbf{a})$ byla minimální. Ve zbytku tohoto odstavce popíšeme naznačený postup podrobněji. Činnost klasifikátoru zapíšeme jako funkci $d(\mathbf{x}, \mathbf{a})$. Parametr \mathbf{a} , který byl ve srovnání s rovnicí (10.1) přidán, naznačuje, že klasifikace je závislá na hodnotách parametrů, které musí být nastaveny. Předpokládejme nejprve, že všechny rozpoznávané objekty jsou z třídy ω_s a určíme střední ztrátu J_s za tohoto předpokladu. Podle definice střední hodnoty máme

$$J_s(\mathbf{a}) = \int_{\chi} \lambda(d(\mathbf{x}, \mathbf{a})|\omega_s) p(\mathbf{x}|\omega_s) d\mathbf{x}, \quad (10.17)$$

kde $p(\mathbf{x}|\omega_s)$ je podmíněná hustota pravděpodobnosti popisující, jak jsou pro objekty náležící do třídy ω_s zastoupeny jednotlivé hodnoty vektoru příznaků. Ve skutečnosti jsou klasifikátoru k rozpoznání předkládány objekty různých tříd. Celkovou střední ztrátu J proto určíme jako střední hodnotu ztrát J_s z jednotlivých tříd. Nechť $P(\omega_s)$ je apriorní pravděpodobnost jevu, že rozpoznávaný objekt patří do

třídy ω_s . Pro celkovou střední ztrátu pak máme (při úpravě následujícího výrazu jsme zaměnili pořadí sumace a integrace)

$$J(\mathbf{a}) = \sum_{s=1}^n J_s(\mathbf{a})P(\omega_s) = \int \sum_{\chi} \lambda(d(\mathbf{x}, \mathbf{a})|\omega_s) p(\mathbf{x}|\omega_s) P(\omega_s) d\mathbf{x}. \quad (10.18)$$

Nastavení klasifikátoru spočívá v nalezení takových hodnot parametrů, při nichž je celková střední ztráta minimální. Hledáme tedy

$$J_{\min} = \min_{\mathbf{a}} J(\mathbf{a}) = \min_{\chi} \int \sum_{s=1}^n \lambda(d(\mathbf{x}, \mathbf{a})|\omega_s) p(\mathbf{x}|\omega_s) P(\omega_s) d\mathbf{x}. \quad (10.19)$$

Ve výrazu (10.19) nabývají všichni součinitelé za znakem sumace pouze kladných hodnot, a proto také jejich součet, tedy integrand, je vždy kladný. Minimální hodnoty $J(\mathbf{a})$ bude proto dosaženo tehdy, bude-li hodnota integrandu v každém bodě \mathbf{x} nejmenší možná. Ať je realizace klasifikátoru a hodnota vektoru parametrů jakákoli, neexistuje jiná možnost než, že klasifikátor každé hodnotě \mathbf{x} přiřazuje některou třídu z množiny Ω všech tříd. Má-li být střední hodnota celkové ztráty minimální, pak pro každé \mathbf{x} musí být $d(\mathbf{x}, \mathbf{a}) = \omega_{\min}$, kde ω_{\min} je ta třída, pro níž integrand nabývá v \mathbf{x} minimální hodnoty. Je

$$\text{tedy} \quad J_{\min} = \int \min_{\chi} \left[\sum_{s=1}^n \lambda(\omega_r|\omega_s) p(\mathbf{x}|\omega_s) P(\omega_s) \right] d\mathbf{x}. \quad (10.20)$$

$$\text{Zavedme označení} \quad L_r(\mathbf{x}) = \sum_{s=1}^n \lambda(\omega_r|\omega_s) p(\mathbf{x}|\omega_s) P(\omega_s). \quad (10.21)$$

$$\text{Pak máme} \quad J_{\min} = \int \min_{\chi} L_r(\mathbf{x}) d\mathbf{x}. \quad (10.22)$$

Z doposud uvedeného výkladu vyplývá, že klasifikace založená na minimalizaci střední hodnoty celkové ztráty je současně klasifikací pomocí diskriminačních funkcí. Uvažujme nějakou konkrétní hodnotu \mathbf{x} . Výběr minima lze snadno převést na výběr maxima

$$\min_r L_r(\mathbf{x}) = -\max_r (-L_r(\mathbf{x})). \quad (10.23)$$

Nyní již právě popsaný postup hledání hodnoty J_{\min} koresponduje s klasifikací s využitím diskriminačních funkcí. Hledání J_{\min} jsme totiž převedli na problém, kdy integrand ve výrazu (10.22) je v každém bodě \mathbf{x} stanoven tak, že ze všech možných je vybrána ta funkce $-L_r(\mathbf{x})$, která dává maximální hodnotu. Tím je současně také pro každé \mathbf{x} určena třída, kterou má klasifikátor odpovídat. Vidíme, že stanovení integrandu vedoucího k minimu ztráty probíhá stejným postupem jako rozpoznání pomocí diskriminačních funkcí. Klasifikátor založený na principu minima střední hodnoty celkové ztráty lze proto realizovat jako klasifikátor využívající diskriminačních funkcí. Diskriminační funkce mají tvar

$$g_r(\mathbf{x}) = -L_r(\mathbf{x}) = -\sum_{s=1}^n \lambda(\omega_r|\omega_s) p(\mathbf{x}|\omega_s) P(\omega_s). \quad (10.24)$$

Ke stanovení diskriminačních funkcí je zapotřebí znát hodnoty $p(\mathbf{x}|\omega_s)$, $P(\omega_s)$, $\lambda(\omega_r|\omega_s)$. Hodnoty $p(\mathbf{x}|\omega_s)$ stanovíme měřením na vhodné trénovací množině. Také hodnoty $P(\omega_s)$ lze zpravidla dosti dobře stanovit (nebo odhadnout). Volbou ztrátové funkce $\lambda(\omega_r|\omega_s)$ lze vzít v úvahu specifické požadavky na chování klasifikátoru. Nejsou-li kladeny žádné speciální požadavky, pak je možné spokojit se s jednoduchou volbou

$$\lambda(\omega_r|\omega_s) = \begin{cases} 0 & \text{pro } r = s \\ 1 & \text{pro } r \neq s. \end{cases} \quad (10.25)$$

Tuto volbu lze interpretovat takto: Jestliže klasifikátor odpověděl správně ($r=s$), pak nedošlo k žádné ztrátě. Je-li odpověď klasifikátoru chybná ($r \neq s$), pak je ztráta vždy 1 bez ohledu na to, kterou z chybných tříd klasifikátor odpověděl. Matice hodnot ztrátové funkce je tedy

$$\mathbf{\Lambda} = \begin{bmatrix} 0 & 1 & \dots & 1 \\ 1 & 0 & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 0 \end{bmatrix}. \quad (10.26)$$

Dále odvodíme tvar diskriminačních funkcí pro tento speciální případ. Dosazením vztahu (10.25) do vztahu (10.21) obdržíme

$$L_r(\mathbf{x}) = \sum_{\substack{s=1 \\ s \neq r}}^n p(\mathbf{x}|\omega_s)P(\omega_s) = \sum_{s=1}^n p(\mathbf{x}|\omega_s)P(\omega_s) - p(\mathbf{x}|\omega_r)P(\omega_r). \quad (10.27)$$

Rovnici upravíme s využitím Bayesova vztahu pro podmíněné pravděpodobnosti $P(A|B)P(B) = P(B|A)P(A)$ (P značí pravděpodobnost; A, B jsou události). Dostaneme

$$L_r(\mathbf{x}) = p(\mathbf{x}) \sum_{s=1}^n P(\omega_s|\mathbf{x}) - p(\mathbf{x}|\omega_r)P(\omega_r) = p(\mathbf{x}) - p(\mathbf{x}|\omega_r)P(\omega_r). \quad (10.28)$$

Konečně ještě uvedeme předpis pro diskriminační funkce. Využijeme vztahů (10.24) a (10.28). Uvážíme také, že se hodnota $p(\mathbf{x})$ vyskytuje ve všech funkcích $L_r(\mathbf{x})$, a nepodílí se proto na rozhodování (na výběru minima nebo maxima). Diskriminační funkci proto můžeme stanovit jako

$$g_r(\mathbf{x}) = p(\mathbf{x}|\omega_r)P(\omega_r). \quad (10.29)$$

V tomto případě je tedy ke stanovení diskriminačních funkcí nutné znát hodnoty $p(\mathbf{x}|\omega_r)$, $P(\omega_r)$. Jak jsme již uvedli v předchozím obecnějším případě, lze tyto hodnoty zpravidla snadno získat. Vztah (10.29) také ukazuje, že volba diskriminačních funkcí v obr. 10.4 byla racionální.

Modifikujeme-li vztah (10.21) podle Bayesova vzorce $p(\mathbf{x}|\omega_s)P(\omega_s) = P(\omega_s|\mathbf{x})p(\mathbf{x})$, dostaneme

$$L_r(\mathbf{x}) = \sum_{s=1}^n \lambda(\omega_r|\omega_s)P(\omega_s|\mathbf{x})p(\mathbf{x}) = p(\mathbf{x}) \sum_{s=1}^n \lambda(\omega_r|\omega_s)P(\omega_s|\mathbf{x}). \quad (10.30)$$

Hustota pravděpodobnosti $p(\mathbf{x})$ nezávisí na klasifikační třídě. Její odstranění neovlivní rozhodování.

Můžeme tedy položit

$$\tilde{L}_r(\mathbf{x}) = \frac{L_r(\mathbf{x})}{p(\mathbf{x})} = \sum_{s=1}^n \lambda(\omega_r|\omega_s)P(\omega_s|\mathbf{x}). \quad (10.31)$$

Při volbě jednotkové ztrátové funkce dle vztahu (10.25) dostaneme

$$\tilde{L}_r(\mathbf{x}) = \sum_{s=1}^n P(\omega_s|\mathbf{x}) - P(\omega_r|\mathbf{x}) = 1 - P(\omega_r|\mathbf{x}). \quad (10.32)$$

Minimum ztráty obdržíme, jestliže pro dané \mathbf{x} odpovídá klasifikátor tou třídou r , pro níž je $P(\omega_r|\mathbf{x})$ maximální. Z toho vyplývá, že diskriminační funkci můžeme také konstruovat podle předpisu (jedná se o tzv. aposteriorní pravděpodobnost třídy ω_r)

$$g_r(\mathbf{x}) = P(\omega_r|\mathbf{x}). \quad (10.33)$$

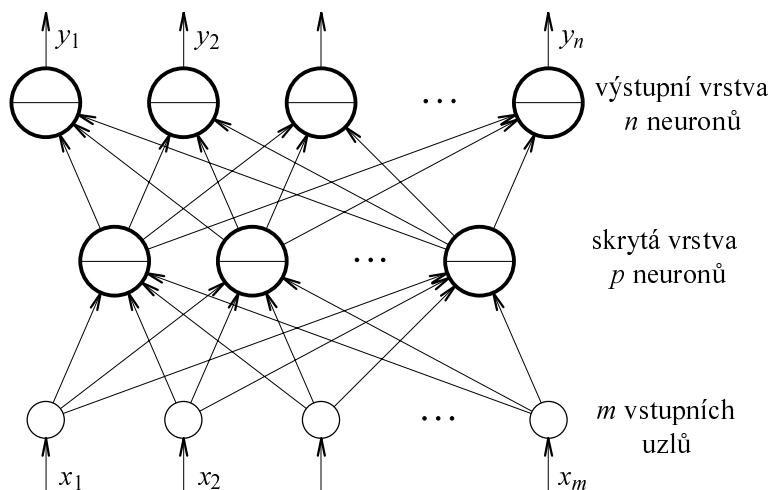
Na závěr tohoto odstavce poznamenejme, že technika stanovení diskriminačních funkcí minimalizací rizika, která byla v tomto odstavci popsána, se často nazývá technikou bayesovských odhadů. Odpovídající klasifikátor pak bývá často nazýván bayesovským strojem. My jsme se zde přidrželi názvu, který je popisnější.

10.2 Neuronové sítě

Neuronové sítě opět provádějí klasifikaci formálně popsanou vztahem (10.1), avšak jsou realizovány poněkud speciálně. Neuronovým sítím je v rámci studia informatiky na FEI VŠB-TU věnován rozsáhlý prostor, a proto se zde omezíme jen na základní případy, které jsou v praxi příznakového rozpoznávání používány nejčastěji.

10.2.1 Třívrstvá (vícevrstvá) síť s učením „back propagation“

Schéma třívrstvé neuronové sítě je znázorněno na obr. 10.6. Do vrstvy vstupních uzlů se zavádí jednotlivé složky vektoru příznaků, a proto je vstupních uzlů tolik, kolik je příznaků. Z výstupní vrstvy se odebrává identifikátor třídy. Třebaže jsou možné i jiné způsoby, často se používá kódování 1 z n . V tomto případě je pak počet neuronů ve výstupní vrstvě roven počtu rozpoznávaných tříd. Počet neuronů ve střední skryté vrstvě (případně vrstev) se zpravidla volí na základě zkušenosti. Neurony po



Obr. 10.6. Schéma třívrstvé neuronové sítě.

sobě jdoucích vrstev jsou spolu pospojovány tak, jak je naznačeno na obr. 10.6.

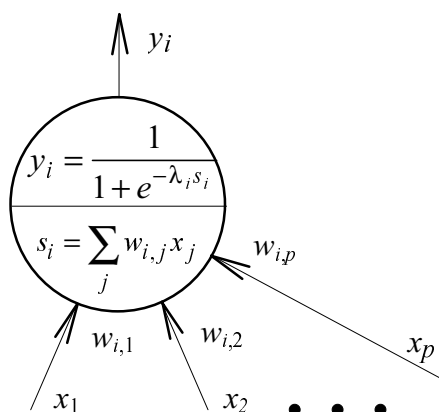
Dopředné šíření vzruchu i -tým neuronem popisují rovnice (obr. 10.7)

$$y_i = \frac{1}{1 + e^{-\lambda_i s_i}}, \quad (10.34)$$

$$\text{kde } s_i = \sum_j w_{i,j} x_j \quad (10.35)$$

Váhy $w_{i,j}$ v předchozí rovnici se stanovují učením, aby síť odpovídala tak, jak je požadováno (koeficienty λ_i lze zpravidla stanovit pevně předem). Před zahájením učení se vahám přiřadí např. ná-

hodné hodnoty z intervalu $\langle -0.5, 0.5 \rangle$. Předpokládáme, že pro učení máme k dispozici trénovací množinu obsahující dvojice $\{(\mathbf{x}_1, \mathbf{t}_1), (\mathbf{x}_2, \mathbf{t}_2), \dots, (\mathbf{x}_q, \mathbf{t}_q)\}$. Vektor \mathbf{x}_i je vstup, \mathbf{t}_i je výstup, kterým by neuronová síť měla na vstup \mathbf{x}_i odpovídat. Je užitečné přikládat na vstup neuronové sítě hodnoty (složky vektoru \mathbf{x} , tj. příznaky) normalizované, např. do intervalu $\langle 0, 1 \rangle$. To proto, aby všechny složky (příznaky) měly na učení a rozpoznání přibližně stejný vliv. Také hodnoty na výstupu nemohou být s ohledem na sigmoidální funkci (10.34) předepsány libovolně. Teoreticky se na výstupu mohou objevit hodnoty z intervalu $(0, 1)$. Prakticky však předepisujeme hodnoty z intervalu užšího, např. z intervalu $\langle 0.1, 0.9 \rangle$. Učení sítě probíhá tak, že na vstup přikládáme postupně všechny vektory \mathbf{x}_i z trénovací množiny. Po přiložení každého vektoru vyhodnotíme rozdíl mezi skutečnou a požadovanou odpovědí neuronové sítě a změnou vah $w_{i,j}$ pro všechny neurony se snažíme dosáhnout zlepšení souhlasu. Nechť $(\mathbf{x}=(x_1, x_2, \dots, x_m), \mathbf{t}=(t_1, t_2, \dots, t_n))$ je jedna dvojice z trénovací množiny. Chyba neuronové sítě pro tuto



Obr.10.7. Realizace neuronu.

dvojici činí (y_i je skutečná, t_i požadovaná odpověď na i -tém výstupním neuronu; sčítáme přes všechny výstupní neurony, kterých je n)

$$E = \frac{1}{2} \sum_{i=1}^n (y_i - t_i)^2. \quad (10.36)$$

Logickým principem, na kterém je založeno učení neuronové sítě, je nalezení takových vah $w_{i,j}$, při kterých je uvedená chyba minimální. Metoda učení „back propagation“ je aplikací hledání minima gradientní metodou. Připomeňme, že gradientní metoda hledá minimum funkce $f(\mathbf{x})$ následujícím iteračním postupem

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \eta \text{grad}(f(\mathbf{x}))|_{\mathbf{x}=\mathbf{x}^{(k)}}, \quad (10.37)$$

kde horní indexy (k) , $(k+1)$ znamenají pořadové číslo iterace. Gradientní metoda tedy postupuje proti směru gradientu „dolů“ ve směru největšího spádu. Délku

roku ovlivňuje vhodně zvolený násobitel η . Pro váhy v neuronové síti na základě gradientní metody máme (\mathbf{w} je vektor všech vah, vektory i gradient jsme v následujícím vztahu rozepsali po složkách)

$$w_{i,j}^{(k+1)} = w_{i,j}^{(k)} - \eta \left. \frac{\partial E(\mathbf{w})}{\partial w_{i,j}} \right|_{\mathbf{w}=\mathbf{w}^{(k)}}. \quad (10.38)$$

Je tedy

$$\Delta w_{i,j} = -\eta \left. \frac{\partial E}{\partial w_{i,j}} \right|_{\mathbf{w}=\mathbf{w}^{(k)}} = -\eta \left. \frac{\partial E}{\partial y_i} \frac{\partial y_i}{\partial s_i} \frac{\partial s_i}{\partial w_{i,j}} \right|_{\mathbf{w}=\mathbf{w}^{(k)}}. \quad (10.39)$$

Dále vyjádříme jednotlivé derivace z předchozí rovnice. Nejprve se omezíme na výstupní vrstvu neuronů. Máme

$$\frac{\partial E}{\partial y_i} = (y_i - t_i), \quad \frac{\partial y_i}{\partial s_i} = \frac{\lambda_i e^{-\lambda_i s_i}}{(1 + e^{-\lambda_i s_i})^2} = \lambda_i y_i (1 - y_i), \quad \frac{\partial s_i}{\partial w_{i,j}} = x_j. \quad (10.40a,b,c)$$

Pro výstupní vrstvu neuronů tedy máme

$$\Delta w_{i,j} = -\eta \delta_i \lambda_i y_i (1 - y_i) x_j, \quad \text{kde } \delta_i = (y_i - t_i). \quad (10.41a,b)$$

Pro vnitřní skrytou vrstvu neuronů (při odvození vztahů pro učení se zde omezíme na síť třívrstvé) je jiná hodnota $\partial E/\partial y_i$. Uvažujme neuron ve skryté vrstvě. Hodnota y_i na výstupu tohoto neuronu ovlivňuje hodnoty na výstupu všech výstupních neuronů. Je proto (hodnoty se stříškou a suma v následujícím výrazu se týkají výstupní vrstvy neuronů)

$$\frac{\partial E}{\partial y_i} = \sum_{k=1}^n \frac{\partial E}{\partial \hat{y}_k} \frac{\partial \hat{y}_k}{\partial s_k} \frac{\partial s_k}{\partial \hat{x}_{k,i}} = \sum_{k=1}^n (\hat{y}_k - t_k) \hat{\lambda}_k \hat{y}_k (1 - \hat{y}_k) \hat{w}_{k,i}. \quad (10.42)$$

Protože pro hodnoty zbývajících derivací $\partial y_i/\partial s_i$, $\partial s_i/\partial w_{i,j}$ platí vztahy (10.40b,c) i pro skrytou vrstvu, můžeme pro modifikaci vah $w_{i,j}$ ve skryté vrstvě opět použít vztahu (10.41a), v němž však položíme

$$\delta_i = \frac{\partial E}{\partial y_i} = \sum_{k=1}^n (\hat{y}_k - t_k) \hat{\lambda}_k \hat{y}_k (1 - \hat{y}_k) \hat{w}_{k,i} = \sum_{k=1}^n \hat{\delta}_k \hat{\lambda}_k \hat{y}_k (1 - \hat{y}_k) \hat{w}_{k,i}, \quad (10.43)$$

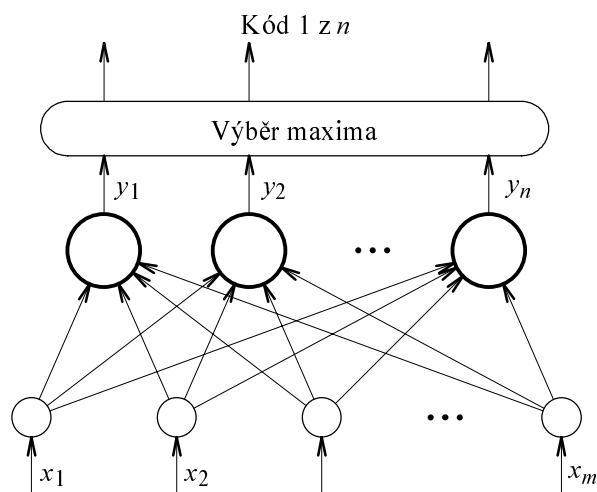
kde $\hat{\delta}_k = (\hat{y}_k - t_k)$ je diference vztahující se k výstupní vrstvě neuronů. Učení probíhá tak, že dvojice $(\mathbf{x}_i, \mathbf{t}_i)$ z trénovací množiny probíráme jednu po druhé a pro každou dvojici provedeme pro všechny váhy jeden krok iterace (10.38). Tento proces se cyklicky opakuje. Ukončení procesu kontrolujeme pomocí zjišťování velikosti chyby

$$Err = \sqrt{\sum_{i=1}^q \sum_{k=1}^n \hat{\delta}_{k,i}^2} / nq, \quad (10.44)$$

kde $\hat{\delta}_{k,i}$ je chyba na výstupu k -tého neuronu při přiložení i -té dvojice z trénovací množiny, n je počet neuronů ve výstupní vrstvě a q je počet prvků (dvojic $(\mathbf{x}_i, \mathbf{t}_i)$) trénovací množiny. Učení využívající trénovací množiny, v níž je ke každému vektoru příznaků zadán také požadovaný výstup neuronové sítě (kód třídy objektu), se nazývá učením za asistence učitele. Právě popsané učení „back propagation“ je příkladem takového postupu. Není-li přiřazení tříd jednotlivým vektorům příznaků dopředu známo, nelze učení za asistence učitele použít.

10.2.2 Neuronová síť s kompeticí

Neuronová síť s kompeticí nevyžaduje ke svému učení asistenci učitele. V trénovací množině nemusí být k jednotlivým vektorům příznaků přiřazeny třídy. Dokonce ani nemusí žádná předem připravená trénovací množina existovat a síť se může učit přímo za provozu. Každý zpracovávaný vektor příznaků reprezentuje nějaký bod v prostoru příznaků. Neuronová síť s kompeticí během svého učení detekuje shluky takových bodů. Rozpoznání je založeno na porovnání předloženého vektoru příznaků s vektory příznaků reprezentujícími jednotlivé shluky. Schéma sítě je uvedeno na obr. 10.8. Síť obsahuje jedinou vrstvu neuronů. Do každého neuronu jsou přivedeny hodnoty jednotlivých složek vektoru



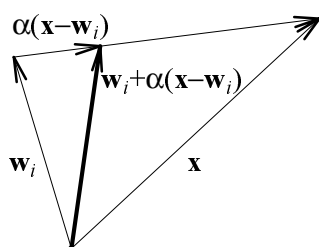
Obr. 10.8. Neuronová síť s kompeticí.

ného objektu - každá třída má svůj neuron, který ji rozpoznává. Neuronová síť tak kóduje rozpoznanou třídu metodou 1 z n. Počet neuronů je proto nejméně roven počtu tříd, které má síť rozpoznávat. Pokud by bylo zapotřebí výsledek rozpoznání reprezentovat pomocí jiného kódu, bylo by možné na výstupy sítě z obr. 10.8 přidat vhodný převodník kódu (k jeho realizaci by bylo možné použít např. další vrstvy neuronů).

Po každém rozpoznání proběhne jeden krok učení, během něhož jsou váhy u vítězného neuronu modifikovány podle předpisu

$$\mathbf{w}_i \leftarrow \mathbf{w}_i + \Delta \mathbf{w}_i, \text{ kde } \Delta \mathbf{w}_i = \alpha(\mathbf{x} - \mathbf{w}_i) \quad (10.46)$$

a kde α ($\alpha \ll 1$) je empiricky stanovený koeficient učení. Je zřejmé, že vztah (10.46) modifikuje váhy vítězného neuronu tak, aby více odpovídaly vektoru příznaků právě rozpoznávaného objektu (obr. 10.9). Na začátku provozu neuronové sítě jsou váhy nastaveny buď na předem zvolené hodnoty nebo náhodně. I když se síť může učit i sama během provozu tím, jak zpracovává další a další vektory příznaků, lze jí prakticky samozřejmě používat teprve tehdy, až bylo dosaženo přiměřeného počátečního naučení. K tomu lze opět použít vhodné reprezentativní množiny příznakových vektorů - trénovací množiny. Při učení lze i v případě neuronové sítě s kompeticí probírat trénovací množinu opakovaně (na vstup sítě se postupně přiloží všechny vektory příznaků trénovací množiny a pak se postup opakuje). Opakovaného probírání trénovací množiny lze využít i k jednoduché kontrole naučení sítě. Síť lze považovat za dostatečně naučenou (s ohledem na zvolenou trénovací množinu), jestliže se při opakovaném probírání trénovací množiny nemění klasifikace jednotlivých vektorů do tříd (klasifikace všech vektorů zůstává stejná, jako byla v předchozím průchodu).



Obr. 10.9. Modifikace vah neuronu při učení.

Alternativně může být signál na výstupu neuronu definován vztahem

$$y_i = \sum_j w_{i,j} x_j = \mathbf{w}_i \mathbf{x}^T = |\mathbf{w}_i| |\mathbf{x}| \cos \varphi. \quad (10.47)$$

Předpokládejme v tomto případě dále, že vektory \mathbf{w}_i , \mathbf{x} jsou normalizovány na jednotkovou délku. Podle vztahu (10.47) je pak hodnota y_i největší tehdy, jestliže vektory \mathbf{w}_i a \mathbf{x} mají stejný směr. Uvedeného postupu lze ovšem použít jen tehdy, jestliže při rozpoznání závisí pouze na směru vektoru příznaků, nikoli na jeho délce.

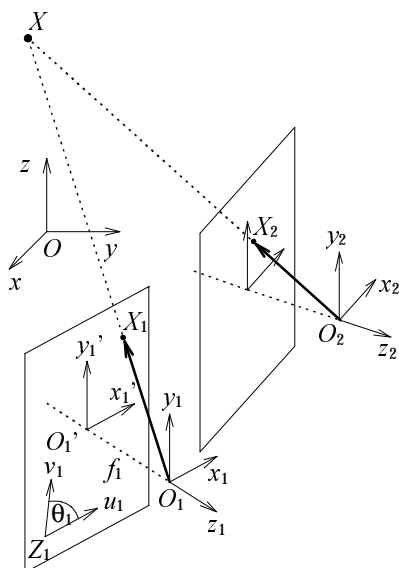
Na závěr této podkapitoly poznamenejme, že problematika neuronových sítí byla v nedávné minulosti studována velmi intenzivně. Toto úsilí vyústilo v řadu zajímavých závěrů, které zde však není možné pro omezený rozsah tohoto textu prezentovat. Pro podrobnější informace proto čtenáře odkazujeme na specializovanou literaturu.

11 Zpětná stereoprojekce

Máme-li k dispozici několik obrazů téže scény, které byly získány několika kamerami umístěnými ve scéně nebo jedinou kamerou umístěnou postupně do různých míst, pak lze za jistých okolností provést rekonstrukci scény. Nejmenší počet kamer (případně míst), které jsou k provedení rekonstrukce zapotřebí, jsou dvě (obr. 11.1). Rekonstrukci při tom rozumíme, že pro zvolené body dokážeme na základě souřadnic jejich obrazů zachycených kamerami vypočítat trojrozměrné souřadnice ve scéně. Řešení naznačeného problému podrobněji popíšeme v následujících podkapitolách.

11.1 Model kamery

Předpokládáme, že vztah mezi objekty pozorovanými kamerou a jejich obrazy je lineární projektivní transformací (kolineací). Modelem kamery je tedy dírková komora. Ve scéně zavedeme globální souřadný systém (O, x, y, z) (obr. 11.1). Uvažujme i -tou kameru (případně i -tou polohu kamery). Pro označení veličin vztahujících se k i -té kamere použijeme symbolů s indexem i . Střed projekce i -té kamery je O_i , π_i je zobrazovací rovina. Kolmice vedená z O_i na rovinu π_i je optická osa kamery. V zobrazovací rovině kamery leží zařízení zachycující obraz (jedná se například o pole snímacích elementů CCD čipu, případně o světlocitlivou vrstvu). Souřadnice bodů v obrazech měříme v souřadném systému (Z_i, u_i, v_i) . V případě CCD kamery jsou často souřadné osy tohoto systému hranami pole snímacích elementů. Ortogonální souřadný systém (O_i, x_i, y_i, z_i) je souřadným systémem kamery. Jeho osa z_i je identická s optickou osou kamery, osa x_i je rovnoběžná s osou u_i (v případě CCD kamery tedy s hranou pole snímacích elementů). Dále v obraze zavedeme souřadný systém (O'_i, x'_i, y'_i) . Bod O'_i leží v místě, kde optická osa kamery protíná zobrazovací rovinu. Osa x'_i je rovnoběžná s osou x_i , osa y'_i je rovnoběžná s osou y_i . Vzdálenost $f_i = \text{dist}(O_i, O'_i)$ je ohniskovou vzdáleností kamery.



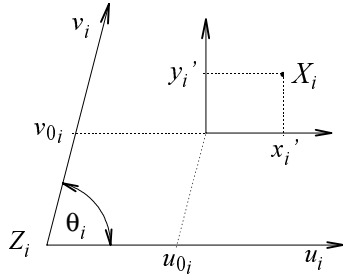
Obr. 11.1. Zpětná stereoprojekce.

Uvažujme nějaký bod X scény. V souřadném systému (O, x, y, z) je bod X reprezentován vektorem $\mathbf{x}=(x, y, z)^T$. V souřadném systému (O_i, x_i, y_i, z_i) je tentýž bod reprezentován vektorem $\mathbf{x}_i=(x_i, y_i, z_i)^T$. Pro transformaci ze souřadného systému (O_i, x_i, y_i, z_i) kamery do globálního souřadného systému (O, x, y, z) scény můžeme psát

$$\mathbf{x} = \mathbf{R}_i \mathbf{x}_i + \mathbf{o}_i, \quad (11.1)$$

kde matice \mathbf{R}_i popisuje rotaci kamery a vektor \mathbf{o}_i (zadaný v souřadném systému (O, x, y, z)) reprezentuje střed projekce O_i . Necht' X_i je bod v zobrazovací rovině i -té kamery a tento bod necht' je obrazem bodu X scény. Souřadnice bodu X_i měřené v souřadném systému (Z_i, u_i, v_i) jsou (u_i, v_i) . Tyto souřadnice uspořádáme do trojrozměrného vektoru $\mathbf{u}_i=(u_i, v_i, 1)^T$. V souřadném systému (O'_i, x'_i, y'_i) je poloha bodu X_i popsána vektorem $\mathbf{x}'_i=(x'_i, y'_i, 1)^T = \mathbf{Q}_i \mathbf{u}_i$, kde \mathbf{Q}_i je matice popisující transformaci ze souřadného systému (Z_i, u_i, v_i) do souřadného systému (O'_i, x'_i, y'_i) . Snadno ověříme, že platí

$$\mathbf{Q}_i = \left(\begin{bmatrix} 1 & 0 & u_{0_i} \\ 0 & 1 & v_{0_i} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & s_i & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -\cotg \theta_i & 0 \\ 0 & 1/\sin \theta_i & 0 \\ 0 & 0 & 1 \end{bmatrix} \right)^{-1} = \begin{bmatrix} 1 & -\cotg \theta_i & u_{0_i} \\ 0 & s/\sin \theta_i & v_{0_i} \\ 0 & 0 & 1 \end{bmatrix}^{-1}. \quad (11.2)$$



Obr. 11.2. K významu prvků matice \mathbf{Q}_i .

Jednotlivé symboly v matici \mathbf{Q}_i mají následující význam (obr.11.2). Hodnoty u_{0i} , v_{0i} jsou souřadnice (měřené v souřadném systému (Z_i, u_i, v_i)) bodu, v němž optická osa protíná zobrazovací rovinu kamery. Úhel θ_i modeluje možnou odchylku pole snímacích elementů CCD senzoru od ortogonality, případně možnou úchylku kolmosti zobrazovací roviny a optické osy. Souřadný systém (Z_i, u_i, v_i) tedy nemusí být ortogonální, ale v praxi je hodnota θ_i obvykle velmi blízká $\pi/2$. Význam měřítkového faktoru s_i je následující: U kamer s CCD čipem zohledňuje možné diference rozměrů snímacích elementů ve směru osy u a v . Jestliže je výstup kamery analogový, pak parametr s_i zohledňuje také možné zkreslení rozměrů snímacích elementů způsobené násled-

nou zpětnou digitalizací analogového signálu. Ze vztahu (11.2) je zřejmé, že parametrem s_i korigujeme pouze poměrné zkreslení délek na osách - skutečná celková velikost obrazu bude vzata v úvahu hodnotou ohniskové vzdálenosti. Poznamenejme, že matice \mathbf{Q}_i^{-1} popisuje transformaci v opačném směru než matice \mathbf{Q}_i , tedy ze souřadného systému (O_i', x_i', y_i') do souřadného systému (Z_i, u_i, v_i) . Zdá se názornější sestavit právě matici \mathbf{Q}_i^{-1} . Matici \mathbf{Q}_i získáme inverzí. Tento postup jsme uplatnili ve vztahu (11.2). Zavedeme dále matici

$$\mathbf{F}_i = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -f_i \end{bmatrix}. \quad (11.3)$$

Snadno ověříme, že platí

$$\mathbf{x}_i = \lambda_i \mathbf{F}_i \mathbf{Q}_i \mathbf{u}_i, \quad (11.4)$$

kde λ_i je reálný parametr. S využitím vztahu (11.1) obdržíme

$$\mathbf{x} = \lambda_i \mathbf{R}_i \mathbf{F}_i \mathbf{Q}_i \mathbf{u}_i + \mathbf{o}_i, \quad (11.5)$$

případně

$$\lambda_i \mathbf{u}_i = (\mathbf{R}_i \mathbf{F}_i \mathbf{Q}_i)^{-1} (\mathbf{x} - \mathbf{o}_i). \quad (11.6)$$

Parametry obsažené v maticích \mathbf{F}_i , \mathbf{Q}_i (f_i , u_{0i} , v_{0i} , θ_i , s_i) jsou vnitřní (tzv. intrinsické) parametry kamery. Matice \mathbf{Q}_i zohledňuje některé rozdíly mezi skutečnou a ideální kamerou. V jednoduchých teoretických úvahách se někdy předpokládá $\mathbf{Q}_i = \mathbf{I}$ (jednotková matice). Hodnoty \mathbf{R}_i , \mathbf{o}_i popisující polohu kamery v prostoru se nazývají vnějšími (tzv. extrinsickými) parametry kamery. Základní úlohou počítačového stereovidění je na základě souřadnic (u_i, v_i) obrazů nějakého bodu scény zjistit jeho prostorové souřadnice (x, y, z) . K řešení této úlohy je zpravidla nejprve nutné provést tzv. kalibraci kamery. Úkolem kalibrace je stanovit vnitřní a vnější parametry kamery. Jen v případech, kdy jsou parametry kamery předem známy, může kalibrace chybět (od výrobce např. můžeme znát hodnoty vnitřních parametrů, umístění kamery v prostoru můžeme někdy dostatečně přesně změřit).

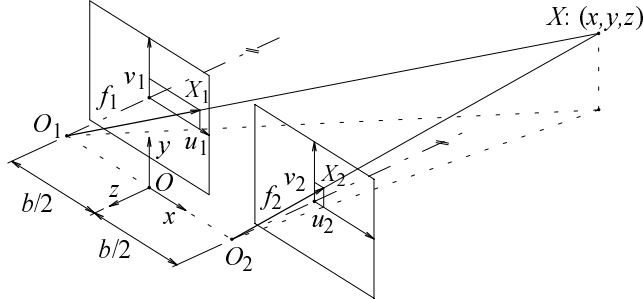
Vztahy popisující promítání realizované kamerou se zjednoduší, jestliže použijeme homogenních souřadnic. (Předpokládáme zde, že je čtenář obeznámen alespoň s elementárními pojmy z projektivní geometrie. V případě potřeby lze potřebné znalosti získat např. z prací (Budínský 83, Penna 86)). Ostatně již dříve jsme vektor \mathbf{u}_i zavedli jako $\mathbf{u}_i = (u_i, v_i, 1)^T$, což lze interpretovat jako homogenní souřadnice bodu X_i v dvojrozměrném projektivním prostoru. Dále navíc připustíme, aby homogenní souřadnice nabývala obecné hodnoty w_i , tedy $\mathbf{u}_i = (w_i u_i, w_i v_i, w_i)^T$. Podobně popíšeme v homogenních souřadnicích i polohu bodu X vzhledem ke globálnímu souřadnému systému scény. Položíme $\mathbf{x} = (x, y, z, 1)^T$. Projektivní transformaci realizovanou kamerou lze nyní zapsat ve tvaru

$$\mathbf{u}_i = \mathbf{P}_i \mathbf{x}. \quad (11.7)$$

Matice \mathbf{P}_i rozměru 3×4 popisuje transformaci. Vztahy (11.5), (11.6), (11.7) popisují různými prostředky tutéž transformaci a jsou vzájemně ekvivalentní. V dalším textu budeme podle potřeby používat vždy ten tvar, který je pro danou úlohu nejvýhodnější. Porovnáním vztahu (11.7) se vztahem (11.6) zjišťujeme, že matice \mathbf{P}_i zahrnuje vliv vnějších i vnitřních parametrů kamery.

11.2 Příklad dvou kamer s rovnoběžnými optickými osami

Nejprve ukážeme jednoduchý případ. Předpokládejme, že máme dvě kamery, které jsou uspořádány takto (obr. 11.3): 1) Obě kamery mají rovnoběžné optické osy. 2) Předpokládáme, že platí $\mathbf{Q}_1 = \mathbf{Q}_2 = \mathbf{I}$,



Obr. 11.3. Kamery s rovnoběžnými optickými osami.

tedy $x_1' = u_1$, $y_1' = v_1$, $x_2' = u_2$, $y_2' = v_2$. 3) Souřadné osy u_1 , u_2 leží v jedné přímce. 4) Souřadné osy v_1, v_2 jsou rovnoběžné. 5) Vzdálenost optických středů kamer je b . 6) Počátek globálního souřadného systému scény pólí spojnicí O_1, O_2 a osy globálního systému jsou rovnoběžné s osami $u_1(u_2)$, $v_1(v_2)$. 7) $f_1 = f_2 = f$. Uvažujme ve scéně bod X o souřadnicích (x, y, z) . Bod X se promítá do bodů X_1, X_2 ležících v zobrazovacích rovinách první a druhé kamery. Souřadnice bodů X_1, X_2 měřené v obrazech získaných kamerami jsou (u_1, v_1) , (u_2, v_2) . Z podobnosti trojúhelníků snadno odvodíme následující vztahy (předpokládáme, že ohnisková vzdálenost f se zadává jako kladná hodnota):

$$\frac{x + b/2}{-z} = \frac{u_1}{f}, \quad \frac{x - b/2}{-z} = \frac{u_2}{f}. \quad (11.8)$$

Analogické vztahy snadno zapíšeme také pro souřadnici y . Máme

$$\frac{y}{-z} = \frac{v_1}{f}, \quad \frac{y}{-z} = \frac{v_2}{f}. \quad (11.9)$$

V rovnicích (11.8) jsou x, z neznámými. Ostatní hodnoty jsou známe. Řešením soustavy pro z získáme

$$z = -\frac{bf}{u_1 - u_2}. \quad (11.10)$$

S využitím rovnic (11.8), (11.9) pak snadno také nalezneme hodnoty x, y . Vychází

$$x = \frac{b(u_1 + u_2)}{2(u_1 - u_2)}, \quad y = \frac{b(v_1 + v_2)}{2(u_1 - u_2)}. \quad (11.11)$$

11.3 Absolutní kalibrace a rekonstrukce

V tomto odstavci ukážeme postup, v němž se kalibrace provádí na základě toho, že pro jistý počet tzv. kalibračních bodů známe jak souřadnice u_i, v_i jejich obrazů tak i jejich souřadnice x, y, z v souřadném systému scény (název absolutní kalibrace je odvozen od požadavku znát souřadnice x, y, z kalibračních bodů). Zaměříme se na případ, kdy máme pouze dvě kamery ($i=1,2$). K odvození výsledných vztahů použijeme homogenních souřadnic a vyjdeme ze vztahu (11.7). Matici \mathbf{P}_i rozepíšeme do tří řádkových vektorů rozměru 1×4 $\mathbf{P}_i = (\mathbf{p}_{i,1}, \mathbf{p}_{i,2}, \mathbf{p}_{i,3})^T$. Rovnici (11.7) lze nyní přepsat takto

$$\begin{bmatrix} w_i u_i \\ w_i v_i \\ w_i \end{bmatrix} = \begin{bmatrix} \mathbf{p}_{i,1} \\ \mathbf{p}_{i,2} \\ \mathbf{p}_{i,3} \end{bmatrix} \mathbf{x}. \quad (11.12)$$

Předpokládejme nyní, že známe projekční matice $\mathbf{P}_1, \mathbf{P}_2$ obou kamer. Dále předpokládejme, že je ve scéně bod, jehož obraz získaný první kamerou má souřadnice u_1, v_1 a obraz získaný druhou kamerou má souřadnice u_2, v_2 . Ukážeme, jak lze rekonstruovat trojrozměrné souřadnice x, y, z tohoto bodu. Položme $\mathbf{x} = (x, y, z, 1)$. Z posledního řádku rovnice (11.12) máme

$$w_1 = \mathbf{p}_{1,3}\mathbf{x}, \quad w_2 = \mathbf{p}_{2,3}\mathbf{x}. \quad (11.13)$$

Dosazením rovnic (11.13) do prvních dvou řádků vztahu (11.12) rozepsaného pro první i druhou kameru dostaneme

$$\begin{bmatrix} \mathbf{p}_{1,1}\mathbf{x} - u_1\mathbf{p}_{1,3}\mathbf{x} \\ \mathbf{p}_{1,2}\mathbf{x} - v_1\mathbf{p}_{1,3}\mathbf{x} \\ \mathbf{p}_{2,1}\mathbf{x} - u_2\mathbf{p}_{2,3}\mathbf{x} \\ \mathbf{p}_{2,2}\mathbf{x} - v_2\mathbf{p}_{2,3}\mathbf{x} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \quad (11.14)$$

Vztah (11.14) je lineární nehomogenní (uvažme, že $\mathbf{x}=(x,y,z,1)$) soustavou rovnic pro neznámé hodnoty souřadnic x,y,z . Protože pro tři neznámé jsou k dispozici čtyři rovnice, jde o soustavu předeterminovanou. Předeterminovanosti lze využít ke zmenšení vlivu chyb, které mohou být do výpočtu vneseny nepřesným zadáním. Hodnoty u_1, v_1, u_2, v_2 souřadnic měřených v obrazech bývají totiž zatíženy šumem, který vzniká nepřesným odečítáním souřadnic v obraze. Do této kategorie nepřesností patří i případ, kdy jsou v digitalizovaném obraze souřadnice měřeny pouze celočíselnými (nikoli reálnými) hodnotami (někdy udáváme souřadnice pixelu, do kterého obraz rekonstruovaného bodu padne, jako celočíselné). Řešení lineárních předeterminovaných systémů metodou nejmenších čtverců je popsáno v dodatku B.

Z předchozího výkladu vyplývá, že jsou-li k dispozici projekční matice \mathbf{P}_i , pak lze provést rekonstrukci souřadnic bodů. Postup, v němž jsou projekční matice stanoveny, se obvykle nazývá kalibrací kamery. Ukážeme metodu, jak lze kalibraci provést. Dosadíme opět (11.13) do prvních dvou řádků vztahu (11.12) a provedeme transpozici obou takto získaných rovnic. Obdržíme

$$\begin{aligned} \mathbf{x}^T \mathbf{p}_{i,1}^T - u_i \mathbf{x}^T \mathbf{p}_{i,3}^T &= 0, \\ \mathbf{x}^T \mathbf{p}_{i,2}^T - v_i \mathbf{x}^T \mathbf{p}_{i,3}^T &= 0. \end{aligned} \quad (11.15)$$

Uspořádejme zatím neznámé prvky projekční matice \mathbf{P}_i do sloupcového vektoru $\mathbf{p}_i = (\mathbf{p}_{i,1}^T, \mathbf{p}_{i,2}^T, \mathbf{p}_{i,3}^T)^T$. Vztahy (11.15) lze pak přepsat maticově takto

$$\begin{bmatrix} \mathbf{x}^T & \mathbf{0}^T & -u_i \mathbf{x}^T \\ \mathbf{0}^T & \mathbf{x}^T & -v_i \mathbf{x}^T \end{bmatrix} \begin{bmatrix} \mathbf{p}_{i,1}^T \\ \mathbf{p}_{i,2}^T \\ \mathbf{p}_{i,3}^T \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \quad (11.16)$$

$$\text{Rozepsáno tedy } \begin{bmatrix} x & y & z & 1 & 0 & 0 & 0 & 0 & -u_i x & -u_i y & -u_i z & -u_i \\ 0 & 0 & 0 & 0 & x & y & z & 1 & -v_i x & -v_i y & -v_i z & -v_i \end{bmatrix} \mathbf{p}_i = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \quad (11.17)$$

Vidíme, že každý kalibrační bod přispívá k nalezení prvků projekční matice dvěma rovnicemi. Stanovení dvanácti neznámých prvků projekční matice provedeme na základě rovnic poskytnutých dostatečným počtem kalibračních bodů. Ukážeme ale, že projekční matici bude možné pomocí kalibračních bodů stanovit až na reálného násobitele. Uvedenou nejednoznačnost lze snadno ilustrovat pomocí rovnice (11.7). Násobme levou i pravou stranu rovnice libovolným nenulovým reálným násobitelem μ . Protože v homogenních souřadnicích vektor $\mu \mathbf{u}_i$ reprezentuje tentýž bod obrazu jako vektor \mathbf{u}_i , musí být také matice $\mu \mathbf{P}_i$ nutně správnou projekční maticí. Nejednoznačnost vyřešíme zavedením doplňující podmínky. Snadno lze např. realizovat postup, kdy některý z prvků projekční matice položíme roven zvolené hodnotě (např. roven jedné). Jistou potíž však v tomto případě působí rozhodnutí, který ze členů projekční matice bychom měli takto nastavit (zejména se nesmíme pokoušet takto nastavit člen, který by měl vyjít nulový). Čistší možností je požadovat, aby vhodná norma projekční matice byla rovna zvolené hodnotě. Můžeme např. požadovat, aby součet čtverců všech prvků projekční matice byl roven jedné. V tomto případě však bohužel obdržíme nelineární úlohu, což je podstatná komplikace. Podkapitolu můžeme uzavřít konstatováním, že máme-li k dispozici pro každou kameru alespoň 6 kalibračních bodů a zvolíme-li doplňující podmínku, můžeme pro řešení prvků projekční matice sestavit systém rovnic a prvky jednoznačně vyřešit. I při použití minimálního počtu šesti bodů je

ovšem tento systém předeterminovaný. Pro eliminaci šumu (šumem opět mohou být zatíženy zejména hodnoty souřadnic u_i, v_i změřené v obrazech) bývá však v praxi počet kalibračních bodů mnohem vyšší než teoreticky nutné minimum. V otázce řešení předeterminovaných lineárních soustav opět odkazujeme na dodatek B.

11.4 Relativní kalibrace a rekonstrukce

Často přichází v úvahu úloha, kdy můžeme měřit pouze souřadnice bodů v obrazech získaných kamerami. Trojrozměrné souřadnice žádného bodu ve scéně při tom nejsou k dispozici. Taková situace může nastat např. v případě vizuálního systému robota, který pracuje v neznámé scéně. Ukážeme, že i za této situace lze scénu rekonstruovat. Uvažujme opět situaci, kdy máme k dispozici dvě kamery, případně dva obrazy sejmuté ze dvou různých míst jedinou kamerou. Zavedeme matici rotace \mathbf{R} a vektor posunutí $\mathbf{b} = O_2 - O_1$, které provádějí transformaci ze souřadného systému druhé kamery do souřadného systému první kamery podle vztahu

$$\mathbf{x}_1 = \mathbf{R}\mathbf{x}_2 + \mathbf{b}. \quad (11.18)$$

Připomeňme, že \mathbf{R} je ortonormální matice a že $\mathbf{R} = \mathbf{R}_z\mathbf{R}_y\mathbf{R}_x$, kde $\mathbf{R}_x, \mathbf{R}_y, \mathbf{R}_z$ jsou matice odpovídající postupně prováděným rotacím kolem os x_2, y_2 a z_2 . Zavedeme vektor $\boldsymbol{\varphi} = (\varphi_x, \varphi_y, \varphi_z)^T$ obsahující odpovídající úhly rotace. Matice $\mathbf{R}_x, \mathbf{R}_y, \mathbf{R}_z$ pak mají tvar

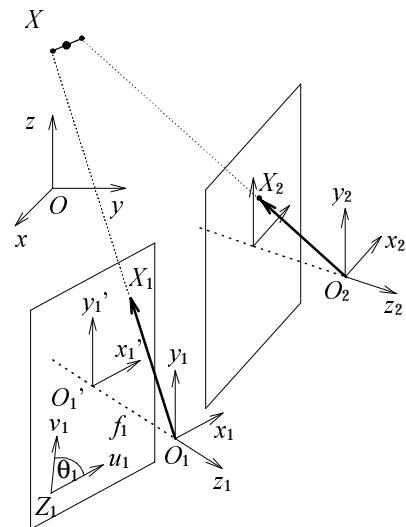
$$\mathbf{R}_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi_x & \sin \varphi_x \\ 0 & -\sin \varphi_x & \cos \varphi_x \end{bmatrix}, \quad \mathbf{R}_y = \begin{bmatrix} \cos \varphi_y & 0 & -\sin \varphi_y \\ 0 & 1 & 0 \\ \sin \varphi_y & 0 & \cos \varphi_y \end{bmatrix}, \quad \mathbf{R}_z = \begin{bmatrix} \cos \varphi_z & \sin \varphi_z & 0 \\ -\sin \varphi_z & \cos \varphi_z & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (11.19)$$

Jak ukážeme později, lze v procesu relativní kalibrace určit právě matici \mathbf{R} a vektor \mathbf{b} . Známe-li matici \mathbf{R}_1 a vektor \mathbf{o}_1 určující polohu první kamery v globálním souřadném systému scény, pak můžeme hodnoty $\mathbf{R}_2, \mathbf{o}_2$ ze vztahu (11.1) určující polohu druhé kamery vyjádřit vztahy

$$\mathbf{o}_2 = \mathbf{R}_1\mathbf{b} + \mathbf{o}_1, \quad \mathbf{R}_2 = \mathbf{R}_1\mathbf{R}. \quad (11.20)$$

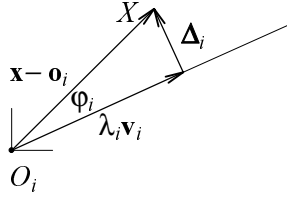
Hodnoty $\mathbf{R}_1, \mathbf{o}_1$ popisují vztah mezi souřadným systémem první kamery a globálním souřadným systémem scény. Tento vztah nelze v procesu relativní kalibrace určit. Někdy mohou být hodnoty $\mathbf{R}_1, \mathbf{o}_1$ známy předem. Jestliže tomu tak není, pak pro mnoho úloh postačí položit $\mathbf{R}_1 = \mathbf{I}$ a $\mathbf{o}_1 = \mathbf{0}$, což znamená, že globální souřadný systém scény ztotožníme se souřadným systémem první kamery a souřadnice bodů ve scéně tak měříme v souřadném systému první kamery.

Předpokládejme nyní nejprve, že kalibrace již byla provedena a že matice \mathbf{R}_i a vektor \mathbf{o}_i ($i=1,2$) z rovnice (11.1) jsou již pro obě kamery známy. Nejprve ukážeme, jak lze provést rekonstrukci trojrozměrných souřadnic bodu, známe-li souřadnice jeho průmětů v obrazech získaných oběma kamerami. Nechť X označuje rekonstruovaný bod. V souřadném systému (O, x, y, z) nechť je bod X reprezentován vektorem $\mathbf{x}=(x, y, z)$. Nechť X_i je obraz bodu X získaný i -tou kamerou. V souřadném systému (Z_i, u_i, v_i) nechť je bod X_i reprezentován vektorem $\mathbf{u}_i=(u_i, v_i, 1)^T$. V souřadném systému (O, x, y, z) je přímka $\langle O_i X_i \rangle$ je popsána rovnicí $\mathbf{o}_i + \lambda_i \mathbf{R}_i \mathbf{F}_i \mathbf{Q}_i \mathbf{u}_i$, kde λ_i je reálný parametr. Obě přímky $\langle O_i X_i \rangle$ ($i=1,2$) by se teoreticky měly protínat v bodě X . Opět je ovšem oprávněné předpokládat, že hodnoty souřadnic v obrazech nelze měřit zcela přesně. Proto nechť v tomto odstavci pojednávajícím o rekonstrukci souřadnic bodu značí \mathbf{u}_i prakticky zjištěné vektory



Obr. 11.4. Rekonstrukce polohy bodu.

souřadnic, jejichž složky jsou zatíženy šumem měření. Podobně bude nyní X_i značit prakticky zjištěnou polohu průmětu bodu X v obraze získaném i -tou kamerou. Je-li měření zatíženo chybami, pak již bod X obecně nebude ležet současně na obou přímkách $\langle O_i X_i \rangle$ (obr. 11.4). Polohu bodu X lze stanovit na základě podmínky, aby součet čtverců vzdáleností bodu X od obou přímek byl minimální. Hledáme tedy minimum



Obr. 11.5. Stanovení vzdálenosti bodu X od průmítacího paprsku.

$$\min_{\mathbf{x}} \sum_{i=1}^2 \text{dist}^2(\mathbf{x}, \mathbf{o}_i + \lambda_i \mathbf{R}_i \mathbf{F}_i \mathbf{Q}_i \mathbf{u}_i). \quad (11.21)$$

Pro stručnost matematických výrazů, které budou následovat, zavedeme vektor

$$\mathbf{v}_i = \frac{\mathbf{R}_i \mathbf{F}_i \mathbf{Q}_i \mathbf{u}_i}{|\mathbf{R}_i \mathbf{F}_i \mathbf{Q}_i \mathbf{u}_i|}. \quad (11.22)$$

Na základě vlastností skalárního součinu pro hodnotu parametru λ_i snadno nalezneme vztah

$$\lambda_i = |\mathbf{x} - \mathbf{o}_i| \cos \varphi_i = (\mathbf{x} - \mathbf{o}_i)^T \mathbf{v}_i. \quad (11.23)$$

Nechť Δ_i označuje vektor vzdálenosti bodu X od přímky $\langle O_i X_i \rangle$ (obr. 11.5). Je zřejmé, že $\Delta_i = (\mathbf{x} - \mathbf{o}_i) - \lambda_i \mathbf{v}_i$. Součet čtverců vzdáleností můžeme nyní vyjádřit pomocí vztahů

$$\begin{aligned} \varepsilon &= \sum_{i=1}^2 \Delta_i^T \Delta_i = \sum_{i=1}^2 [(\mathbf{x} - \mathbf{o}_i) - \lambda_i \mathbf{v}_i]^T [(\mathbf{x} - \mathbf{o}_i) - \lambda_i \mathbf{v}_i] = \\ &= \sum_{i=1}^2 [\mathbf{x}^T \mathbf{x} - 2\mathbf{o}_i^T \mathbf{x} - 2\lambda_i \mathbf{v}_i^T \mathbf{x} + \mathbf{o}_i^T \mathbf{o}_i + 2\lambda_i \mathbf{o}_i^T \mathbf{v}_i + \lambda_i^2 \mathbf{v}_i^T \mathbf{v}_i]. \end{aligned} \quad (11.24)$$

Derivováním získáme

$$\frac{\partial \varepsilon}{\partial \mathbf{x}} = 2 \sum_{i=1}^2 \left[\mathbf{x} - \mathbf{o}_i - \frac{\partial \lambda_i}{\partial \mathbf{x}} \mathbf{v}_i^T \mathbf{x} - \lambda_i \mathbf{v}_i + \frac{\partial \lambda_i}{\partial \mathbf{x}} \mathbf{o}_i^T \mathbf{v}_i + \lambda_i \frac{\partial \lambda_i}{\partial \mathbf{x}} \mathbf{v}_i^T \mathbf{v}_i \right]. \quad (11.25)$$

Z rovnice (11.23) vyplývá, že $\partial \lambda_i / \partial \mathbf{x} = \mathbf{v}_i$. Položíme-li $\partial \varepsilon / \partial \mathbf{x} = 0$, pak po úpravách obdržíme vztah

$$\sum_{i=1}^2 [\mathbf{x} - (\mathbf{v}_i^T \mathbf{x}) \mathbf{v}_i] = \sum_{i=1}^2 [\mathbf{o}_i - (\mathbf{o}_i^T \mathbf{v}_i) \mathbf{v}_i]. \quad (11.26)$$

Řešením rovnice (11.26) získáme pro hledaný vektor \mathbf{x} předpis

$$\mathbf{x} = \left[\sum_{i=1}^2 (\mathbf{I} - \mathbf{v}_i^T \mathbf{v}_i) \right]^{-1} \sum_{i=1}^2 [\mathbf{o}_i - (\mathbf{o}_i^T \mathbf{v}_i) \mathbf{v}_i]. \quad (11.27)$$

Nyní se vrátíme k problému kalibrace a ukážeme, jak lze zjistit matici \mathbf{R} a vektor \mathbf{b} , kterých lze podle rovnice (11.20) použít ke zjištění matice \mathbf{R}_2 a vektoru \mathbf{o}_2 (jak jsme již dříve uvedli, předpokládáme, že matici \mathbf{R}_1 a vektor \mathbf{o}_1 známe). Uvažujme bod X scény. V obrazech získaných první a druhou kamerou se tento bod zobrazuje na body X_1, X_2 . V souřadném systému (O_1, x_1, y_1, z_1) jsou směry přímek $\langle O_1 X_1 \rangle$, $\langle O_2 X_2 \rangle$ reprezentovány vektory $\mathbf{x}_1, \mathbf{R}\mathbf{x}_2$. Jestliže jsou polohy bodů X_1, X_2 v obrazech určeny přesně, pak leží přímky $\langle O_1 X_1 \rangle$, $\langle O_2 X_2 \rangle$, $\langle O_1 O_2 \rangle$ v jedné rovině. Podmínku koplanarity lze matematicky zapsat vztahem $\mathbf{x}_1 \cdot (\mathbf{b} \times (\mathbf{R}\mathbf{x}_2)) = 0$. Zápisem $[\mathbf{b}]_{\times}$ označíme antisymetrickou matici realizující vektorový součin, tj. matici, která pro každý vektor \mathbf{a} vyhovuje podmínce $[\mathbf{b}]_{\times} \mathbf{a} = \mathbf{b} \times \mathbf{a}$. Lze snadno ověřit, že platí

$$[\mathbf{b}]_{\times} = \begin{bmatrix} 0 & -b_z & b_y \\ b_z & 0 & -b_x \\ -b_y & b_x & 0 \end{bmatrix}. \quad (11.28)$$

Přepíšeme-li nyní podmínku koplanarity do maticového tvaru, obdržíme rovnici $\mathbf{x}_1^T[\mathbf{b}]_{\times}\mathbf{R}\mathbf{x}_2 = 0$. Dosaďme-li do podmínky rovnici (11.4), dostaneme

$$\mathbf{u}_1^T\mathbf{Q}_1^T\mathbf{F}_1^T[\mathbf{b}]_{\times}\mathbf{R}\mathbf{F}_2\mathbf{Q}_2\mathbf{u}_2 = 0, \quad (11.29)$$

případně
$$\mathbf{u}_1^T\mathbf{Q}_1^T\mathbf{F}_1^T[\mathbf{b}]_{\times}\mathbf{R}_z\mathbf{R}_y\mathbf{R}_x\mathbf{F}_2\mathbf{Q}_2\mathbf{u}_2 = 0 \quad (11.30)$$

nebo také
$$\mathbf{u}_1^T\mathbf{C}\mathbf{u}_2 = 0, \quad \text{kde } \mathbf{C} = \mathbf{Q}_1^T\mathbf{F}_1^T[\mathbf{b}]_{\times}\mathbf{R}\mathbf{F}_2\mathbf{Q}_2 = \mathbf{Q}_1^T\mathbf{F}_1^T[\mathbf{b}]_{\times}\mathbf{R}_z\mathbf{R}_y\mathbf{R}_x\mathbf{F}_2\mathbf{Q}_2. \quad (11.31)$$

Rovnice (11.29), (11.30), (11.31) bývají nazývány rovnicemi koplanarity. Podobně bývá matice \mathbf{C} z rovnice (11.31) nazývána maticí koplanarity. Každá dvojice obrazů jednoho kalibračního bodu popsaná souřadnicemi \mathbf{u}_1 v prvním obraze a souřadnicemi \mathbf{u}_2 v druhém obraze dává vzniknout jedné rovnici koplanarity. Je-li k dispozici q kalibračních bodů, pak lze sestavit soustavu q rovnic koplanarity. Této soustavě lze pak použít k řešení neznámých vnějších a vnitřních parametrů kamery, které jsou obsaženy v maticích $\mathbf{Q}_1, \mathbf{Q}_2, \mathbf{F}_1, \mathbf{F}_2, [\mathbf{b}]_{\times}, \mathbf{R}$.

Z rovnice (11.29) vyplývá, že řešení bude možné provést až na měřítko vektoru \mathbf{b} . Jestliže je totiž \mathbf{b} vektor, který vyhovuje rovnici (11.29), pak také vektor $\mu\mathbf{b}$, kde μ je reálný násobitel, rovnici vyhovuje. Abychom úlohu specifikovali jednoznačně, zvolíme doplňující podmínku. Můžeme např. požadovat $|\mathbf{b}|=d$, kde d je zvolená hodnota (např. $d=1$). V případě euklidovské metriky lze pak podmínku rozepsat do tvaru

$$b_x^2 + b_y^2 + b_z^2 = d^2. \quad (11.32)$$

Ze skutečnosti, že délku vektoru \mathbf{b} je zapotřebí zvolit, vyplývá, že rekonstrukci scény bude možné provést až na měřítko. To znamená, že pro různé hodnoty d budeme dostávat podobné avšak různě veliké scény. Problém lze jednoduše vyřešit tak, že kalibraci provedeme při $|\mathbf{b}|=1$. Známe-li pak skutečnou vzdálenost mezi některými dvěma body ve scéně, můžeme stanovit měřítko, kterým je třeba dodatečně násobit všechny souřadnice ve scéně (rozměry scény) tak, aby požadované vzdálenosti mezi body bylo skutečně dosaženo.

Z rovnice (11.29) vidíme, že při relativní kalibraci lze stanovit vzájemnou polohu kamer a vnitřní parametry kamery. Vzájemná poloha kamer je popsána maticí \mathbf{R} a vektorem \mathbf{b} . Protože však může existovat nejvýše 7 nezávislých rovnic koplanarity (podrobný důkaz tohoto tvrzení neuvádíme, poněvadž by přesáhl rámec tohoto textu), lze pomocí rovnic koplanarity a s využitím doplňující podmínky (11.32) zjistit nejvýše osm parametrů. To však nemusí být na závadu. Často lze totiž předpokládat, že se některé parametry (zejména vnitřní parametry kamery obsažené v maticích $\mathbf{Q}_1, \mathbf{Q}_2$) v čase nemění a lze je proto předem stanovit předchozím měřením (kalibrací) v laboratorních podmínkách.

Řešení soustavy kalibračních rovnic je obtížné, protože se jedná o nelineární problém. V posledním desetiletí bylo prezentováno několik přístupů k jeho řešení. Zde nastíníme jednu z možných variant. Vyjdeme z rovnic ve tvaru (11.30). Předpokládáme, že máme k dispozici q kalibračních bodů. Souřadnice průmětů i -tého kalibračního bodu v obraze první a druhé kamery označíme postupně $\mathbf{u}_1^{(i)}, \mathbf{u}_2^{(i)}$. Každá dvojice $\mathbf{u}_1^{(i)}, \mathbf{u}_2^{(i)}$ dává vzniknout jedné rovnici koplanarity. Dostaneme tak celkem q rovnic koplanarity, kterých využijeme spolu s podmínkou (11.32) ke stanovení hodnoty parametrů $f_1, f_2, \varphi_x, \varphi_y, \varphi_z, b_x, b_y, b_z$. Je zřejmé, že nejmenší hodnota q , při níž lze úlohu řešit, je 7. Prakticky však bývá q vyšší. Předeterminovaností soustavy lze opět využít k redukci vlivu šumu, kterým může být zatíženo měření souřadnic vektorů $\mathbf{u}_1^{(i)}, \mathbf{u}_2^{(i)}$. Hledané hodnoty parametrů uspořádáme do vektoru $\mathbf{p} = (f_1, f_2, \varphi_x, \varphi_y, \varphi_z, b_x, b_y, b_z)^T$. Pro dvojici obrazů i -tého kalibračního bodu můžeme pak rovnici (11.30) zapsat přehledně ve tvaru

$$F(\mathbf{u}_1^{(i)}, \mathbf{u}_2^{(i)}, \mathbf{p}) = 0. \quad (11.33)$$

Podmínka (11.32) spolu s q ($q>7$) rovnicemi (11.33) tvoří předeterminovaný nekonzistentní systém (nekonzistence vyplývá z přítomnosti šumu ve vstupních hodnotách). Nekonzistentní systém nemá přísně vzato žádné řešení. Lze pouze najít řešení, které systému v jistém smyslu vyhovuje nejlépe.

Můžeme například najít vektor parametrů \mathbf{p} tak, aby součet čtverců reziduí rovnic (11.33) byl minimální. Hledáme tedy takovou hodnotu vektoru parametrů, která minimalizuje funkci

$$\Phi = \sum_{i=1}^q \left[F(\mathbf{u}_1^{(i)}, \mathbf{u}_2^{(i)}, \mathbf{p}) \right]^2, \quad (11.34)$$

a to za současného splnění podmínky (11.32). Hledání minima funkce za doplňujících podmínek lze provést metodou Lagrangeových multiplikátorů, která je popsána v dodatku C. Metoda převádí problém na problém nalezení prostého minima (bez doplňující podmínky) nové funkce, která však má větší počet argumentů. V našem případě je jistou komplikací, že je úloha nelineární. Nelineární úlohu minimalizace lze řešit např. gradientní metodou. Postup je popsán v dodatku D.

11.5 Epipolára

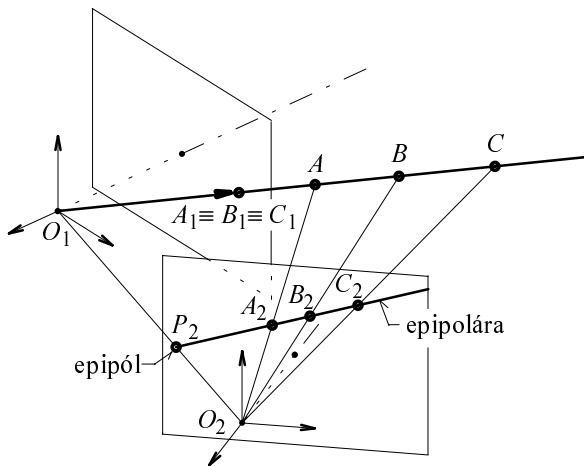
Uvažujme bod X scény. V obrazech získaných první a druhou kamerou jsou souřadnice průmětů tohoto bodu popsány vektory $\mathbf{u}_1 = (u_1, v_1, 1)$, $\mathbf{u}_2 = (u_2, v_2, 1)$. Dosadíme-li do rovnice (11.18) rovnici (11.4) zapsanou pro první a druhou kameru, snadno získáme vztah

$$\lambda_1 \mathbf{F}_1 \mathbf{Q}_1 \mathbf{u}_1 = \lambda_2 \mathbf{R} \mathbf{F}_2 \mathbf{Q}_2 \mathbf{u}_2 + \mathbf{b}. \quad (11.35)$$

Odtud máme

$$\lambda_2 \mathbf{u}_2 = \lambda_1 \mathbf{Q}_2^{-1} \mathbf{F}_2^{-1} \mathbf{R}^T \mathbf{F}_1 \mathbf{Q}_1 \mathbf{u}_1 - \mathbf{Q}_2^{-1} \mathbf{F}_2^{-1} \mathbf{R}^T \mathbf{b}. \quad (11.36)$$

Součin $\lambda_2 \mathbf{u}_2$ ve vztahu (11.36) reprezentuje v homogenních souřadnicích průmět bodu X v obraze druhé kamery. Rovnice (11.36) je homogenní parametrickou rovnicí přímky, v níž je parametrem hodnota λ_1 . Můžeme tedy učinit tento závěr: Jestliže máme v obraze první kamery bod, jehož souřadnice jsou $\mathbf{u}_1 = (u_1, v_1, 1)$, pak poloha jemu odpovídajícího bodu v obraze druhé kamery závisí na hodnotě λ_1



Obr. 11.6. Epipolára a epipól.

(použitím vztahu (11.4) lze snadno ověřit, že platí $\lambda_1 = -z_1/f_1$). Všechny možné polohy vytvoří v obraze získaném druhou kamerou přímku. Tato přímka se nazývá epipolárou. Pro $\lambda_1 = 0$ rovnice (11.36) určuje v obraze získaném druhou kamerou polohu průmětu středu projekce první kamery. Tento bod se nazývá epipól. Epipolára i epipól jsou vyobrazeny na obr. 11.6. Poznamenejme, že odvozené vztahy je samozřejmě možné také obrátit, což znamená, že k bodu v obraze získaném druhou kamerou je možné nalézt epipoláru v obraze získaném první kamerou. Totéž platí i pro epipól.

V případě kamer s rovnoběžnými osami (podkapitola 11.2) se rovnice (11.36) podstatně zjednoduší. V tomto případě totiž máme $\mathbf{F}_1 = \mathbf{F}_2$ (proto dále píšeme jen \mathbf{F}), $\mathbf{Q}_1 = \mathbf{Q}_2 = \mathbf{I}$ (jednotková matice) a $\mathbf{R} = \mathbf{I}$. Z rovnice (11.36) obdržíme vztah $\lambda_2 \mathbf{u}_2 = \lambda_1 \mathbf{u}_1 - \mathbf{F}^{-1} \mathbf{b}$. Uvážíme-li, že $\mathbf{b} = (b_x, 0, 0)$, pak z rovnice pro souřadnici z zjišťujeme, že $\lambda_2 = \lambda_1$ (dále proto píšeme jen λ). Parametrické rovnice epipoláry pak jsou $u_2 = u_1 - b_x / \lambda$, $v_2 = v_1$.

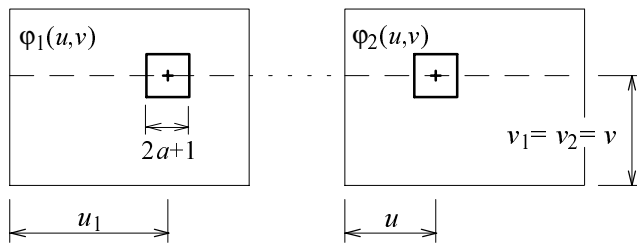
11.6 Automatizované hledání korespondence

Pro rekonstrukci polohy bodu je zapotřebí identifikovat jeho průměty v obraze získaném první a druhou kamerou. Zdá se, že je možné uvažovat i o automatizaci této úlohy. Lze postupovat např. v následujících krocích: 1) V obraze získaném první kamerou identifikujeme body zájmu např. některým z

detektorů rohů (kapitola 8.4). 2) Pro každý bod zájmu nalezený v prvním obraze určíme korespondující bod v druhém obraze. Druhý krok popíšeme podrobněji. Necht' X_1 je bod zájmu nalezený v prvním obraze a jeho souřadnice necht' jsou u_1, v_1 . Z předchozí podkapitoly víme, že korespondující bod X_2 bude v druhém obraze ležet na epipoláře. Předpokládáme, že kromě hodnoty λ_1 , která je parametrem, jsou všechny zbývající hodnoty na pravé straně rovnice (11.36) k dispozici. Můžeme proto stanovit rovnici epipoláry, na níž bod X_2 leží. Přesnou polohu bodu X_2 nalezneme tak, že systematicky prověříme všechny body na epipoláře (při praktické implementaci můžeme postupovat např. po celých pixelech). Označme $\varphi_1(u, v)$, $\varphi_2(u, v)$ obrazové funkce prvního a druhého obrazu. Necht' u, v jsou souřadnice prověřovaného bodu ve druhém obraze. Otázku, zda prověřovaný bod koresponduje s X_1 , můžeme rozhodnout na základě porovnání okolí bodu X_1 s okolím bodu prověřovaného. K porovnání lze použít např. vztahu (obr. 11.7)

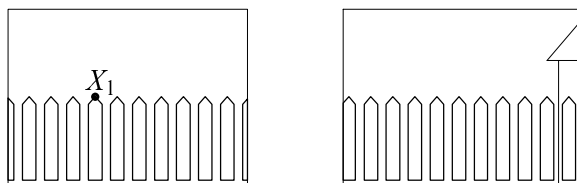
$$\text{Diff}(u, v) = \sum_{\Delta u = -a}^a \sum_{\Delta v = -a}^a [\varphi_1(u_1 + \Delta u, v_1 + \Delta v) - \varphi_2(u + \Delta u, v + \Delta v)]^2. \quad (11.37)$$

Je zřejmé, že vztah (11.37) porovnává okolí čtvercového tvaru o straně $2a+1$. Čím menší hodnota vyjde, tím jsou si okolí podobnější. Za bod korespondující s X_1 lze považovat takový bod X_2 , který splňuje následující kriteria: 1) hodnota vypočítaná podle vztahu (11.37) je pro X_2 minimální a menší než



Obr. 11.7. Hledání korespondence porovnáváním okolí bodu (kamery s rovnoběžnými optickými osami).

obrazy sejmuté z různých míst jsou přesně vzato různé a nelze obecně očekávat shodu příliš velkých oblastí ve smyslu vztahu (11.37). Vztah (11.37) dále předpokládá, že okénku velikosti $2a+1$ v prvním obraze odpovídá ve druhém obraze okénko téže velikosti a že hrany okének jsou v obou obrazech rovnoběžné se souřadnými osami. To je splněno pouze ve speciálních případech (např. v případě kamer s rovnoběžnými optickými osami z podkapitoly 11.2).



Obr. 11.8. K nejednoznačnosti řešení problému korespondence.

s bodem X_1 . Tím méně lze u podobných úloh očekávat řešení automatizované. Automatizované hledání korespondence je z výše uvedených důvodů úlohou, jejíž obecné řešení je obtížné. I když bylo pro automatizované hledání korespondence publikováno více postupů, nelze žádný z nich považovat za zcela úspěšný.

nějaký zvolený práh, 2) pro všechny ostatní body na epipoláře je hodnota (11.37) významně větší než pro bod X_2 . Druhé kritérium zohledňuje požadavek jednoznačnosti detekce.

Ačkoli právě popsany postup vypadá dosti přímočaře, skrývá mnohá úskalí. Prvním problémem je volba vhodné velikosti okénka. Je-li okénko malé, hrozí, že detekce nebude jednoznačná. Při velkém okénku se naopak významněji uplatní fakt, že

Otázka jednoznačnosti detekce korespondence je klíčová. Přísně vzato je zapotřebí se smířit s faktem, že mohou existovat případy, kdy jednoznačnou detekci zajistit nelze. Uvažme např. příklad dle obr. 11.8. Ani člověku se zde bez doplňujících informací nepodaří stanovit, který bod v druhém obraze koresponduje

12 Analýza obrazů proměnných v čase

Až dosud jsme předpokládali, že obrazy, které zpracováváme, jsou nepohyblivé. Některé úlohy však mají povahu dynamickou. V takovém případě pak často pracujeme nikoli s jedním, ale s celou sekvencí obrazů pořízených např. kamerou. Škála úloh, které přicházejí v úvahu, je rozmanitá. Úkolem například může být detekovat ve scéně pohybující se objekty, stanovit parametry jejich pohybu, případně objekty i rozpoznat. Jindy je naopak scéna statická, ale pohybuje se kamera, která scénu sleduje. Taková situace nastává např. u autonomních robotů, u nichž se žádá, aby se sami dokázali ve scéně orientovat. Cílem také může být stanovení geometrických parametrů sledovaných objektů nebo stanovení trajektorie kamery ve scéně. Velké šíři problémů, které při analýze obrazů proměnných v čase vyvstávají, odpovídá i široká škála metod, které jsou v literatuře pro jejich řešení popisovány. V tomto textu se omezíme pouze na některé častěji používané typické přístupy. Jedním z nástrojů, který se při zpracování obrazů proměnných v čase uplatňuje téměř pravidelně, je Kalmanův filtr. Popis Kalmanova filtru a jeho možných aplikací v diskutované třídě úloh uvádíme v podkapitolách 12.1 až 12.4. Kapitola 12.5 je věnována tzv. optickému toku, dalšímu nástroji, který lze dnes již považovat za klasický.

12.1 Diskrétní Kalmanův filtr

Uvažujme dynamický systém, tedy systém, jehož stav se mění v čase. Chování systému budeme vyšetřovat v okamžicích $t=0,1,2,\dots,\infty$. Předpokládáme, že stav systému v libovolném čase t je možné zcela popsat vektorem \mathbf{x}_t stavových proměnných. Posloupnost $\{\mathbf{x}_t\}$ je vektorovým stochastickým procesem. Dále předpokládáme, že chování systému v čase je popsáno rovnicí

$$\mathbf{x}_{t+1} = \Phi_{t+1|t}\mathbf{x}_t + \mathbf{D}_t\mathbf{d}_t + \Gamma_t\mathbf{w}_t. \quad (12.1)$$

Slovně lze rovnici (12.1) formulovat takto: Stav systému v čase $t+1$ popsán vektorem \mathbf{x}_{t+1} je závislý na stavu \mathbf{x}_t , na deterministickém budícím signálu \mathbf{d}_t a na šumu \mathbf{w}_t v čase t . Je zřejmé, že vektor \mathbf{x}_{t+1} je lineární kombinací prvků vektorů \mathbf{x}_t , \mathbf{d}_t , \mathbf{w}_t . Matice $\Phi_{t+1|t}$, \mathbf{D}_t , Γ_t obsahují koeficienty lineární kombinace. Indexem t je při tom zdůrazněno, že také tyto matice obecně mohou být proměnné v čase. Je zajímavé promyslet význam posledního členu na pravé straně rovnice (12.1). Neformálně řečeno vyjadřuje tento člen nedůvěru, v to, že pro popis systému platí rovnice $\mathbf{x}_{t+1} = \Phi_{t+1|t}\mathbf{x}_t + \mathbf{D}_t\mathbf{d}_t$. Čím větší je nedůvěra, tím více šumu přidáváme. Vidíme tedy, že Kalmanův filtr počítá s tím, že chování vyšetřovaného systému nemusí být známo zcela přesně. Neúplná znalost je vzata v úvahu přidáním šumu dynamiky systému. Pro úplnost poznamenejme, že z předpokladu (12.1) vyplývá, že posloupnost $\{\mathbf{x}_t\}$ je tzv. markovský proces. (V markovských procesech závisí stav v čase t pouze na stavu v čase $t-1$ a nikoli na historii, jak se systém do stavu v čase $t-1$ dostal.) Předpokládejme dále, že v každém čase t provádíme měření, kterým zjišťujeme, v jakém stavu se systém právě nachází. Obecně lze měřit libovolný počet hodnot, z nichž každá je nějakou lineární kombinací stavových proměnných. Hodnoty naměřené v čase t uspořádáme do vektoru \mathbf{z}_t . Modelem měření je rovnice

$$\mathbf{z}_t = \mathbf{H}_t\mathbf{x}_t + \mathbf{v}_t. \quad (12.2)$$

Matice \mathbf{H}_t je maticí měření, která popisuje, jak měřené veličiny závisí na hodnotách stavových proměnných. Měření je zkresleno šumem \mathbf{v}_t . Předpokládáme, že šumy $\mathbf{w}_t, \mathbf{v}_t$ jsou gaussovské vektorové náhodné procesy s nulovou střední hodnotou a s charakterem bílého šumu. Uvedené předpoklady jsou vyjádřeny rovnicemi

$$E\{\mathbf{w}_t\} = E\{\mathbf{v}_t\} = \mathbf{0}, \quad (12.3)$$

$$\text{cov}\{\mathbf{w}_t, \mathbf{w}_s\} = \mathbf{V}_{\mathbf{w},t}\delta(t-s), \quad \text{cov}\{\mathbf{v}_t, \mathbf{v}_s\} = \mathbf{V}_{\mathbf{v},t}\delta(t-s). \quad (12.4)$$

Výše uvedené rovnice platí pro všechna $t, s = 0, 1, 2, \dots$; $\mathbf{V}_{\mathbf{w},t}$ a $\mathbf{V}_{\mathbf{v},t}$ jsou disperzní matice šumu dynamiky systému a šumu měření

$$\mathbf{V}_{\mathbf{w},t} = \text{cov}\{\mathbf{w}_t, \mathbf{w}_t\}, \quad \mathbf{V}_{\mathbf{v},t} = \text{cov}\{\mathbf{v}_t, \mathbf{v}_t\}. \quad (12.5)$$

Předpokládáme, že šумы \mathbf{w}, \mathbf{v} jsou na sobě nezávislé a že jsou nezávislé také na počátečním stavu \mathbf{x}_0 . Pro všechna $t, s = 0, 1, 2, \dots$ tedy máme

$$\text{cov}\{\mathbf{w}_t, \mathbf{v}_s\} = \mathbf{0}, \quad (12.6)$$

$$\text{cov}\{\mathbf{x}_0, \mathbf{w}_t\} = \mathbf{0}, \quad \text{cov}\{\mathbf{x}_0, \mathbf{v}_t\} = \mathbf{0}. \quad (12.7)$$

Cílem výpočtu je provést v každém okamžiku t odhad $\hat{\mathbf{x}}_t$ stavového vektoru \mathbf{x}_t . Pro provedení odhadu by při tom měly být vzaty v úvahu všechny hodnoty, které byly až do doby t získány měřením. Chybu $\tilde{\mathbf{x}}_t$ odhadu můžeme stanovit jako rozdíl

$$\tilde{\mathbf{x}}_t = \mathbf{x}_t - \hat{\mathbf{x}}_t. \quad (12.8)$$

Odhad $\hat{\mathbf{x}}_t$ stanovíme tak, aby byly splněny následující vlastnosti

$$\mathbb{E}\{\tilde{\mathbf{x}}_t\} = \mathbf{0}, \quad \mathbb{E}\{\tilde{\mathbf{x}}_t^T \tilde{\mathbf{x}}_t\} = \min. \quad (12.9)$$

Vztah (12.9) je předpisem, v němž se odhad $\hat{\mathbf{x}}_t$ hledá na základě minimalizace tzv. bayesovského rizika. Zavedme podmíněnou hustotu pravděpodobnosti $p(\mathbf{x}_t | \mathbf{z}_0, \mathbf{z}_1, \dots, \mathbf{z}_t)$, což je hustota pravděpodobnosti vektorové náhodné proměnné \mathbf{x}_t za předpokladu, že v časech $0, 1, \dots, t$ byly naměřeny hodnoty $\mathbf{z}_0, \mathbf{z}_1, \dots, \mathbf{z}_t$. Tuto hustotu označíme zkráceně $p(\mathbf{x}_t | \mathbf{Z}_t)$. Lze ukázat, že v našem případě je požadavku (12.9) ekvivalentní postup, kdy je odhad $\hat{\mathbf{x}}_t$ nalezen jako hodnota maximalizující hodnotu $p(\mathbf{x}_t | \mathbf{Z}_t)$. Věříme, že čtenář shledá, že tento postup je racionální i sám o sobě, bez zdůrazňování skutečnosti, že jde o důsledek požadavku (12.9), což jsme zde pouze uvedli, ale nedokázali. Protože podrobná realizace naznačeného postupu a odvození výsledných vztahů pro Kalmanův filtr by přesahovaly možnosti tohoto textu, uvedeme zde pouze výsledek. Filtr počítá odhad $\hat{\mathbf{x}}_{t+1}$ v čase $t+1$ pomocí rekurentního předpisu využívajícího hodnot v čase t . Protože je předpis poněkud komplikovanější, bývá rozdělen do několika dílčích kroků. Kromě výpočtu odhadu $\hat{\mathbf{x}}_{t+1}$ je také stanovována disperzní matice $\mathbf{V}_{\tilde{\mathbf{x}}, t+1|t}$ chyby odhadu. Jako pomocná hodnota je dále v každém kroku počítána predikce $\mathbf{V}_{\tilde{\mathbf{x}}, t+1|t}$ disperzní matice chyby odhadu a matice \mathbf{K}_{t+1} zesílení. Výpočet pak může být realizován v následujících krocích

$$\mathbf{V}_{\tilde{\mathbf{x}}, t+1|t} = \Phi_{t+1|t} \mathbf{V}_{\tilde{\mathbf{x}}, t} \Phi_{t+1|t}^T + \Gamma_t \mathbf{V}_{\mathbf{w}, t} \Gamma_t^T, \quad (12.10)$$

$$\mathbf{K}_{t+1} = \mathbf{V}_{\tilde{\mathbf{x}}, t+1|t} \mathbf{H}_{t+1}^T \left[\mathbf{H}_{t+1} \mathbf{V}_{\tilde{\mathbf{x}}, t+1|t} \mathbf{H}_{t+1}^T + \mathbf{V}_{\mathbf{v}, t+1} \right]^{-1}, \quad (12.11)$$

$$\hat{\mathbf{x}}_{t+1|t} = \Phi_{t+1|t} \hat{\mathbf{x}}_t + \mathbf{D}_t \mathbf{d}_t, \quad (12.12)$$

$$\hat{\mathbf{x}}_{t+1} = \hat{\mathbf{x}}_{t+1|t} + \mathbf{K}_{t+1} \left[\mathbf{z}_{t+1} - \mathbf{H}_{t+1} \hat{\mathbf{x}}_{t+1|t} \right], \quad (12.13)$$

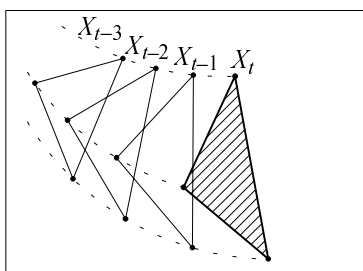
$$\mathbf{V}_{\tilde{\mathbf{x}}, t+1} = (\mathbf{I} - \mathbf{K}_{t+1} \mathbf{H}_{t+1}) \mathbf{V}_{\tilde{\mathbf{x}}, t+1|t}. \quad (12.14)$$

Jak je z uvedených vzorců vidět, je pro realizaci výpočtu zapotřebí znát hodnoty $\Phi_{t+1|t}$, \mathbf{D}_t , Γ_t , \mathbf{H}_t , $\mathbf{V}_{\mathbf{w}, t}$, $\mathbf{V}_{\mathbf{v}, t}$. Tyto hodnoty získáme na základě analýzy úlohy, v níž má být filtr použit (příklad je uveden v následující podkapitole). Pro spuštění algoritmu je dále nutné zadat počáteční hodnotu $\hat{\mathbf{x}}_0$ odhadu stavového vektoru a počáteční hodnotu $\mathbf{V}_{\tilde{\mathbf{x}}, 0}$ disperzní matice chyby odhadu v čase $t=0$. Protože je odhad neustále korigován měřením, není Kalmanův filtr příliš citlivý na přesnost nastavení počátečních hodnot. Nejsou-li k dispozici přesnější údaje, můžeme položit $\hat{\mathbf{x}}_0 = \mathbf{0}$ a za $\mathbf{V}_{\tilde{\mathbf{x}}, 0}$ vzít diagonální matici, v níž jsou na diagonále dostatečně veliké odhady počátečních disperzí jednotlivých prvků stavového vektoru.

12.2 Aplikace Kalmanova filtru při sledování pohybujících se objektů

Předpokládejme, že v časech $t = 0, 1, \dots$ kamera snímá obrazy scény a produkuje tak sekvenci obrazů I_1, I_2, \dots . Zatímco kamera je pevná, vyskytují se ve scéně pohybující se objekty. Pro jednoduchost

předpokládejme, že objekty zájmu jsou izolované body. Úkolem pak je tyto body detekovat a sledovat jejich pohyb v jednotlivých obrazech sekvence. Problém je ilustrován na obr. 12.1. Necht' X je bod zájmu. Průmět bodu X v obraze I_t označíme X_t ; (u_t, v_t) necht' jsou souřadnice bodu X_t v I_t . Předpokládáme, že hodnoty souřadnic lze získat měřením. Měření je však zatíženo chybami, a to např. přinejmenším proto, že souřadnice bodů v obrazech často měříme celočíselným počtem pixelů, zatímco přesné hodnoty jsou reálné (na chybách měření se ovšem mohou podílet i další vlivy). Pokud lze pro pohyb bodu v obraze formulovat nějaký model, pak je možné k odhadu přesnějších hodnot souřadnic bodů použít Kalmanova filtru. Filtr postupně odhaduje souřadnice bodu v čase $t=1,2,\dots$. Při tom pro



Obr.12.1. Sledování pohybu význačných bodů.

každé t bere v úvahu nejen všechna měření provedená až do doby t včetně, ale také teoretická očekávání o pohybu bodu, která jsou vtělena do modelu systému. Lze proto očekávat, že odhad poskytnutý Kalmanovým filtrem bude lepší než jednotlivé měření.

Nyní ukážeme konkrétní návrh Kalmanova filtru. Řekněme, že jsme zjistili, že pro všechna t jsou souřadnice u_t, v_t obrazu bodu zájmu v I_t nekorelované, a že proto můžeme pro každou souřadnici použít samostatný filtr (tento předpoklad je dosti častý a většinou také oprávněný). Každý sledovaný bod tak bude vyžadovat dva filtry: Jeden pro souřadnici u , druhý pro souřadnici v . Oddělené zpracování souřadnic je výhodné. Oproti případu, kdy by pro obě souřadnice byl použit filtr jediný, mají matice ve vztazích (12.1), (12.2), (12.10) až (12.14) poloviční rozměry, a výpočet je proto rychlejší, a to i přesto, že se v případě dvou filtrů provádí dvakrát (uvažte např. násobení čtvercových matic - zvětší-li se rozměr matice dvakrát, vzroste počet operací osmkrát). Protože pro obě souřadnice použijeme dvou shodně konstruovaných filtrů, budeme v dalším textu uvádět pouze vztahy pro souřadnici u . Vztahy pro souřadnici v budou analogické. Řekněme dále, že o tom, jakými zákonitostmi se pohyb sledovaných bodů v obrazech řídí, nemáme přesné poznatky. Musíme proto navrhnout pokud možno dosti obecný model. Takovým může být např. model, v němž jako stavové proměnné volíme souřadnici u , její první diferenci u' a druhou diferenci u'' a v němž předpokládáme, že druhá diference (zrychlení) je konstantní. Stavový vektor je tedy $\mathbf{x}_t = (u_t, u'_t, u''_t)^T$. Předpokládáme, že chování systému dostatečně výstižně popisují rovnice

$$u_{t+1} = u_t + \Delta t u'_t + \frac{1}{2} (\Delta t)^2 u''_t, \quad (12.15)$$

$$u'_{t+1} = u'_t + \Delta t u''_t, \quad (12.16)$$

$$u''_{t+1} = u''_t. \quad (12.17)$$

V našem případě je $\Delta t=1$. Měřenou hodnotou je pouze hodnota souřadnice (vektor \mathbf{z}_t obsahuje pouze jedinou složku). Deterministické buzení systému v našem případě není, a proto lze ve všech vzorcích členy $\mathbf{D}_t \mathbf{d}_t$ vypustit. Matice $\Phi_{t+1|t}$, \mathbf{H}_t jsou v popisovaném případě konstantní v čase a mají tvar

$$\Phi_{t+1|t} = \begin{bmatrix} 1 & 1 & 0.5 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{H}_t = [1 \ 0 \ 0]. \quad (12.18)$$

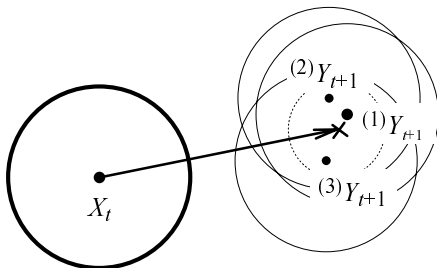
Podle míry, jak přesně rovnice (12.15) až (12.17) vystihují skutečný pohyb bodů v obraze, nastavíme hodnoty v maticích $\mathbf{V}_{w,t}$ a Γ_t , které popisují nejistotu modelu. Obě matice zde volíme diagonální a v čase konstantní

$$\mathbf{V}_{w,t} = \begin{bmatrix} \sigma_x^2 & 0 & 0 \\ 0 & \sigma_v^2 & 0 \\ 0 & 0 & \sigma_a^2 \end{bmatrix}, \quad \Gamma_t = \begin{bmatrix} \gamma_x & 0 & 0 \\ 0 & \gamma_v & 0 \\ 0 & 0 & \gamma_a \end{bmatrix}. \quad (12.19)$$

Hodnoty σ_x^2 , σ_v^2 , σ_a^2 jsou variance šumů, který je podle vztahu (12.1) přičítán k pravé straně rovnice pro souřadnici, k pravé straně rovnice pro rychlost a k pravé straně rovnice pro zrychlení. Jak je vidět z tvaru matice $\mathbf{V}_{w,t}$, předpokládáme v našem případě, že jednotlivé složky šumového vektoru jsou na sobě nezávislé. Jak je dále vidět z rovnice (12.1), upravují hodnoty γ_x , γ_v , γ_a velikost šumu. Matice $\mathbf{V}_{v,t}$ obsahuje pouze hodnotu variance σ_m^2 šumu měření a je tedy $\mathbf{V}_{v,t} = [\sigma_m^2]$. Návrh konkrétních hodnot obsažených v maticích $\mathbf{V}_{w,t}$, $\mathbf{\Gamma}_b$, $\mathbf{V}_{v,t}$ je obtížný. I když se návrh může opírat o nějakou logickou úvahu, je vždy na místě experimentální ověření. Nevhodně zvolené hodnoty mohou vést ke zbytečnému snížení účinnosti filtru. Pro rozběhnutí filtru je dále zapotřebí zvolit počáteční hodnoty. Počáteční hodnotu stavového vektoru můžeme volit např. $\hat{\mathbf{x}}_0 = (u_0, 0, 0)$, kde u_0 je souřadnice zjištěná při prvním výskytu sledovaného bodu. Jako počáteční hodnotu $\mathbf{V}_{\hat{\mathbf{x}},0}$ disperzní matice chyby odhadu můžeme vzít diagonální matici, kde na diagonálu postupně umístíme předpokládané disperze chyby počátečního odhadu souřadnice, rychlosti a zrychlení. Není-li k dispozici žádný apriorní předpoklad, pak na diagonálu jednoduše umístíme dostatečně velké hodnoty.

Nakonec ještě popíšeme samotný algoritmus sledování pohybu bodů. Uvažujme po sobě jdoucí obrazy I_t , I_{t+1} sejmuté v časech t , $t+1$. Řekneme, že bod X_t z obrazu I_t koresponduje s bodem X_{t+1} z obrazu I_{t+1} , jestliže X_t , X_{t+1} jsou obrazy téhož bodu scény. Sledování bodů lze realizovat pomocí hledání korespondujících párů bodů v po sobě jdoucích obrazech. Předpokládáme, že časový interval, v němž jsou obrazy snímány, je dostatečně krátký, a že jsou proto po sobě následující obrazy dosti podobné. Algoritmus sledování může pracovat např. takto: V obrazech I_t , I_{t+1} nalezneme body zájmu. K tomu lze použít např. některého z detektorů rohů (kapitola 8.4). Nechť X_t je bod zájmu nalezený v obraze I_t . Předpokládejme, že odhad \hat{u}_t, \hat{v}_t souřadnic tohoto bodu získaný Kalmanovým filtrem podle vztahu (12.13) je již znám. V obraze I_{t+1} nyní hledáme bod X_{t+1} korespondující s X_t . Vztahy (12.12), (12.15)

poskytují predikci $\hat{u}_{t+1|t}, \hat{v}_{t+1|t}$ polohy bodu X_{t+1} v



Obr.12.2. Výběr korespondujícího bodu z množiny kandidátů.

obraze I_{t+1} . Tato predikce je verifikována nalezením skutečné polohy korespondujícího bodu. Skutečnou polohu korespondujícího bodu lze najít takto: Body zájmu nalezené v I_{t+1} , jejichž vzdálenost od predikované polohy je menší než jistá prahová hodnota, jsou považovány za kandidáty na korespondenci a korespondence mezi X_t a každým z kandidátů je dále prověřována. (Na obr. 12.2 se v blízkosti predikované polohy vyznačené křížkem vyskytují celkem tři kandidáti na korespondenci s bodem X_t ; prahová hodnota vzdálenosti je vyznačena čárkovanou kružnicí.) Vhodnou prahovou hodnotu vzdálenosti lze odhadnout s využitím disperzní matice $\mathbf{V}_{\hat{\mathbf{x}},t}$ chyby odhadu

stavového vektoru. Nechť Y_{t+1} je možným kandidátem. Posouzení zda Y_{t+1} s X_t skutečně koresponduje je provedeno na základě porovnání hodnot zvolených jasových příznaků (kapitola 9.8, 9.9) vypočítaných pro jisté okolí každého z bodů testované dvojice (na obr. 12.2 jsou znázorněna okolí kruhová). Alternativně by též bylo možné přímo porovnat okolí bodů X_t , Y_{t+1} podobně, jak jsme prováděli v kapitole 11.6. Nechť $\mathbf{a}(X_t)$, $\mathbf{a}(Y_{t+1})$ jsou vektory příznaků vypočítané pro body X_t , Y_{t+1} . Hodnota $|\mathbf{a}(X_t) - \mathbf{a}(Y_{t+1})|$ umožňuje posoudit míru korespondence. Za bod korespondující s X_t považujeme toho z kandidátů, pro kterého je uvedená délka dostatečně malá a významně menší než pro kandidáty zbývající. Jestliže je bod korespondující s X_t nalezen, jsou známy jeho souřadnice, což v terminologii Kalmanova filtru znamená, že bylo provedeno měření. Na základě takto provedeného měření pak vztah (12.13) poskytuje zpřesněný odhad $\hat{u}_{t+1}, \hat{v}_{t+1}$, v němž je redukován vliv šumu. Je samozřejmé, že v průběhu času mohou některé body zájmu z obrazu zmizet. To může nastat buď proto, že bod opustí zorné pole kamery nebo proto, že je zakryt nějakým objektem scény. Nové body se naopak mohou objevit.

12.3 Alfa, beta filtr pro sledování pohybu bodu

Předpokládejme nyní, že matice $\Phi_{t+1|t}$, \mathbf{D}_t , Γ_t , \mathbf{H}_t , $\mathbf{V}_{w,t}$, $\mathbf{V}_{v,t}$ v Kalmanově filtru jsou v čase konstantní (ve zbytku této podkapitoly již proto u těchto matic index t nebudeme používat). Praktická zkušenost s použitím Kalmanova filtru ukazuje, že se hodnoty v maticích \mathbf{K}_t , $\mathbf{V}_{\hat{x},t}$, $\mathbf{V}_{\hat{x},t+1|t}$ mění nejvíce na začátku výpočtu a s rostoucím t se postupně ustalují. Nabízí se proto myšlenka provádět výpočet s pevnou hodnotou matice \mathbf{K} , která odpovídá ustálenému stavu. Očekáváme, že tento postup sice na začátku výpočtu povede k jistému snížení účinnosti filtru, to však ale bude na druhé straně vyváženo jednodušším (a tudíž rychlejším) výpočtem. Zavedme tedy

$$\mathbf{V}_{\hat{x}} = \lim_{t \rightarrow \infty} \mathbf{V}_{\hat{x},t}, \quad \hat{\mathbf{V}}_{\hat{x}} = \lim_{t \rightarrow \infty} \mathbf{V}_{\hat{x},t+1|t}, \quad \mathbf{K} = \lim_{t \rightarrow \infty} \mathbf{K}_t. \quad (12.20)$$

Položme dále $\mathbf{Q} = \Gamma_t \mathbf{V}_{w,t} \Gamma_t^T$. Rovnice (12.10) až (12.14) můžeme nyní přepsat do tvaru

$$\hat{\mathbf{V}}_{\hat{x}} = \Phi \mathbf{V}_{\hat{x}} \Phi^T + \mathbf{Q}, \quad (12.21)$$

$$\mathbf{K} = \hat{\mathbf{V}}_{\hat{x}} \mathbf{H}^T (\mathbf{H} \hat{\mathbf{V}}_{\hat{x}} \mathbf{H}^T + \mathbf{V}_v)^{-1}, \quad (12.22)$$

$$\hat{\mathbf{x}}_{t+1|t} = \Phi \hat{\mathbf{x}}_t + \mathbf{D} \mathbf{d}_t, \quad (12.23)$$

$$\hat{\mathbf{x}}_{t+1} = \hat{\mathbf{x}}_{t+1|t} + \mathbf{K} [\mathbf{z}_{t+1} - \mathbf{H} \hat{\mathbf{x}}_{t+1|t}], \quad (12.24)$$

$$\mathbf{V}_{\hat{x}} = (\mathbf{I} - \mathbf{K} \mathbf{H}) \hat{\mathbf{V}}_{\hat{x}}. \quad (12.25)$$

Do rovnice (12.21) postupně dosadíme rovnice (12.25) a (12.22). Dostaneme

$$\hat{\mathbf{V}}_{\hat{x}} = \Phi (\mathbf{I} - \mathbf{K} \mathbf{H}) \hat{\mathbf{V}}_{\hat{x}} \Phi^T + \mathbf{Q} = \Phi \left[\mathbf{I} - \hat{\mathbf{V}}_{\hat{x}} \mathbf{H}^T (\mathbf{H} \hat{\mathbf{V}}_{\hat{x}} \mathbf{H}^T + \mathbf{V}_v)^{-1} \mathbf{H} \right] \hat{\mathbf{V}}_{\hat{x}} \Phi^T + \mathbf{Q}. \quad (12.26)$$

Úpravou rovnice (12.26) pak snadno dostáváme rovnici

$$\hat{\mathbf{V}}_{\hat{x}} = \Phi \hat{\mathbf{V}}_{\hat{x}} \Phi^T - \Phi \hat{\mathbf{V}}_{\hat{x}} \mathbf{H}^T (\mathbf{H} \hat{\mathbf{V}}_{\hat{x}} \mathbf{H}^T + \mathbf{V}_v)^{-1} \mathbf{H} \hat{\mathbf{V}}_{\hat{x}} \Phi^T + \mathbf{Q}. \quad (12.27)$$

Poznamenejme, že rovnice ve tvaru, ke kterému jsme dospěli, bývá nazývána rovnicí Ricattiho. Rovnice (12.27) je rovnicí, z níž lze stanovit matici $\hat{\mathbf{V}}_{\hat{x}}$ (ostatní hodnoty jsou známé). Známe-li tuto matici, pak dle vztahu (12.22) již snadno stanovíme matici Kalmanova zisku.

Ilustrujme popsany postup na konkrétním případě. Podobně jako v předchozí podkapitole navrhneme filtr pro odhad polohy bodů v obraze. Opět předpokládáme, že souřadnice bodů jsou nekorelované, a navrhneme proto pro každou souřadnici oddělený filtr. Protože jsou filtry pro obě souřadnice konstruovány shodně, omezíme se při výkladu pouze na jedinou souřadnici, kterou označíme u . Pro jednoduchost tentokrát uvažujeme model s konstantní rychlostí. Stavový vektor volíme $\mathbf{x}_t = (u_t, u_t')^T$. Předpokládáme, že pohyb bodů dostatečně přesně popisují rovnice $u_{t+1} = u_t + \Delta t u_t'$, $u_{t+1}' = u_t'$. Označme u_t^* hodnotu souřadnice získanou měřením v čase t . Vektor měření pak je $\mathbf{z}_t = (u_t^*)$. Matice filtru mají tvar

$$\Phi = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{H} = [1 \quad 0], \quad \mathbf{Q} = \begin{bmatrix} q_{11} & q_{12} \\ q_{12} & q_{22} \end{bmatrix}, \quad \mathbf{V}_v = [\sigma_m^2]. \quad (12.28)$$

Zatím neznámou matici $\hat{\mathbf{V}}_{\hat{x},t}$ předpokládáme ve tvaru

$$\hat{\mathbf{V}}_{\hat{x}} = \begin{bmatrix} v_{11} & v_{12} \\ v_{12} & v_{22} \end{bmatrix}. \quad (12.29)$$

Dále vypočteme

$$\Phi \hat{\mathbf{V}}_{\hat{x}} \Phi^T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_{11} & v_{12} \\ v_{12} & v_{22} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} v_{11} + 2v_{12} + v_{22} & v_{12} + v_{22} \\ v_{12} + v_{22} & v_{22} \end{bmatrix}. \quad (12.30)$$

Z rovnice (12.27) máme

$$\mathbf{Q} = \begin{bmatrix} -(2v_{12} + v_{22}) & -v_{22} \\ -v_{22} & 0 \end{bmatrix} + \frac{1}{v_{11} + \sigma_m^2} \Phi \hat{\mathbf{V}}_{\tilde{\mathbf{x}}} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \hat{\mathbf{V}}_{\tilde{\mathbf{x}}} \Phi^T. \quad (12.31)$$

Úpravou (12.31) dále dostaneme

$$\mathbf{Q} = \begin{bmatrix} -2v_{12} - v_{22} + \frac{(v_{11} + v_{12})^2}{v_{11} + \sigma_m^2} & -v_{12} + \frac{(v_{11} + v_{12})v_{12}}{v_{11} + \sigma_m^2} \\ -v_{22} + \frac{(v_{11} + v_{12})v_{12}}{v_{11} + \sigma_m^2} & \frac{v_{12}^2}{v_{11} + \sigma_m^2} \end{bmatrix}. \quad (12.32)$$

Známe-li hodnoty prvků matice \mathbf{Q} , můžeme z rovnice (12.32) vypočítat hodnoty v_{11}, v_{12}, v_{22} , což jsou prvky matice $\hat{\mathbf{V}}_{\tilde{\mathbf{x}}}$. Jistou komplikací ovšem je, že rovnice (12.32) je nelineární, a výpočet je proto poněkud nepříjemný. Jestliže je matice $\hat{\mathbf{V}}_{\tilde{\mathbf{x}}}$ vypočtena, můžeme snadno stanovit matici Kalmanova zisku. Z rovnice (12.22) dostaneme

$$\mathbf{K} = \frac{\hat{\mathbf{V}}_{\tilde{\mathbf{x}}} \mathbf{H}^T}{\mathbf{H} \hat{\mathbf{V}}_{\tilde{\mathbf{x}}} \mathbf{H}^T + \sigma_m^2} = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}. \quad (12.33)$$

Položme $\alpha = k_1$, $\beta = \Delta t k_2$. Rozepsáním vztahů (12.23) (člen $\mathbf{D}d_t$ opět odpadá) a (12.24) obdržíme

$$u_{t+1|t} = u_t + \Delta t u'_t, \quad (12.34)$$

$$u_{t+1} = u_{t+1|t} + \alpha \left[u_{t+1}^* - u_{t+1|t} \right], \quad (12.35)$$

$$u'_{t+1} = u'_{t+1|t} + \frac{\beta}{\Delta t} \left[u_{t+1}^* - u_{t+1|t} \right]. \quad (12.36)$$

Prakticky se při sledování pohybu bodů hodnoty α, β počítají pouze při zřizování filtru pro nově se objevivší bod. Navíc lze často stejných hodnot α, β využít pro celou skupinu bodů nebo dokonce pro celou třídu úloh. Výpočet prováděný během sledování se tak redukuje na pouhé opakované vyčíslování vztahů (12.34) až (12.36). Nízká složitost výpočtu je výhodná zejména pro úlohy v reálném čase. Na závěr této podkapitoly poznamenejme, že podobným postupem, jako jsme odvodili vztahy pro případ, kdy rychlost předpokládáme konstantní, by bylo možné odvodit také vztahy pro případ, kdy je konstantní zrychlení. Obdrželi bychom tzv. α, β, γ filtr. Odvození je však bohužel poněkud zdlouhavé a výsledné vztahy dosti komplikované.

12.4 Sledování objektů v obrazech získaných pohybující se kamerou

Předpokládejme nyní, že sledovaná scéna je pevná, ale že se pohybuje kamera, která scénu sleduje. Nejprve promysleme, jakými zákonitostmi se řídí pohyb průmětů bodů sledované scény po obraze. K tomu využijeme terminologie a symboliky zavedené v kapitole 11.1. Nechť bod O_t (reprezentovaný v souřadném systému scény vektorem \mathbf{o}_t) popisuje polohu počátku souřadného systému kamery v čase t (tj. polohu středu projekce) a matice \mathbf{R}_t nechť popisuje natočení souřadného systému kamery v čase t vzhledem k souřadnému systému scény. Zavedeme vektor posunutí $\mathbf{b}_{t,s}$ a matici rotace $\mathbf{R}_{t,s}$ tak, aby pro transformaci ze souřadného systému kamery v čase s do souřadného systému kamery v čase t platilo (následující vztah je zobecněním vztahu (11.18))

$$\mathbf{x}_t = \mathbf{R}_{t,s} \mathbf{x}_s + \mathbf{b}_{t,s}. \quad (12.37)$$

Uvažujme bod X scény. V obrazech získaných kamerou v časech $t, t+1$ jsou souřadnice průmětů tohoto bodu popsány vektory $\mathbf{u}_t, \mathbf{u}_{t+1}$ (kapitola 11.1). Z rovnic (11.4) a (12.37) snadno získáme vztah

$$\lambda_t \mathbf{F}_t \mathbf{Q}_t \mathbf{u}_t = \lambda_{t+1} \mathbf{R}_{t,t+1} \mathbf{F}_{t+1} \mathbf{Q}_{t+1} \mathbf{u}_{t+1} + \mathbf{b}_{t,t+1}. \quad (12.38)$$

Odtud máme

$$\mathbf{u}_{t+1} = \frac{\lambda_t}{\lambda_{t+1}} \mathbf{Q}_{t+1}^{-1} \mathbf{F}_{t+1}^{-1} \mathbf{R}_{t,t+1}^{-1} \mathbf{F}_t \mathbf{Q}_t \mathbf{u}_t - \frac{1}{\lambda_{t+1}} \mathbf{Q}_{t+1}^{-1} \mathbf{F}_{t+1}^{-1} \mathbf{R}_{t,t+1}^{-1} \mathbf{b}_{t,t+1}. \quad (12.39)$$

Je patrné, že rovnice (12.39) je ve tvaru, jaký je předpokládán Kalmanovým filtrem. Stavovým vektorem je vektor \mathbf{u}_t , vektorem deterministického buzení je vektor $\mathbf{b}_{t,t+1}$. Porovnáním rovnice (12.39) s rovnicí (12.1) máme

$$\Phi_{t+1|t} = \frac{\lambda_t}{\lambda_{t+1}} \mathbf{Q}_{t+1}^{-1} \mathbf{F}_{t+1}^{-1} \mathbf{R}_{t,t+1}^{-1} \mathbf{F}_t \mathbf{Q}_t, \quad \mathbf{D}_t = -\frac{1}{\lambda_{t+1}} \mathbf{Q}_{t+1}^{-1} \mathbf{F}_{t+1}^{-1} \mathbf{R}_{t,t+1}^{-1}. \quad (12.40)$$

Praktické použití právě odvozeného exaktního modelu však bohužel není jednoduché. Ke stanovení matic $\Phi_{t+1|t}$ a \mathbf{D}_t jsou potřebné matice \mathbf{Q}_{t+1} , \mathbf{F}_{t+1} , $\mathbf{R}_{t,t+1}$ a vektor $\mathbf{b}_{t,t+1}$, které však v okamžiku, kdy je odhad vektoru \mathbf{u}_{t+1} zapotřebí provést, obvykle nejsou k dispozici. Často jsou naopak teprve na základě tohoto odhadu (provedeného pro dostatečný počet bodů) vypočítány. Proto jsou i v případě pohybující se kamery častěji používány přibližné modely popsané v podkapitolách 12.2, 12.3.

12.5 Optický tok

Uvažujeme opět obrazy, které se mění v čase, a to buď proto, že se pohybují objekty ve scéně nebo proto, že se pohybuje kamera. Na rozdíl od předchozích podkapitol však nyní předpokládáme, že se obrazy v čase mohou měnit zcela spojitě. Necht' X je nějaký bod scény. V čase t necht' se tento bod promítá do bodu, který je v souřadné soustavě obrazu reprezentován vektorem (u, v) . V čase $t+\delta t$ se tentýž bod X scény zobrazuje do bodu $(u+\delta u, v+\delta v)$, kde $\delta u = u'\delta t$, $\delta v = v'\delta t$ a $u' = du/dt$, $v' = dv/dt$. Představujeme si, že protože je bod $(u+\delta u, v+\delta v)$ průmětem téhož bodu X , jako byl dříve bod (u, v) , nabývá bod $(u+\delta u, v+\delta v)$ v čase $t+\delta t$ stejného jasu, jako byl jas v bodě (u, v) v čase t . Necht' $f(u, v, t)$ označuje funkci jasu (obrazovou funkci), která je nyní proměnná v čase. Rozvíňme tuto funkci v okolí bodu (u, v, t) v Taylorovu řadu. Dostaneme:

$$f(u+\delta u, v+\delta v, t+\delta t) = f(u, v, t) + \frac{\partial f}{\partial u} \delta u + \frac{\partial f}{\partial v} \delta v + \frac{\partial f}{\partial t} \delta t + \dots \quad (12.41)$$

Tečky v rovnici (12.41) naznačují členy vyšších řádů. Podle předchozí úvahy má být

$$f(u+\delta u, v+\delta v, t+\delta t) = f(u, v, t). \quad (12.42)$$

Dosadíme-li do rovnice (12.42) vztah (12.41), v němž zanedbáme členy vyšších řádů, vydělíme-li rovnicí hodnotou δt a předpokládáme-li $\delta t \rightarrow 0$, pak z (12.42) obdržíme rovnici

$$\frac{\partial f}{\partial u} \frac{du}{dt} + \frac{\partial f}{\partial v} \frac{dv}{dt} + \frac{\partial f}{\partial t} = 0. \quad (12.43)$$

Zaved'me dále označení $\frac{\partial f}{\partial u} = f_u$, $\frac{\partial f}{\partial v} = f_v$, $\frac{\partial f}{\partial t} = f_t$. (12.44)

Rovnici (12.43) lze pak přepsat do tvaru

$$f_u u' + f_v v' + f_t = 0. \quad (12.45)$$

Rovnice (12.45) bývá nazývána rovnicí optického toku. Její použití je následující: Funkce f_u , f_v , f_t jsou známy. Známe totiž funkci $f(u, v, t)$ (je sejmuta např. kamerou) a zřejmě ji také dokážeme derivovat, a to podle okolností buď analyticky nebo spíše numericky. Neznámými jsou tedy hodnoty rychlostí $u'(u, v)$, $v'(u, v)$ v jednotlivých bodech obrazu. Nalezneme-li funkce $u'(u, v)$, $v'(u, v)$, pak na jejich základě můžeme usuzovat např. na tvar pohybujících se objektů. Objekty různých tvarů totiž produkují různá rychlostní pole. Rovnice (12.45) však bohužel sama o sobě ke zjištění rychlostního pole nestačí, a je proto nutné formulovat doplňující podmínky. Často bývá jako doplňující uváděna podmínka, aby se rychlostní pole po ploše obrazu měnilo co nejmírněji. (Přísně vzato je však uvedený požadavek poněkud spekulativní. Např. na kontuře, kde je pohyblivý objekt překrýván objektem pevným, určitě může být změna rychlostního pole náhlá.) Požadavek „mírné změny“ rychlostního pole lze vzít v úvahu tak, že hledáme takové funkce $u'(u, v)$, $v'(u, v)$, které vyhovují rovnici (12.45) a při tom současně minimalizují funkcionál

$$F(u', v') = \iint_{\Omega} \left[(u_u'^2 + u_v'^2) + (v_u'^2 + v_v'^2) \right] dx dy. \quad (12.46)$$

Minimalizaci funkcionálu s doplňující podmínkou ve tvaru rovnosti (rovnice (12.45)) lze, jak známo, provést metodou Lagrangeových multiplikátorů. S ohledem na charakter úlohy, kdy můžeme připustit i přibližné řešení, lze postup zjednodušit. Zavedme označení

$$G(u', v') = \iint_{\Omega} (f_u u' + f_v v' + f_t)^2 \, dudv. \quad (12.47)$$

Nyní minimalizujeme funkcionál $F(u', v') + \lambda G(u', v')$, kde λ je zvolená reálná kladná konstanta. V tomto případě již ovšem není zajištěno, že rovnice (12.45) bude splněna přesně. Velikost konstanty λ určuje váhu, jakou má požadavek na splnění rovnice (12.45). Velké hodnoty λ způsobují, že je dobře splněna rovnice (12.45), ale hůře požadavek (12.46). U malých hodnot λ je tomu naopak. Konečně ještě poznamenejme, že ať již řešíme problém minimalizace funkcionálu jakoukoli metodou, je nutné zadat okrajové podmínky pro u', v' (tedy hodnoty těchto funkcí na okrajích obrazu). Okrajové podmínky bývají zpravidla opět zadávány poněkud spekulativně (např. $u'(u, v) = v'(u, v) = 0$).

Na závěr podkapitoly o optickém toku ještě ukážeme několik příkladů rychlostního pole jednoduchých objektů. Předpokládejme nejprve izolovaný bod X pohybující se ve scéně. Poloha bodu ve scéně je popsána vektorem $(x(t), y(t), z(t))$. Scénu pozorujeme kamerou, jejíž střed promítání je umístěn do počátku souřadného systému a jejíž průmětna má rovnici $z = -f$, kde f je ohnisková vzdálenost (obr. 12.3a). Pro rychlost pohybu obrazu bodu X z podobnosti trojúhelníků snadno dostaneme

$$u'(t) = \frac{d}{dt} \left[-f \frac{x(t)}{z(t)} \right], \quad v'(t) = \frac{d}{dt} \left[-f \frac{y(t)}{z(t)} \right]. \quad (12.48)$$

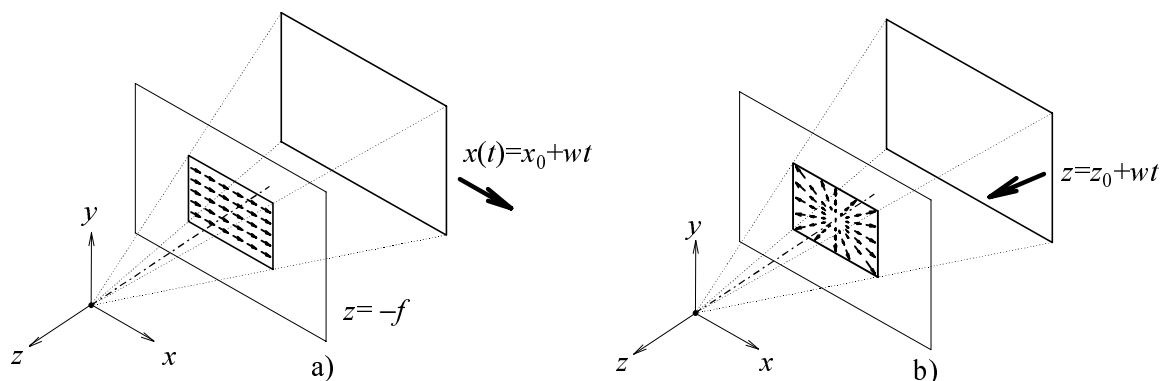
Nyní uvažujme rovinnou plochu, která se pohybuje rovnoběžně s průmětnou (obr. 12.3a). Necht' X je bodem této plochy. V čase $t = 0$ je poloha bodu X popsána vektorem (x_0, y_0, z_0) . V čase t je poloha tohoto bodu $x(t) = x_0 + wt$, $y(t) = y_0$, $z(t) = z_0$. Pro rychlost pohybu obrazu bodu X máme

$$u'(t) = \frac{d}{dt} \left[-f \frac{x_0 + wt}{z_0} \right] = -\frac{fw}{z_0}, \quad v'(t) = \frac{d}{dt} \left[-f \frac{y_0}{z_0} \right] = 0. \quad (12.49)$$

Rychlostní pole odpovídající popsanému případu je znázorněno na obr. 12.3a. Podobně dále uvažujme případ, kdy je pohyb popsán rovnicemi $x(t) = x_0$, $y(t) = y_0$, $z(t) = z_0 + wt$ ($w > 0$, jedná se tedy o pohyb roviny směrem ke kameře). V tomto případě obdržíme

$$u'(t) = \frac{d}{dt} \left[-f \frac{x_0}{z_0 + wt} \right] = \frac{fx_0 w}{(z_0 + wt)^2}, \quad v'(t) = \frac{d}{dt} \left[-f \frac{y_0}{z_0 + wt} \right] = \frac{fy_0 w}{(z_0 + wt)^2}. \quad (12.50)$$

Odpovídající rychlostní pole je znázorněno na obr. 12.3b.



Obr. 12.3. Rychlostní pole vytvořené pohybující se rovinou.

Dodatek A: Základní pojmy a vztahy počtu pravděpodobnosti

Pro zavedení potřebné terminologie, symboliky a pro pohodlí čtenáře uvádíme v tomto dodatku přehled základních vztahů z počtu pravděpodobnosti, kterých využíváme v tomto textu.

A.1 Náhodná proměnná

Výsledkem statistického experimentu je nějaký elementární jev ω . Množina všech elementárních jevů, které mohou při experimentu nastat, je $\Omega = \{\omega_1, \omega_2, \omega_3, \dots\}$. Ke každému elementárnímu jevu $\omega_i \in \Omega$ můžeme přiřadit reálné číslo (pravděpodobnost) p_i tak, že $p_i \geq 0$ a $\sum p_i = 1$. Necht' A je podmnožinou množiny Ω . Množinu A nazýváme událostí. Také samotná množina Ω je událostí (jistá událost). Prázdná množina je rovněž událostí. Jestliže A_i, A_j jsou události, pak také $\neg A_i, A_i \cap A_j, A_i \cup A_j$ jsou události. Dvě události A_i, A_j se vzájemně vylučují právě když $A_i \cap A_j = \emptyset$. Přiřaďme události A pravděpodobnost $\mathcal{P}(A)$ podle pravidel

$$\mathcal{P}(A) \geq 0, \quad \mathcal{P}(\Omega) = 1, \quad \text{je-li } A_i \cap A_j = \emptyset, \text{ pak } \mathcal{P}(A_i \cup A_j) = \mathcal{P}(A_i) + \mathcal{P}(A_j). \quad (\text{A.1})$$

Je zřejmé, že pravděpodobnost $\mathcal{P}(A)$ je rovna součtu pravděpodobností těch elementárních jevů, které jsou v A obsaženy. Uvažujme nyní funkci $\mathbf{f}(\omega_i)$, která ke každému elementárnímu jevu přiřadí reálné číslo. Funkce $\mathbf{f}(\omega_i)$ se nazývá náhodná proměnná. Zápis $\{\mathbf{f} \leq z\}$ označuje událost $\{\omega_i \mid \mathbf{f}(\omega_i) \leq z\}$. Podobně označují $\{z_1 < \mathbf{f} \leq z_2\}, \{\mathbf{f} = z\}$ události $\{\omega_i \mid z_1 < \mathbf{f}(\omega_i) \leq z_2\}$ nebo $\{\omega_i \mid \mathbf{f}(\omega_i) = z\}$. Pravděpodobnost $\mathcal{P}\{\mathbf{f} \leq z\}$ události $\{\mathbf{f} \leq z\}$ je funkcí z . Tuto funkci nazveme distribuční funkcí náhodné proměnné \mathbf{f} a označíme $P_{\mathbf{f}}(z)$. Je tedy

$$P_{\mathbf{f}}(z) = \mathcal{P}\{\mathbf{f} \leq z\}. \quad (\text{A.2})$$

Distribuční funkce má následující vlastnosti

$$P_{\mathbf{f}}(-\infty) = \mathcal{P}\{\mathbf{f} \leq -\infty\} = 0, \quad P_{\mathbf{f}}(\infty) = \mathcal{P}\{\mathbf{f} \leq +\infty\} = 1, \quad (\text{A.3a})$$

$$P_{\mathbf{f}}(z_1) \leq P_{\mathbf{f}}(z_2) \Leftrightarrow z_1 \leq z_2, \quad (\text{A.3b})$$

$$\mathcal{P}\{z_1 < \mathbf{f} \leq z_2\} = P_{\mathbf{f}}(z_2) - P_{\mathbf{f}}(z_1). \quad (\text{A.3c})$$

Derivace $p_{\mathbf{f}}(z)$ distribuční funkce $P_{\mathbf{f}}(z)$ se nazývá hustota pravděpodobnosti náhodné proměnné \mathbf{f} . Máme tedy

$$p_{\mathbf{f}}(z) = \frac{dP_{\mathbf{f}}(z)}{dz}, \quad P_{\mathbf{f}}(z) = \int_{-\infty}^z p_{\mathbf{f}}(z) dz. \quad (\text{A.4,5})$$

Z vlastností (A.3c) distribuční funkce a z rovnice (A.5) máme

$$\mathcal{P}\{z_1 < \mathbf{f} \leq z_2\} = \int_{z_1}^{z_2} p_{\mathbf{f}}(z) dz. \quad (\text{A.6})$$

Pro dostatečně malé Δz lze psát $\mathcal{P}\{z < \mathbf{f} \leq z + \Delta z\} \approx p_{\mathbf{f}}(z) \Delta z$. (A.7)

Odtud máme
$$p_{\mathbf{f}}(z) = \lim_{\Delta z \rightarrow 0} \frac{\mathcal{P}\{z < \mathbf{f} \leq z + \Delta z\}}{\Delta z}. \quad (\text{A.8})$$

Rovnice naznačuje možný postup při praktickém stanovení hustoty pravděpodobnosti: Provedme N (velký počet) experimentů. N_z necht' je počet experimentů, kdy náhodná proměnná padla do intervalu $(z, z + \Delta z)$. Hodnota $p_{\mathbf{f}}(z)$ v bodě z je pak přibližně dána podílem $N_z / (N \Delta z)$. Střední (očekávaná) hodnota $E\{\mathbf{f}\}$ náhodné proměnné \mathbf{f} je definována vztahem

$$E\{\mathbf{f}\} = \int_{-\infty}^{\infty} z p_{\mathbf{f}}(z) dz. \quad (\text{A.9})$$

Význam uvedeného vztahu je lépe patrný, aproximujeme-li integrál sumací

$$E\{\mathbf{f}\} = \sum_i z_i p_f(z_i) \Delta z = \sum_i z_i \mathcal{P}\{z_i < \mathbf{f} \leq z_i + \Delta z\}. \quad (\text{A.10})$$

Nechť \mathbf{g} je náhodná proměnná, která je funkcí náhodné proměnné \mathbf{f} . Je tedy $\mathbf{g} = \mathcal{L}(\mathbf{f})$. Pak platí

$$E\{\mathbf{g}\} = E\{\mathcal{L}(\mathbf{f})\} = \int_{-\infty}^{\infty} \mathcal{L}(z) p_f(z) dz. \quad (\text{A.11})$$

Nechť μ_f označuje střední hodnotu náhodné proměnné \mathbf{f} . Pak variance (rozptyl) σ_f^2 náhodné proměnné \mathbf{f} je definován vztahem

$$\sigma_f^2 = E\{(\mathbf{f} - \mu_f)^2\} = \int_{-\infty}^{\infty} (z - \mu_f)^2 p_f(z) dz. \quad (\text{A.12})$$

Odmocnina z variance se nazývá směrodatná odchylka. Náhodná proměnná \mathbf{f} má normální rozložení, jestliže její hustota $p_f(z)$ pravděpodobnosti je Gaussova křivka. Máme proto

$$p_f(z) = \frac{1}{\sigma_f \sqrt{2\pi}} \exp\left[-\frac{(z - \mu_f)^2}{2\sigma_f^2}\right]. \quad (\text{A.13})$$

A.2 Podmíněná pravděpodobnost

Uvažujme dvě události A, B . Jak bylo uvedeno dříve, je $A \cap B$ také událost. Podmíněná pravděpodobnost $\mathcal{P}\{A | B\}$ události A za předpokladu, že nastala událost B , je definována vztahem

$$\mathcal{P}(A | B) = \mathcal{P}(A \cap B) / \mathcal{P}(B). \quad (\text{A.14})$$

Distribuční funkce $P_f(z | B)$ podmíněné pravděpodobnosti náhodné proměnné \mathbf{f} za předpokladu, že nastala událost B , je definována jako podmíněná pravděpodobnost události $\{\mathbf{f} \leq z\}$. Je tedy

$$P_f(z | B) = \mathcal{P}\{\mathbf{f} \leq z | B\} = \mathcal{P}(\{\mathbf{f} \leq z\} \cap B) / \mathcal{P}(B). \quad (\text{A.15})$$

(Čitatel tedy zahrnuje všechny elementární události ω_i , pro které platí $\mathbf{f}(\omega_i) \leq z$ a současně $\omega_i \in B$). Podmíněná hustota pravděpodobnosti je definována vztahem

$$p_f(z | B) = \frac{dP_f(z | B)}{dz}. \quad (\text{A.16})$$

Podmíněná střední hodnota $E\{\mathbf{f} | B\}$ je pak ve shodě s výrazem (A.9) dána předpisem

$$E\{\mathbf{f} | B\} = \int_{-\infty}^{\infty} z p_f(z | B) dz. \quad (\text{A.17})$$

Pro pravděpodobnost $\mathcal{P}(A \cap B)$ součinu událostí A, B platí

$$\mathcal{P}(A \cap B) = \mathcal{P}(A | B) \mathcal{P}(B) = \mathcal{P}(B | A) \mathcal{P}(A). \quad (\text{A.18})$$

Vztah (A.18) bývá nazýván Bayesovým vztahem. Jestliže jsou události A, B nezávislé, pak $\mathcal{P}(A | B) = \mathcal{P}(A)$, $\mathcal{P}(B | A) = \mathcal{P}(B)$ a vztah (A.18) tak přechází na tvar

$$\mathcal{P}(A \cap B) = \mathcal{P}(A) \mathcal{P}(B). \quad (\text{A.19})$$

A.3 Sdružená pravděpodobnost

Nechť $\Omega = \{\omega_1, \omega_2, \dots\}$ je opět množina všech možných elementárních jevů. Na této množině definujme nyní n funkcí $\mathbf{f}_1(\omega_i), \mathbf{f}_2(\omega_i), \dots, \mathbf{f}_n(\omega_i)$. Každá z těchto funkcí zobrazuje množinu Ω elementárních jevů do množiny reálných čísel, a každá je proto náhodnou proměnnou. Podle podkapitoly A.1 je každá z množin $\{\mathbf{f}_1 \leq z_1\}, \{\mathbf{f}_2 \leq z_2\}, \dots, \{\mathbf{f}_n \leq z_n\}$ událostí. Uvažujme událost, která vznikne součinem

$$\{\mathbf{f}_1 \leq z_1\} \cap \{\mathbf{f}_2 \leq z_2\} \cap \dots \cap \{\mathbf{f}_n \leq z_n\} \equiv \{\mathbf{f}_1 \leq z_1, \mathbf{f}_2 \leq z_2, \dots, \mathbf{f}_n \leq z_n\}.$$

Tato událost obsahuje ty elementární jevy ω_i , pro které platí $\mathbf{f}_1(\omega_i) \leq z_1, \mathbf{f}_2(\omega_i) \leq z_2, \dots, \mathbf{f}_n(\omega_i) \leq z_n$. Pravděpodobnost této události (v závislosti na z_1, z_2, \dots, z_n) nazveme sdruženou distribuční funkcí. Je tedy

$$P_{\mathbf{f}_1, \dots, \mathbf{f}_n}(z_1, z_2, \dots, z_n) = \mathcal{P}\{\mathbf{f}_1 \leq z_1, \mathbf{f}_2 \leq z_2, \dots, \mathbf{f}_n \leq z_n\}. \quad (\text{A.20})$$

Sdružená hustota pravděpodobnosti je opět derivací distribuční funkce:

$$p_{\mathbf{f}_1, \dots, \mathbf{f}_n}(z_1, z_2, \dots, z_n) = \frac{\partial^n P_{\mathbf{f}_1, \dots, \mathbf{f}_n}(z_1, z_2, \dots, z_n)}{\partial z_1 \partial z_2 \dots \partial z_n}. \quad (\text{A.21})$$

Někdy chceme vyšetřovat situace, kdy na hodnotách některých náhodných proměnných nezáleží (řekněme, že se jedná o proměnné $\mathbf{f}_{i+1}, \dots, \mathbf{f}_n$, kde $i < n$). V takovém případě můžeme událost $\{\mathbf{f}_1 \leq z_1\} \cap \{\mathbf{f}_2 \leq z_2\} \cap \dots \cap \{\mathbf{f}_i \leq z_i\}$ zapsat ve tvaru $\{\mathbf{f}_1 \leq z_1\} \cap \{\mathbf{f}_2 \leq z_2\} \cap \dots \cap \{\mathbf{f}_i \leq z_i\} \cap \{\mathbf{f}_{i+1} \leq \infty\} \cap \dots \cap \{\mathbf{f}_n \leq \infty\}$. Proto platí

$$P_{\mathbf{f}_1, \dots, \mathbf{f}_i}(z_1, z_2, \dots, z_i) = P_{\mathbf{f}_1, \dots, \mathbf{f}_n}(z_1, z_2, \dots, z_i, \infty, \dots, \infty). \quad (\text{A.22})$$

Analogicky ke vztahu (A.5) pro jednu náhodnou proměnnou platí i zde

$$P_{\mathbf{f}_1, \dots, \mathbf{f}_n}(z_1, z_2, \dots, z_n) = \int_{-\infty}^{z_1} \int_{-\infty}^{z_2} \dots \int_{-\infty}^{z_n} p_{\mathbf{f}_1, \dots, \mathbf{f}_n}(\xi_1, \xi_2, \dots, \xi_n) d\xi_1 d\xi_2 \dots d\xi_n. \quad (\text{A.23})$$

Dále platí

$$p_{\mathbf{f}_1, \dots, \mathbf{f}_i}(z_1, z_2, \dots, z_i) = \frac{\partial^i P_{\mathbf{f}_1, \dots, \mathbf{f}_i}(z_1, z_2, \dots, z_i)}{\partial z_1 \partial z_2 \dots \partial z_i} \quad (\text{A.24})$$

a

$$p_{\mathbf{f}_1, \dots, \mathbf{f}_i}(z_1, z_2, \dots, z_i) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p_{\mathbf{f}_1, \dots, \mathbf{f}_n}(z_1, z_2, \dots, z_n) dz_{i+1} dz_{i+2} \dots dz_n. \quad (\text{A.25})$$

Nechť náhodná proměnná \mathbf{g} je funkcí náhodných proměnných $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n$, tedy $\mathbf{g} = \mathcal{L}\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n\}$, Pak střední hodnota náhodné proměnné \mathbf{g} je dána výrazem

$$E\{\mathbf{g}\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \mathcal{L}(z_1, z_2, \dots, z_n) p_{\mathbf{f}_1, \dots, \mathbf{f}_n}(z_1, z_2, \dots, z_n) dz_1 dz_2 \dots dz_n. \quad (\text{A.26})$$

Náhodné proměnné $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n$ nazveme nezávislými, když platí

$$p_{\mathbf{f}_1, \dots, \mathbf{f}_n}(z_1, z_2, \dots, z_n) = p_{\mathbf{f}_1}(z_1) p_{\mathbf{f}_2}(z_2) \dots p_{\mathbf{f}_n}(z_n). \quad (\text{A.27})$$

Náhodné proměnné $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n$ nazveme nekorelovanými, když pro každé $i \neq j$ platí

$$E\{\mathbf{f}_i \mathbf{f}_j\} = E\{\mathbf{f}_i\} E\{\mathbf{f}_j\}. \quad (\text{A.28})$$

Kovarianci $C_{\mathbf{f}_i, \mathbf{f}_j}$ dvou náhodných proměnných $\mathbf{f}_i, \mathbf{f}_j$ definujeme vztahem

$$C_{\mathbf{f}_i, \mathbf{f}_j} = E\{(\mathbf{f}_i - \mu_{\mathbf{f}_i})(\mathbf{f}_j - \mu_{\mathbf{f}_j})\}. \quad (\text{A.29})$$

Roznásobením členů v závorkách a uplatněním střední hodnoty na každý ze sčítanců lze dokázat, že je

$$C_{\mathbf{f}_i, \mathbf{f}_j} = E\{\mathbf{f}_i \mathbf{f}_j\} - \mu_{\mathbf{f}_i} \mu_{\mathbf{f}_j}. \quad (\text{A.30})$$

Odtud vidíme, že náhodné proměnné $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n$ jsou nekorelované právě tehdy, jestliže kovariance každých dvou různých z nich se rovná nule.

Dodatek B: Řešení předeterminovaných systémů lineárních rovnic

Mějme systém m rovnic o n neznámých. Jestliže platí $m > n$, pak je systém předeterminovaný. Předeterminované systémy bývají v praxi často nekonzistentní. K nekonzistenci dojde např. tehdy, jestliže jednotlivé rovnice byly sestaveny na základě měření, která byla zatížena nepřesnostmi. Přísně vzato si rovnice v nekonzistentním systému odporují a systém by tak teoreticky neměl mít žádné řešení. Prak-

tický může být ale zajímavé najít hodnoty neznámých, které v jistém smyslu všem rovnicím systému vyhovují co nejlépe. Tyto hodnoty budeme nazývat řešením předeterminovaného nekonzistentního systému. Uvažujme systém tvaru

$$\mathbf{Ax} = \mathbf{b}, \quad (\text{B.1})$$

kde \mathbf{x} je vektor hledaných hodnot, \mathbf{b} je vektor pravých stran a \mathbf{A} je matice soustavy. Rozměry vektorů \mathbf{x}, \mathbf{b} a matice \mathbf{A} jsou postupně $n, m, m \times n$. Reziduum (míru nesplnění jednotlivých rovnic) soustavy (B.1) můžeme vyjádřit vztahem

$$\mathbf{r} = \mathbf{Ax} - \mathbf{b}. \quad (\text{B.2})$$

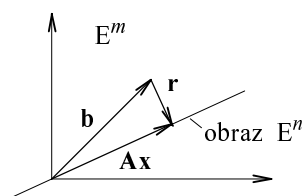
Řešení \mathbf{x} soustavy (B.1) můžeme najít tak, aby minimalizovalo délku vektoru \mathbf{r} . Hledáme tedy

$$\min_{\mathbf{x}} [\mathbf{r}^T \mathbf{r}] = \min_{\mathbf{x}} [(\mathbf{Ax} - \mathbf{b})^T (\mathbf{Ax} - \mathbf{b})]. \quad (\text{B.3})$$

Ukážeme, že řešení získané na základě vztahu (B.3) je ekvivalentní s řešením získaným tzv. zobecněnou inverzí, které lze získat jednoduchou úpravou vztahu (B.1). Dostaneme

$$\mathbf{x}^* = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b}. \quad (\text{B.4})$$

Je tedy \mathbf{x}^* řešením, které ve smyslu vztahu (B.3) nejlépe vyhovuje dané soustavě. Zdůvodnění uvedené tvrzení je následující: Všechna možná rezidua vyplní m -rozměrný euklidovský prostor E^m ,



Obr. B.1. K řešení předeterminovaných soustav.

všechna řešení vyplní n -rozměrný euklidovský prostor E^n . Součin \mathbf{Ax} zobrazuje prostor E^n do prostoru E^m . Obecně obrazy všech prvků z E^n vytvoří v E^m n -rozměrný podprostor prostoru E^m . Uvažujme například následující systém dvou rovnic o jedné neznámé: $2x=1, 3x=2$. V tomto případě má matice \mathbf{A} tvar $\mathbf{A}=(2,3)^T$. Jednorozměrný prostor E^1 je prostor všech možných hodnot x . Tento prostor se zobrazuje do dvojrozměrného prostoru E^2 jako přímka $(2x,3x)$ o směrnici $3/2$. Délka vektoru \mathbf{r} z rovnice (B.2) je vzdáleností mezi body, které jsou v E^m reprezentovány vektory \mathbf{Ax}, \mathbf{b} (obr. B.1). Reziduum \mathbf{r} bude nejmenší, jestliže \mathbf{x} nabude

takové hodnoty \mathbf{x}^* , při níž je vektor \mathbf{r} kolmý k vektoru \mathbf{Ax} . Podmínku kolmosti vektorů $\mathbf{Ax}^*, \mathbf{r}$ můžeme vyjádřit pomocí skalárního součinu ve tvaru

$$(\mathbf{Ax}^*)^T (\mathbf{Ax}^* - \mathbf{b}) = 0. \quad (\text{B.5})$$

Úpravou dostaneme

$$(\mathbf{x}^*)^T (\mathbf{A}^T \mathbf{Ax}^* - \mathbf{A}^T \mathbf{b}) = 0. \quad (\text{B.6})$$

Rovnice (B.6) bude splněna, když výraz ve druhé závorce na levé straně rovnice bude nulovým vektorem (varianta, kdy by nulovým vektorem byl samotný vektor \mathbf{x}^* , není prakticky zajímavá). Řešením rovnice $\mathbf{A}^T \mathbf{Ax}^* - \mathbf{A}^T \mathbf{b} = \mathbf{0}$ snadno získáváme hledaný vztah (B.4).

Dodatek C: Hledání minima s podmínkou

Úlohou je najít vektor $\mathbf{x}=(x_1, x_2, \dots, x_n)$ hodnot, v nichž daná funkce $f(\mathbf{x})$ nabývá lokálního minima. Při tom mají být dále splněny podmínky $g_1(\mathbf{x})=0, g_2(\mathbf{x})=0, \dots, g_p(\mathbf{x})=0$, kde $g_i(\mathbf{x})$ jsou známé funkce. Úlohu lze řešit metodou Lagrangeových multiplikátorů. Metoda se opírá o tvrzení, že hledané řešení minimalizuje funkci

$$F(\mathbf{x}, \lambda_1, \lambda_2, \dots, \lambda_p) = f(\mathbf{x}) + \sum_{i=1}^p \lambda_i g_i(\mathbf{x}). \quad (\text{C.1})$$

Původní problém s doplňujícími podmínkami je tak převeden na problém bez podmínek. Nyní je třeba najít hodnoty $\mathbf{x}, \lambda_1, \lambda_2, \dots, \lambda_p$, pro které funkce F nabývá minima. Minimum lze nalézt uplatněním obvyklého postupu tak, že položíme $\partial F / \partial \mathbf{x} = \mathbf{0}, \partial F / \partial \lambda_i = 0$. Dostaneme

$$\frac{\partial F(\mathbf{x}, \lambda_1, \lambda_2, \dots, \lambda_p)}{\partial \mathbf{x}} = \frac{f(\mathbf{x})}{\partial \mathbf{x}} + \sum_{i=1}^p \lambda_i \frac{\partial g_i(\mathbf{x})}{\partial \mathbf{x}} = \mathbf{0}, \quad (\text{C.2})$$

$$\frac{\partial F(\mathbf{x}, \lambda_1, \lambda_2, \dots, \lambda_p)}{\partial \lambda_i} = g_i(\mathbf{x}) = 0. \quad (\text{C.3})$$

Vztah (C.2) dává n rovnic, vztah (C.3) dává p rovnic. Z uvedené soustavy můžeme řešit $n+p$ neznámých $x_1, x_2, \dots, x_n, \lambda_1, \lambda_2, \dots, \lambda_p$. Komplikace nastanou, jestliže soustava (C.2), (C.3) vyjde nelineární, což je v praxi bohužel dosti časté. Pro tento případ je v dodatku D popsána metoda, kterou lze pro minimalizaci funkce F snadno použít, aniž by na její derivace byl kladen požadavek linearity. Rovnice (C.3) také ukazuje, že minimalizace nově zavedené funkce F skutečně zajistí splnění požadovaných doplňujících podmínek.

Dodatek D: Hledání minima gradientní metodou

Úlohou je najít vektor $\mathbf{x}=(x_1, x_2, \dots, x_n)$, v němž daná reálná funkce $f(\mathbf{x})$ nabývá minima. Gradientní metoda je metodou iterační. Její princip vychází z jednoduché představy: Během iterací počítáme posloupnost $\mathbf{x}^{(i)}$ hodnot vektoru \mathbf{x} tak, že vždy postupujeme proti směru gradientu funkce $f(\mathbf{x})$ o nějaký jeho přiměřený násobek tak dlouho, dokud se délka gradientu dostatečně nepřiblíží nule (vždy tedy postupujeme ve směru, v němž hodnota funkce nejvíce klesá). Nechť $\mathbf{x}^{(i)}$ označuje hodnotu vypočítanou v i -tém kroku. Hodnotu $\mathbf{x}^{(i+1)}$ v $(i+1)$ -vém kroku počítáme podle předpisu

$$\mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} - k \operatorname{grad}[f(\mathbf{x})]_{\mathbf{x}=\mathbf{x}^{(i)}}, \quad (\text{D.1})$$

kde k je reálná dostatečně malá hodnota (tak malá, abychom minimum nevhodně velkým krokem „nepřeskočili“). Výhodou gradientní metody je, že ji lze jednoduše aplikovat zejména v nelineárních úlohách, v nichž by uplatnění obvyklé podmínky $\partial f/\partial \mathbf{x}=\mathbf{0}$ vedlo na soustavu nelineárních rovnic. Gradientní metodu lze obvykle velmi snadno implementovat. Nevýhody jsou obdobné jako u ostatních iteračních metod: Je nutné řešit otázky konvergence a otázku vhodné volby výchozí hodnoty \mathbf{x} . Z principu metody je též zřejmé, že gradientní metoda hledá minimum lokální, ačkoli při řešení praktických úloh bychom často požadovali minimum globální (odtud opět vyplývá potřeba dostatečně dobrého odhadu počáteční hodnoty \mathbf{x}). Gradientní metoda je známá svojí pomalou konvergencí. Rychlost konvergence navíc klesá s tím, jak se délka gradientu blíží k nule, když se \mathbf{x} přibližuje k hledanému řešení. I přes uvedené nedostatky však může být gradientní metoda dobrou metodou pro první experimentování s řešeným problémem.

Literatura odkazovaná z textu

- M.F. Barnsley, L.P. Hurd, *Fractal Image Compression*, AK Peters, Wellesley, MA, 1993.
- B. Budínský, *Analytická a diferenciální geometrie*, SNTL Praha, 1983.
- J.F. Canny, Finding edges and lines in images, *Technical report AI-TR-720*, MIT Artificial Intelligence Laboratory, 1983.
- J.F. Canny, A computational approach to edge detection, *IEEE PAMI* **8** (1986), 679—698.
- K.R. Castleman, *Digital Image Processing*, Prentice-Hall, 1996 (ISBN 0-13-211467-4).
- CCITT, *Recommendation T.81*, Information technology - digital compression and coding of continuous-tone still images - requirements and guidelines, 1992.
- C.K. Chui, *An Introduction to Wavelets*, Academic Press, San Diego, 1992.
- R. Deriche, O.D. Faugeras, Tracking line segments, *Proc. First ECCV*, Antibes, 1990, Springer-Verlag, Berlin, 259—269.
- R.E. Dougherty, *An Introduction to Morphological Image Processing*, SPIE Press, Bellingham, WA, 1992.
- R.O. Duda, P.E. Hart, Use of Hough transformation to detect lines and curves in pictures. *Comm. ACM* **15** (1972), 11—15.
- D.E. Dudgeon, R.M. Mersereau, *Multidimensional Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1984.
- C.G. Harris, M.J. Stephens, A combined corner and edge detector, *Proc. 4th Alvey Vision Conference*, 1988, 147—152.
- L. Kitchen, A. Rosenfeld, Gray-level corner detection, *Pattern Recognition Letters* **1** (1982), 95—102.
- D. Maar, E. Hildreth, Theory of edge detection, *Proc. Roy. Soc. London B* **207**, 1980, 187—217.
- N. Otsu, Discriminant and least square threshold selection, *Proc. 4th IJCP*, Tokyo, 1978, 592—596.
- M.A. Penna, R.R. Patterson, *Projective Geometry and its Applications to Computer Graphics*, Prentice-Hall, Englewood Cliffs, NJ, 1986 (ISBN 0-13-730649-0).
- W.K. Pratt, *Digital Image Processing*, John Wiley & Sons, New York, Toronto, 1978 (ISBN 0-471-01888-0).
- J. Serra, *Image Analysis and Mathematical Morphology, Vol.1*, Academic Press, New York, 1982.
- J. Serra (ed.), *Image Analysis and Mathematical Morphology, Vol.2*, Academic Press, New York, 1988.

Další doporučená literatura

- Y. Anzai, *Pattern Recognition and Machine Learning*, Academic Press, 1992 (ISBN 0-12-058830-7).
- E.R. Dougherty, *Digital Image Processing*, Marcel Dekker, 1994 (ISBN 0-8247-8927-X).
- O. Faugeras, *Three-Dimensional Computer Vision*, MIT Press, 1994 (ISBN 0-262-06158-9).
- L.J. Galbiati, *Machine Vision and Digital Image Processing Fundamentals*, Prentice Hall, 1990 (ISBN 0-13-542044-X).
- G.H. Granlund, H. Knutsson, *Signal Processing for Computer Vision*, Kluwer Academic Publishers, 1994 (ISBN 0-7923-9530-1).

- B. Jahne, *Digital Image Processing: Concepts, Algorithms, and Scientific Applications* (4th ed.), Springer-Verlag, Berlin, Heidelberg, 1997 (ISBN 3-540-62724-3).
- A.K. Jain, *Fundamentals of Digital Image Processing*, Prentice Hall, 1990 (ISBN 0-13-332578-4).
- J.R. Jensen, *Introductory Digital Image Processing: A Remote Sensing Perspective* (2nd ed.), Prentice Hall, 1995 (ISBN 0-13-205840-5).
- R. Johnsonbaugh, E. Gose, S. Jost, *Pattern Recognition with Image Analysis*, Prentice Hall, 1996 (ISBN 0-13-236415-8).
- R. Klette, K. Schluns, A. Koschan, *Computer Vision: Three-Dimensional Data from Images*, Springer-Verlag, Singapore, 1998 (ISBN 981-3083-71-9).
- W. Kou, *Digital Image Compression: Algorithms and Standards*, Kluwer Academic Publishers, 1995 (ISBN 0-7923-9626-X).
- C.T. Leondes (ed.), *Image Processing and Pattern Recognition*, AP Professional, 1998 (ISBN 0-12-443860-1).
- C.G. Looney, *Pattern Recognition Using Neural Networks*, Oxford University Press, 1997 (ISBN 0-19-507920-5).
- G. McLachlan, *Discriminant Analysis and Statistical Pattern Recognition*, John Wiley & Sons, 1992 (ISBN 0-471-61531-5).
- J.L. Mitchell, W.B. Pennebaker, *MPEG Digital Video Compression Standard*, Kluwer Academic Publishers, 1996 (ISBN 0-412-08771-5).
- M. Nadler, E. Smith, R.W. Connors, *Pattern Recognition Engineering*, John Wiley & Sons, 1993 (ISBN 0-471-62293-1).
- A. Nigrin, *Neural Networks for Pattern Recognition*, MIT Press, 1993 (ISBN 0-262-14054-3).
- J.R. Parker, *Algorithms for Image Processing and Computer Vision*, John Wiley & Sons, 1996 (ISBN 0-471-14056-2).
- M. Pavel, *Fundamentals of Pattern Recognition* (2nd ed.), Marcel Dekker, 1993 (ISBN 0-8247-8883-4).
- W.B. Pennebaker, J.L. Mitchell, *JPEG: Still Image Compression Standard*, Kluwer Academic Publishers, 1992 (ISBN 0-442-01272-1).
- I. Pitas, *Digital Image Processing Algorithms*, Prentice Hall, 1993 (ISBN 0-13-145814-0).
- W.K. Pratt, *Digital Image Processing* (2nd ed.), John Wiley & Sons, 1991 (ISBN 0-471-85766-1).
- J.A. Richards, *Remote Sensing Digital Image Analysis: An Introduction*, Springer-Verlag, Berlin, Heidelberg, 1994 (ISBN 3-540-58219-3).
- B.D. Ripley, *Pattern recognition and Neural Network*, Cambridge University Press, 1996 (ISBN 0-521-46086-7).
- R. Schalkoff, *Digital Image Processing and Computer Vision*, John Wiley & Sons, 1989 (ISBN 0-471-85718-1, 0-471-50536-6).
- R.J. Schalkoff, *Pattern Recognition: Statistical, Structural and Neural Approaches*, John Wiley & Sons, 1991 (ISBN 0-471-52974-5).
- A.M. Tekalp, *Digital Video Image Processing*, Prentice Hall, 1995 (ISBN 0-13-190075-7).
- S. Theodoridis, K. Koutroumbas, *Pattern Recognition*, Academic Press, 1998 (ISBN 0-12-686140-4).
- C. Torras, *Computer Vision: Theory and Industrial Applications*, Springer-Verlag, Berlin, Heidelberg, 1992 (ISBN 3-540-52036-8).