

## **Metody strojového vnímání**

Základy zpracování signálů, obrazů a přirozené řeči. Počítačové vidění, vznik obrazu, analýza scény, stereovidění, systémy více kamer, analýza pohybu.

Vypracování státnicové otázky č. 7 okruhu IV nových otázek

# Obsah

<b>I</b>	<b>Zpracování přirozené řeči</b>	<b>1</b>
<b>1</b>	<b>Analýza akustického řečového signálu</b>	<b>1</b>
1.1	Model vytváření řeči . . . . .	1
1.2	Akustická analýza řeči . . . . .	2
1.3	Fonetická analýza . . . . .	2
<b>2</b>	<b>Automatické rozpoznávání řeči</b>	<b>2</b>
2.1	Metody rozpoznávání založené na porovnávání vzorů . . . . .	3
2.2	Statistický přístup k rozpoznávání řeči . . . . .	3
2.3	Znalostní přístup . . . . .	3
<b>II</b>	<b>Počítačové vidění</b>	<b>4</b>
<b>3</b>	<b>Úvod</b>	<b>4</b>
3.1	Dvě úrovně zpracování obrazu . . . . .	4
3.2	Základní kroky zpracování obrazu . . . . .	4
3.3	Proč je porozumění v úlohách 3D počítačového vidění těžké? . . . . .	5
<b>4</b>	<b>Vznik obrazu</b>	<b>5</b>
4.1	Formování obrazu . . . . .	5
4.2	Snímání obrazu . . . . .	6
4.3	Digitalizace obrazu . . . . .	6
<b>5</b>	<b>Zpracování digitálního obrazu na nízké úrovni</b>	<b>7</b>
5.1	Bodové jasové transformace . . . . .	7
5.2	Geometrické transformace . . . . .	8
5.3	Lokální předzpracování . . . . .	8
<b>6</b>	<b>Analýza 2D obrazu</b>	<b>10</b>
6.1	Segmentace . . . . .	10
6.2	Popis objektů . . . . .	11
<b>7</b>	<b>Analýza pohybu</b>	<b>12</b>
7.1	Rozdílové metody analýzy pohybu . . . . .	13
7.2	Optický tok . . . . .	13
7.3	Detekce významných bodů . . . . .	13
<b>8</b>	<b>Analýza 3D objektů</b>	<b>14</b>
8.1	Řízení obrazovými daty – řízení zdola nahoru. . . . .	14
8.2	Řízení podle vnitřního modelu – řízení shora dolů. . . . .	15
<b>9</b>	<b>Geometrie pro počítačové vidění</b>	<b>16</b>
9.1	Jedna kamera . . . . .	16
9.2	Systém více kamer . . . . .	16

## Část I

# Zpracování přirozené řeči

## 1 Analýza akustického řečového signálu

### 1.1 Model vytváření řeči

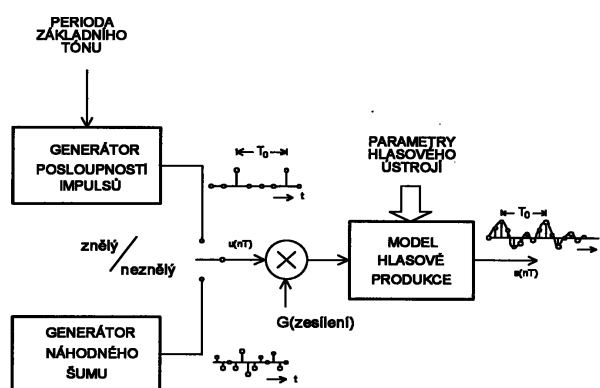
Zdrojem řečových kmitů jsou **řečové orgány**: hlasivky, dutina hrdelní, ústní a nosní, měkké a tvrdé patro, zuby a jazyk. Zdrojem hlasové energie jsou plíce a dýchací svaly. Zdrojem znělých zvuků jsou hlasivky vydávající **základní tón hlasu** (150 až 400 Hz).

**Samohlásky** vytváří volný průchod vzduchu hlasovým traktem. V akustickém spektru se objevují zesílené tóny vznikající rezonancí v dutinách hlasového traktu – **formanty**.

**Souhlásky** jsou vytvářeny turbulencí – třením vzduchu o překážku (jazyk, zuby, rty). Je-li překážka úplná jedná se o **závěrové** souhlásky (p, t, t', k, b, d, d', g, m, n, ň). Další typ je **úžínová** souhlásky, vznikající šumem způsobeným zúžením cesty vzduchu (f, v, s, š, z, ž, j, ch, h, l, r, ř). Oba typy překážek (postupně) se objevují u souhlásek **polozávěrových** (c, č). Souhlásky lze dále rozdělit podle znělosti (znělá–neznělá), **párové** (p-b, t-d, t'-d', k-g, s-z, š-ž, f-v, ch-h, c-dz, č-dž) a pouze znělé **nepárové** (m, n, ň, l, j, r, ř).

Střídáním poloh hlasového ústrojí vznikají akustické signály, z nichž se formují zvukové elementy – **fonémy** (v češtině odpovídají přibližně vysloveným hláskám). Fonémy se mění podle kontextu – závisí na předchozím a následujícím zvuku – **koartikulace**.

Cílem modelování produkce řeči je nalézt matematické vztahy použitelné pro reprezentaci fyzikálních dějů spojených s touto produkcí. Lidská řeč je nelineární časově proměnný proces. Lze ale udělat jisté zjednodušující předpoklady: pro mnohé zvuky platí, že typ buzení a vlastnosti hlasového traktu jsou téměř konstantní v intervalu 10–20 ms. Pak lze vytvořit model (obrázek 1) složený z lineárního modelu hlasového traktu s pomalu se měnícími parametry, který je buzen vhodným budícím signálem (periodický sled impulsů pro znělou, či šumový signál pro neznělou řeč).



Obrázek 1: Model vytváření řeči.

**Model hlasové produkce** na obrázku 1 respektuje chování hlasivek, hlasového traktu i proces vyzařování

zvuku ze rtů a nozder. Jeho z-přenos je ve tvaru celopólového filtru Q-tého řádu:

$$H(z) = \frac{1}{1 + \sum_{i=1}^Q a_i z^{-i}},$$

kde  $a_i$  jsou parametry modelu a  $Q$  v rozsahu 8–14.

## 1.2 Akustická analýza řeči

Předpoklad pomalých změn vlastností signálu vede na metody **krátkodobé analýzy**, kdy se signál rozdělí na samostatné **mikrosegmenty** délky 10 až 30 ms. Signál se vzorkuje a kvantuje, používá se **pulsní kódová modulace**.

Analýza v **časové oblasti** lze vyjádřit vztahem

$$Q_n = \sum \tau(s(k))w(n - k),$$

$Q_n$  je krátkodobá charakteristika,  $s(k)$  je vzorek akustického signálu v čase  $k$ ,  $\tau(\cdot)$  je transformační funkce a  $w(n)$  je váhová posloupnost – okénko, vybírající příslušné vzorky signálu.

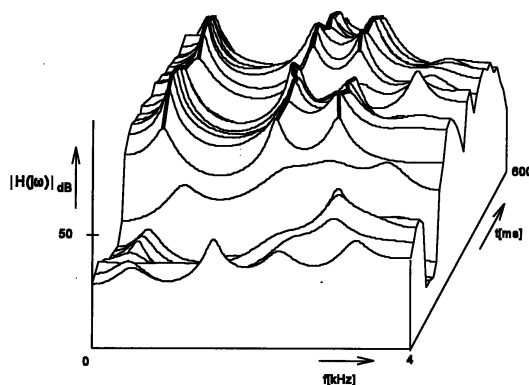
Analýza ve **frekvenční oblasti** používá krátkodobou Fourierovu transformaci signálu rovněž v okénku. Další metody analýzy jsou **kepstrální analýza** a **lineární prediktivní kódování**.

## 1.3 Fonetická analýza

Metody odhadu **fundamentální frekvence** – základní hlasivkový tón odpovídající kmitům hlasivek. Využívá se autokorelační funkce. Problém je odlišit základní frekvenci od formantů.

Informace o **formantových frekvencích** – ve spektrální obálce signálu. Používá se vyhlazené signálové spektrum. Příklad vývoje spektrální obálky je na obrázku 2.

Při dalším zpracování se často užívá **vektorová kvantizace** – aproximace analogové hodnoty popisující mikrosegment celočíselnou. **Kódová kniha** – reprodukční abeceda typových spektrálních vzorů, konstruuje se z trénovací množiny.



Obrázek 2: Časový průběh změn tvarů spektrálních obálek slova „čtyři“.

## 2 Automatické rozpoznávání řeči

Používají se metody založené na porovnávání vzorů, na statistickém přístupu a na znalostním přístupu.

## 2.1 Metody rozpoznávání založené na porovnávání vzorů

Zejména klasifikace izolovaných slov. Každé slovo ve slovníku má několik vzorů vyjádřených řetězci příznaků. Používá se porovnání se vzory, často kritérium minimální vzdálenosti nebo pravidlo nejbližšího souseda. Problém je porovnání obrazu vzorového a neznámého slova, protože mezi slovy jsou odlišnosti v časovém členění. K porovnávání se používá dynamické programování – nalezení funkce která nelineárně upraví časovou osu tak, aby si slova byla co nejvíce podobná (bortivá funkce, dynamic time warping).

## 2.2 Statistický přístup k rozpoznávání řeči

Používají se Markovské modely, využívá se vektorová kvantizace.

## 2.3 Znalostní přístup

Je třeba provést formalizaci znalostí o vytváření řeči (báze znalostí) a údajů o konkrétní řeči (báze dat). Používají se principy expertních systémů. Akusticko-fonetické znalosti vyjadřují vztah mezi fonetickým přepisem promluvy. Lexikální a fonologické znalosti představují v podstatě slovník. Prozodické znalosti slouží k lokalizaci přízvukných slabik a segmentaci promluvy do syntaktických frází. Syntaktické znalosti jsou reprezentovány vhodnou gramatikou, vyjadřují přípustnou konstrukci vět. Sémantické znalosti se využívají při oceňování smyslu nebo přípustnosti obsahu analyzovaných vět. Pragmatické znalosti reprezentují kontext.

Používá se několik modelů organizace a řízení znalostí:

**Hierarchický model.** Používá se režim zdola-nahoru (zpracování od akustické úrovně), režim shora-dolu (generování a ověřování hypotéz) a režim kombinovaný.

**Heterarchický model.** Umožňuje komunikaci mezi libovolnými dvěma zdroji znalostí. Nevýhoda – množství informačních spojení.

**Model tabule.** Zdroje dat komunikují pomocí společné datové struktury.

## Část II

# Počítačové vidění

## 3 Úvod

**Počítačové vidění** je disciplína, která se snaží technickými prostředky alespoň částečně napodobit lidské vidění – zdroj převážné většiny informací o okolním světě. Při vyhodnocení vizuální informace hraje obrovskou roli inteligence člověka, která umožňuje reprezentovat dlouho nabývané znalosti a zkušenosti o okolním světě. Výzkum počítačového vidění se snaží o řešení analogických úloh. Teoreticky i technicky jsou ale zvládnuty jen velmi jednoduché problémy. Pro počítačové vidění je typická snaha porozumět obecné trojrozměrné scéně.

Zpracování a rozpoznávání obrazu počítačem se liší od počítačové grafiky. Cílem grafiky je zobrazit informaci z počítače, počítačové vidění se naopak věnuje vstupu obrazové informace o skutečném světě a jejímu vyhodnocení.

### 3.1 Dvě úrovně zpracování obrazu

Ve zpracování obrazu se dají rozlišit dvě hlavní úrovně. Postupy počítačového vidění jsou značně složité, s těžištěm v interpretaci obrazových dat, která jsou nejčastěji reprezentována symbolicky. Jádrem pokročilejších postupů jsou znalostní systémy a techniky umělé inteligence. Této části počítačového vidění se říká **vyšší úroveň** a jde v ní o porozumění obrazu.

Předpokladem porozumění je však práce s obrazem na **nižší úrovni**. Cílem nižší úrovně počítačového vidění je analyzovat vstupní dvojrozměrná obrazová data číselného charakteru a najít kvalitativní symbolickou informaci potřebnou pro vyšší úroveň. Postupy nižší úrovně jsou také bezprostředně prakticky použitelné pro odstraňování šumu z obrazu, rozpoznávání jednoduchých objektů v obraze apod. Pro nižší úroveň se také používá název zpracování obrazu počítačem (image processing). Jednotlivé úrovně zpracování obrazu spolu však souvisejí, ovlivňují se a lze mezi nimi najít zpětné vazby.

### 3.2 Základní kroky zpracování obrazu

Postup zpracování a rozpoznávání obrazu reálného světa se daří obvykle rozložit do **posloupnosti základních kroků**:

1. **Snímání, digitalizace a uložení obrazu v počítači.** Snímání je převod vstupní optické veličiny na elektrický signál spojitý v čase i úrovni. Vstupní informací může být jas (z TV kamery, scanneru), intenzita rentgenového záření, ultrazvuk, tepelné záření aj. Snímat se může v jednom nebo více spektrálních pásmech. Pro barevné snímání stačí tři spektrální složky (červená, zelená, modrá), zpracováváme-li obraz, který bude pozorován člověkem (odpovídá typům čidel na sítnici oka).

Digitalizace převádí vstupní spojitý signál do diskrétního tvaru. Signál je popsán funkcí  $f(i, j)$  souřadnic v obraze. Funkční hodnota odpovídá např. jasu. Vstupní signál je vzorkován a kvantován. Výsledkem je matice přirozených čísel popisujících obraz. Prvek matice = bod, pixel, dále nedělitelná jednotka.

Existují i jiné možnosti reprezentace vstupního obrazu v počítači. Častým případem je popis obrazu

koeficienty dvourozměrné Fourierovy transformace. Výhodou je to, že Fourierovu transformaci lze provést okamžitě optickými prostředky již před digitalizací.

2. **Předzpracování.** Cílem je potlačit šum a zkuslení vzniklé při digitalizaci a přenosu obrazu nebo zvýraznění určitých rysů obrazu podstatných pro další zpracování (například hledání hran).
3. **Segmentace obrazu na objekty.** je jeden s nejtěžších kroků. Za objekty se podobně jako v obecné teorii systémů považují ty části obrazu, které nás z hlediska dalšího zpracování zajímají. Při segmentaci se tedy zhusta využívá znalosti interpretace obrazu (sémantika).
4. **Popis objektů.** Lze je popsat buď kvantitativně pomocí souboru číselných charakteristik a/nebo kvalitativně pomocí relací mezi objekty.
5. **Porozumění obsahu obrazu (často jen klasifikace objektů).** Jednoduchý případ je klasifikace objektů do několika předem známých tříd. V obecném případě představuje porozumění **interpretaci obrazových dat**, o kterých se předem nic nepředpokládá. Porozumění obrazu je potom založeno na znalosti, cílech, tvorbě plánů a využití zpětných vazeb mezi různými úrovněmi zpracování. Používají se postupy vyšší úrovně zpracování, využívá se technik znalostních systémů.

Analýza obrazu je složitější, když je při vyhodnocení třeba brát v úvahu nejen skalární obrazovou funkci závislou na dvou souřadnicích v rovině, která odpovídá jednomu statickému monochromatickému obrazu. **Obrazová funkce** nemusí být statická, ale **v čase proměnná**. Další komplikací je, když obrazová funkce není skalární, ale **vektorová** (vícerozměrná) – vícespektrální obrazová funkce, velmi často ale pro jednoduchost zpracovávána jako nezávislé monochromatické obrazy.

### 3.3 Proč je porozumění v úlohách 3D počítačového vidění těžké?

Je-li senzorem jediná kamera, informace se ztrácí perspektivním zobrazením. Zpětná úloha odvození 3D vlastností objektů z obrazu kamery má tedy nekonečně řešení. Lze řešit jen za použití dalších omezení. Druhou komplikací je vztah mezi měřeným jasnem a tvarem povrchu 3D objektů v obraze – jas závisí na mnoha vlivech (odrazivost povrchu, poloha a vlastnosti zdrojů světla, orientace povrchu). Úloha určení 3D vlastností objektů na základě pozorovaného jasu je také nedostatečně určená úloha. Třetí překážkou je veliké množství obrazových dat a v neposlední řadě šum, který je v reálném obraze vždy přítomen.

V praktických aplikacích se snažíme, abychom úlohu porozumění neřešili v její plné obecnosti. Jde-li to, převedeme trojrozměrnou úlohu na dvojrozměrnou nebo např. je možné vhodnou volbou osvětlení objektů nahradit složité postupy segmentace obrazu na jednoduché prahování podle jasu.

Některé prakticky zajímavé vědní a technické obory zpracovávají obrazy, které mají přímo 2D povahu. Řešení úloh, které mají 2D povahu, je mnohem jednodušší než snaha porozumět 3D scéně.

## 4 Vznik obrazu

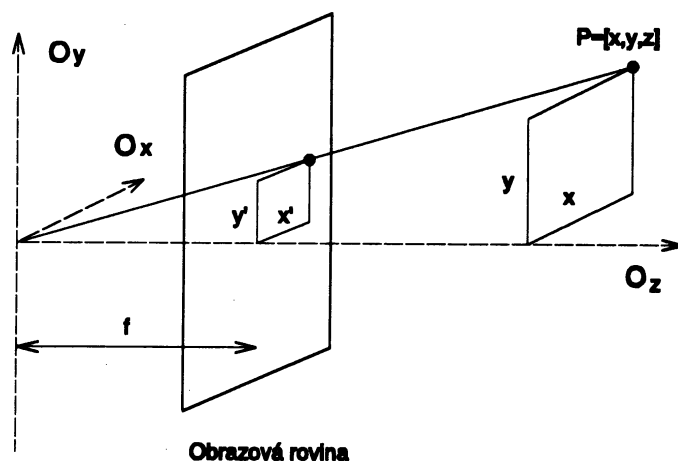
### 4.1 Formování obrazu

**Matematickým modelem obrazu** může být spojitá funkce  $f(i, j)$  dvou argumentů, souřadnic v rovině. Funkci  $f(i, j)$  se obvykle říká **obrazová funkce**. Hodnotou obrazová funkce je nejčastěji jas (intenzita), – měří většina čidel pro vstup obrazu. **Jas** souhrnně vyjadřuje vlastnosti obrazového signálu způsobem, který odpovídá jeho vnímání člověkem. Někdy může obrazová funkce reprezentovat i jiné fyzikální veličiny, jako distribuci tlaku, vzdálenost od pozorovatele, teplotu.

Prostředí, ve kterém se běžně pohybujeme, má trojrozměrnou (3D) povahu. Dvojrzměrná (2D) obrazová funkce je výsledkem **perspektivního zobrazení** části 3D prostředí – realistický model odpovídající získání obrazu v dírkové komoře. Geometrie viz obrázek 3. Necht'  $(x, y, z)$  jsou souřadnice bodu ve 3D scéně a  $f$  je ohnisková vzdálenost. Potom bod po perspektivním zobrazení má v obrazové rovině souřadnice:

$$x' = \frac{xf}{z}; \quad y' = \frac{yf}{z}.$$

Při snaze o **zjednodušení** se někdy nelineární perspektivní zobrazení nahrazuje **pravouhlym** (ortografickým) **zobrazením**, kde se předpokládá, že ohnisková vzdálenost objektivu je nekonečná.



Obrázek 3: Perspektivní zobrazení.

Popisem vzniku obrazu se podrobně zabývá fotometrie a radiometrie, pracující s pojmy jako světelný a zářivý tok a ozáření. Při analýze obrazu jsou často brány zjednodušující předpoklady na vznik obrazu. Hlavní z nich je předpoklad **Lambertovského povrchu** – ideálně matný povrch, který odráží světlo rovnoměrně do všech směrů.

## 4.2 Snímání obrazu

Pro zpracování obrazu počítačem je třeba ho nejprve sejmut – převod optické veličiny na elektrickou. Nejběžnějšími zařízeními pro snímání optického obrazu jsou: scanner, TV kamera a v poslední době též CCD kamera.

## 4.3 Digitalizace obrazu

Čidla pro vstup obrazové funkce jsou (např. TV kamera) jsou většinou zdrojem spojitého signálu (neplatí pro CCD kameru). Abychom obrazovou fci mohli zpracovat v PC, musíme nejdříve získat její digitální ekvivalent – **vzorkování** obrazu v matici  $M \times N$  bodů a **kvantování** jasové úrovně do  $K$  intervalů. Jasová fce tedy nabývá celočíselných hodnot.

Otázku vzdálenosti vzorků (plošná vzorkovací frekvence) řeší Shannonova věta (frekvence musí být alespoň dvakrát větší než nejvyšší frekvence ve vzorkovaném signálu, tedy je přizpůsobena nejmenším detailům v obraze). Dále je třeba vybrat plošné uspořádání bodů pro vzorkování (mřížka). Obvykle se používá čtvercová (častější) nebo hexagonální mřížka (výhodou je pravidelnost vzhledem k okolí bodu).



Kvantování je ovlivněno požadavkem na přesnost reprezentace obrazu. Počet kvantovacích úrovní má být dostatečně velký, aby byly přesně vyjádřeny jemné detaily obrazu, nevznikaly falešné obrysy a aby se citlivost zařízení blížila citlivosti lidského oka (člověk rozpozná najednou asi 50 úrovní jasu). Většina systémů pro digitální zpracování obrazu používá kvantování do  $k$  stejných intervalů. Použijeme-li  $b$  bitů, je počet úrovní jasu  $k = 2^b$ . Obvykle se používá 8 bitů na obrazový element. Použijeme-li pouze 1 bit, hovoříme o **binárních obrazech**. Někdy je výhodné použít nelineární kvantování, které zvětšuje rozsah těch intervalů jasu, jejichž zastoupení není v obraze pravděpodobné.

## 5 Zpracování digitálního obrazu na nízké úrovni

Uvažujme čtvercovou mřížku, reprezentovanou v PC dvojrozměrnou maticí celých čísel. Představu o rozdělení jasových úrovní v digitálním obraze získáme pomocí **histogramu** – vektor s počtem složek rovným počtu jasových úrovní. Hodnota každé složky odpovídá četnosti bodů příslušného jasu v obraze. Histogram často bývá jedinou globální informací o obraze. Můžeme ho použít při nastavování podmínek pro snímání a digitalizaci, při změnách jasové stupnice a při segmentaci.

Obraz může být zatížen různými nežádoucími poruchami, které se obvykle nazývají **šum**. Ten může vzniknout při snímání, přenosu i zpracování obrazu. Může být na obrazovém signálu buď závislý, nebo nezávislý. Šum se většinou popisuje pravděpodobnostními a frekvenčními charakteristikami (bílý šum – všechny frekvence, Gaussův šum – normální rozdělení pravděpodobnosti). Šum nezávislý na signálu je označen jako aditivní šum ( $f(x, y) = g(x, y) + v(x, y)$ ,  $v$  je šum a  $g$  je vstupní obraz). V řadě případů závisí velikost šumu na velikosti obrazového signálu – multiplikativní šum ( $f = g + vg$ ). Kvantizační šum se objeví pokud není použit dostatečný počet jasových úrovní.

Metody (před)zpracování obrazu slouží ke zlepšení obrazu z hlediska dalšího zpracování. Tato hlediska budou jiná, pokud bude výsledný obraz pozorovat na monitoru člověka, nebo když budeme usilovat o automatické zpracování obrazu. Cílem předzpracování je potlačit šum vzniklý při digitalizaci a přenosu obrazu, odstranit zkreslení dané vlastnostmi snímacího zařízení (např. korekce zakřivení zemského povrchu u družicových snímků) nebo potlačit či zvýraznit jiné rysy obrazu důležité z hlediska dalšího zpracování.

Je důležité si uvědomit, že v průběhu předzpracování nezískáme z hlediska Shannonovy teorie žádnou novou informaci (měřenou entropií). Lze jen některou informaci potlačit nebo zvýraznit.

### 5.1 Bodové jasové transformace

**Jasové korekce.** Jas v bodě výstupního obrazu závisí pouze na jasu bodu vstupního obrazu se stejnými souřadnicemi. Slouží k odstranění systematické poruchy převodní charakteristiky. Nejčastěji se předpokládá porušení obrazu multiplikativním koeficientem  $e(i, j)$ . Pro každý bod původního obrazu  $g(i, j)$  získáme na výstupu zkreslený bod  $f(i, j) = e(i, j)g(i, j)$ . Funkci  $e(i, j)$  můžeme získat sejmutím obrazu o známém průběhu jasové funkce.

**Změna jasové stupnice.** Transformace jasové stupnice nezávisí na poloze v obraze. Transformace  $T$  výchozí stupnice jasu  $p$  na novou stupnici  $q$  je dána vztahem  $q = T(p)$ . Touto transformací lze realizovat fce „negativ“, zvýšení kontrastu (vyrovnáním histogramu), prahování. Transformace jasové stupnice se realizuje pomocí tzv. vyhledávací tabulky (look up table). V terminologii PC se používá pojem paleta.

## 5.2 Geometrické transformace

Při snímání může být obraz geometricky zkreslen, je proto třeba toto zkreslení korigovat. Geometrická transformace  $\mathbf{T}_G$  je vektorová fce, která transformuje bod v rovině  $(x, y)$  do bodu  $(x', y')$ . Transformace  $\mathbf{T}_G$  je definována složkovými vztahy

$$x' = T_x(x, y), y' = T_y(x, y).$$

Transformační vztahy  $T_x$  a  $T_y$  mohou být buď známi předem (rotace, posun, zvětšení), nebo je možné hledat transformační vztah na základě znalostí původního i transformovaného obrazu (několik známých bodů).

Geometrická transformace se skládá ze dvou kroků: plošná transformace (najde k bodu vstupního obrazu odpovídající bod výstupního obrazu) a nalezení úrovně jasu. Transformační vztahy se obvykle aproximují polynomy.

## 5.3 Lokální předzpracování

Tyto metody využívají pro výpočet jasu bodu výstupního obrazu lokálního okolí odpovídajícího bodu ve vstupním obraze. Lze je rozdělit do dvou velkých skupin: vyhlazování obrazu (potlačení vyšších frekvencí obrazové fce – potlačení šumu) a gradientní operace (zdůraznění vyšších frekvencí).

Používají se lineární a nelineární metody. Lineární počítají výsledný jas pomocí diskrétní konvoluce vstupního obrazu s jádrem  $h$ :

$$f(i, j) = \sum_{(m, n) \in O} h(i - m, j - n)g(m, n).$$

Lokální operace předzpracování obvykle jen málo využívají předběžné znalosti obsahu obrazu. Vzhledem k malým rozměrům okolí zpracovávaného bodu si tuto sémantickou znalost nemohou odvodit ani v průběhu své činnosti. Zvláště pro operace vyhlazování je vhodné znát charakter poruch nebo šumu a podle něho přizpůsobit zpracování. Konvoluční lineární operace (filtry) jsou použitelné pro vyhlazování, detekci hran nebo čar.

### Filtrace

Filtrace je soubor transformací, jež mají za cíl zvýraznit nebo potlačit některé vlastnosti obrazové fce. Často požadujeme potlačení rozdílu jasu uvnitř oblastí, který zapříčiňuje šum. Volba transformace je závislá na velikosti objektů. Ve frekvenčním spektru odpovídá filtraci potlačení vysokých plošných frekvencí (výsledek dvojrozměrné Fourierovy transformace obrazu). S filtrací obrazu je spojen problém rozmazání hran.

**Obyčejné průměrování** filtruje obraz tak, že jako nový jas bodu přiřadí aritmetický průměr jasu bodů obdélníkového okolí. Nevýhodou je velké rozmazávání hran. K potlačení šumu dojde proto, že uvažujeme-li aditivní šum s náhodným rozdělením se střední hodnotou 0 a směrodatnou odchylkou  $\sigma$ , pak směrodatná odchylka zprůměrované náhodné veličiny je  $\frac{\sigma}{\sqrt{n}}$ . Pokud máme k dispozici  $n$  obrazů téže předlohy, lze filtrovat šum bez rozmazání hran průměrováním přes tytéž body v různých obrazech.

Obyčejné průměrování je speciální případ diskrétní konvoluce s jádrem (pro okolí 3x3):

$$h = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Někdy se zvětšuje váha některých bodů masky aby se lépe aproximoval šum s Gaussovským rozdělením.

$$h = \frac{1}{10} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad h = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}.$$

Potíže s rozmazáváním hran eliminují další filtrační metody, které průměrují jen tu část okolí, ke které bod pravděpodobně patří. Často se používá **filtrace metodou mediánu** (50% kvantil, vychází z metod robustní statistiky). Nevýhodou je porušení tenkých čar a ostrých rohů. Tento nedostatek se dá obejít použitím jiného tvaru okolí, např. kříž. Metoda mediánu je speciální případ **OS (order statistic) filtrace** – statistika počítaná z posloupnosti seřazených bodů okolí (medián, min, max, lineární kombinace).

**Filtrace metodou rotující masky** se snaží podle homogenity jasu najít k filtrovanému bodu část jeho okolí, ke které pravděpodobně patří. Pro výpočet průměrováním je pak použita jen tato část. Metoda nerozmazává hrany a má mírně ostřicí charakter.

### Detekce hran, ostření

Hrana je vektorová veličina určená velikostí a směrem (vychází z gradientu obrazové fce) a indikuje body obrazu ve kterých dochází ke změnám obrazové fce. Hrana je vlastnost každého pixelu. Operátory pro detekci hran vycházejí z parciálního diferenciálního operátoru. Zajímáme-li se pouze o velikost gradientu používá se často Laplaceův operátor, aproximující druhou derivaci.

Gradientních operátorů lze využít i pro ostření obrazu – zdůraznění vysokých frekvencí. Pro obraz  $f$ , který je výsledkem ostření obrazu  $g$ , platí

$$f(i, j) = g(i, j) - C \quad S(i, j),$$

kde  $C$  je kladný součinitel udávající sílu ostření a  $S(i, j)$  je hranový operátor. Hranové operátory lze rozdělit do dvou skupin:

- Operátory aproximující derivace obrazové fce diferencí pomocí diskrétní konvoluce (Laplaceův operátor, Robertsův operátor).
- Operátory založené na hledání hran v místech, kde druhá derivace obrazové fce prochází nulou – inflexe (operátor Marra a Hildrethové, Cannyho detektor).

**Robertsův operátor** používá okolí 2x2 (dostí malé a tudíž je citlivý na šum):

$$h_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad h_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

a velikost gradientu se určí vztahem:

$$|\nabla g(i, j)| = |g(i, j) - g(i + 1, j + 1)| + |g(i, j + 1) - g(i + 1, j)|.$$

**Sobelův operátor** lépe aproximuje první parciální derivace, proto je směrově závislý. Dvě masky pro okolí 3x3 jsou:

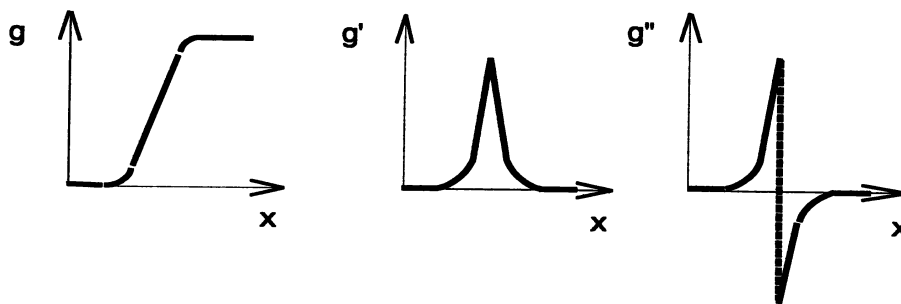
$$h_1 = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}, \quad h_2 = \begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \\ -2 & -1 & 0 \end{bmatrix}.$$

Ostatních 8 masek vznikne jen pootočením.

Předchozí operátory mají hlavní nevýhodu – závislost na šumu a na velikosti objektů v obraze. Nový hranový detektor se nazývá hledání průchodů nulou. Prakticky je jednodušší hledat průchody druhé derivace nulou než maxima prvních derivací. Druhá derivace obrazové fce se dá velmi dobře najít pomocí filtru s normálním rozdělením. Druhou derivaci nám poskytne Laplaceův operátor. Při jeho přímém použití jsou problémy se šumem, díky linearitě operací lze však zaměnit pořadí vyhlazování Gaussovým filtrem  $G$  pomocí konvoluce  $\circ$  a druhé derivace. Pro průchody nulou potom platí:

$$\nabla^2(G \circ g) = (\nabla^2 G) \circ g = 0.$$

Derivace Gaussova filtru  $\nabla^2 G$  se dá spočítat předem analyticky.



Obrázek 4: Ilustrace průchodu nulou.

**Cannyho hranový detektor** (podobný  $\nabla^2 G$ ) realizuje hledání nejlepšího rozlišení hran. Je optimální pro skokové hrany vůči třem kritériím. *Detekční kritérium* zajišťuje neopomenutí významných hran. *Lokalizační kritérium* minimalizuje rozdíl mezi skutečnou a nalezenou pozicí hrany. *Kritérium jednoznačnosti* zajišťuje, aby detektor nereagoval vícekrát na jednu hranu. Cannyho detektor se užívá dodnes uzavřel snahu o nalezení ideálního detektoru.

## 6 Analýza 2D obrazu

Tato část popisuje metody analýzy obrazu. Jsou vynechány metody klasifikace (automatického třídění), které jsou obsahem jiné otázky.

### 6.1 Segmentace

Segmentace slouží k nalezení objektů v obraze – oddělení objektu od pozadí. Pokud soubor nalezených oblastí jednoznačně koresponduje s objekty vstupního obrazu, jde o **kompletní segmentaci**, pokud

přímo nesouhlasí, jde o **částečnou segmentaci**. Pro kompletní segmentaci je nezbytná spolupráce s vyšší úrovní zpracování a znalost řešeného problému. Podle dominantní vlastnosti využívané pro segmentaci lze metody rozdělit do tří skupin: metody využívající globální znalosti obrazu (reprezentované obvykle histogramem), postupy vycházející z určování hranic mezi oblastmi a postupy vytvářející oblasti.

### Segmentace prahováním

Prahování je nejrychlejší segmentační metodou, lze ji provádět v reálném čase. Je to transformace vstupního obrazu  $f$  na výstupní binární obraz  $g$  dle vztahu:

$$g(i, j) = \begin{cases} 1 & \text{pro } f(i, j) \geq T, \\ 0 & \text{pro } f(i, j) < T, \end{cases}$$

kde  $T$  je předem určená konstanta – práh a  $g(i, j) = 1$  pro obrazové elementy náležící objektům. Správná volba prahu je zásadní problém – lze ji určit interaktivně nebo automaticky (analýzou histogramu). Někdy je třeba pracovat s různým prahem v různých částech obrazu.

### Segmentace pomocí detekce hran

Používá se některý z hranových operátorů. Nalezený obraz hran je třeba nejprve prahovat. Poté následuje další zpracování, které spojuje hrany do řetězců, které odpovídají průběhům hranic objektů. Známe-li vlastnosti tvarů objektů v obraze, je výhodné jich využít – heuristické sledování hranice. Hrany nalezené hranovým operátorem tvoří uzly grafu a k hledání hranic je možné využít postupů umělé inteligence (heuristické prohledávání grafu,  $A^*$ ).

K vyhledávání hranice se také používá Houghova transformace. Ta se používá v případě, že obraz obsahuje předměty známého tvaru (a velikosti). Hledaný předmět je parametrizován (např. přímka má rovnici se dvěma parametry) a každý bod hrany „hlasuje“ do prostoru parametrů (akumulátor) pro všechny možné objekty jimž může patřit. Objekty (jejich parametry) jsou potom určeny maximy v akumulátoru.

### Segmentace narůstání oblastí

Metody narůstání oblastí se používají v obrazech se šumem, v nichž se hranice určují obtížně. Významnou vlastností oblasti je její homogenita. Základní myšlenkou segmentace narůstáním oblastí je rozčlenit obraz do maximálních homogenních souvislých oblastí. Kritérium homogenity se může opírat o jasové vlastnosti, texturu či o vytvářený model obrazu. Používá se algoritmus spojování oblastí, které jsou dle kritéria homogenity „podobné“. Další možností je algoritmus štěpení oblastí, který velké nehomogenní oblasti dělí na menší. Rovněž je možná kombinace obou postupů.

## 6.2 Popis objektů

K rozpoznání objektů je potřebný jejich popis.

**Popis hranice objektů.** Používají se různé metody:

- aproximace úsečkami – seznam souřadnic krajních bodů,
- řetězové kódy – kroky sledování hranice,

- křivka  $\varphi - s$  – spojitý případ řetězového kódu,
- Fourierovské popisovače – koeficienty F-transformace,
- délka hranice,
- popis křivosti hranice,
- popis segmentů hranice

**Reprezentace oblastí.** Segmentované oblasti je třeba nejprve identifikovat barvením – přiřazení identifikátorů jednotlivým oblastem. Dále se používají tyto charakteristiky:

- lineární pole příslušnosti k oblasti,
- velikost oblasti,
- jas – střední hodnota a rozptyl,
- podlouhlost, pravoúhlost, výstřednost,
- projekce – výška, šířka,
- směr,
- kompaktnost,
- momenty, definované vztahem  $m_{pq} = \iint_{-\infty}^{\infty} x^p y^q f(x, y) dx dy$ ,
- centrální momenty,  $m_{pq} = \iint_{-\infty}^{\infty} (x - x_c)^p (y - y_c)^q f(x, y) dx dy$ ,
- konvexní obal.

## 7 Analýza pohybu

Z hlediska praxe lze nalézt tři typy úloh:

1. Detekce pohybu – optický hlídač, ostraha objektů apod.
2. Sledování trajektorie – cílem je nalézt umístění pohybujících se objektů, případně sledování jejich trajektorií či predikce další dráhy.
3. Určení 3D vlastností objektů

Probíhá-li analýza postupem závislým na detekci pohybujících se objektů, je výhodné používat následujících předpokladů o pohybu tuhých těles:

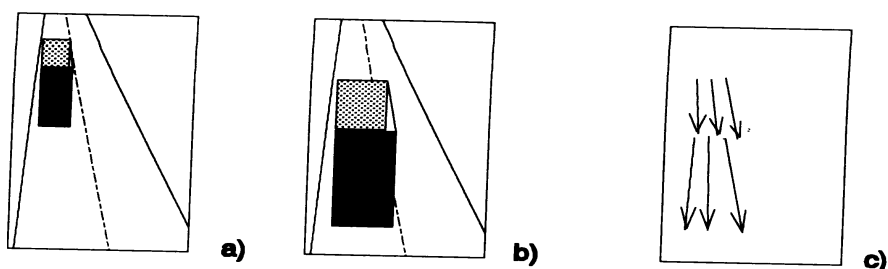
- a) **Předpoklad maximální rychlosti.**
- b) **Předpoklad malého zrychlení** vzhledem k nenulové hmotnosti.
- c) **Předpoklad společného pohybu** tuhých těles.

## 7.1 Rozdílové metody analýzy pohybu

Jednoduchá detekce pohybu se provádí výpočtem rozdílů mezi obrazy. Rozdílový obraz je binární obraz takový, že hodnoty nula představují sobě odpovídající místa dvou obrazů, v nichž nedošlo k významné změně (větší než stanovený práh) jasu mezi okamžiky snímání. Tím lze detekovat pohyb objektu jasově odlišného od pozadí. Z rozdílového obrazu nelze rozhodnout o směru pohybu objektů, proto se užívá akumulativně rozdílový obraz. Hodnota akumulativního rozdílu říká, kolikrát se hodnota jasu na snímcích posloupnosti lišila od hodnoty jasu v prvním obraze. Jednotlivé obrazy jsou před připočtením k akumulátoru váženy váhovými koeficienty. Potom směr monotónního nárůstu hodnot v akumulátoru vypovídá o směru pohybu.

## 7.2 Optický tok

Optický tok zachycuje všechny změny obrazu za čas  $dt$ . Každému bodu obrazu optického toku odpovídá dvojrozměrný vektor rychlosti, vypovídající o směru a velikosti rychlosti pohybu v daném místě obrazu (obrázek 5). Výpočet optického toku je nutným předpokladem zpracování vyšší úrovně, které dovoluje pracovat se statickým i pohyblivým umístěním pozorovatele, umožňuje určit parametry pohybu, relativní vzdálenosti předmětů v obraze apod. Určení optického toku vychází z lokálních vlastností obrazu.



Obrázek 5: Optický tok.

Druhy pohybu, které se v dynamických obrazech mohou vyskytnout, lze popsat jako kombinaci čtyř základních pohybů – translace v konstantní vzdálenosti, translace do dálky, rotace v konstantní vzdálenosti a rotace kolmá na osu pohledu. Při analýze pohybu z optického toku lze uvedené základní pohyby od sebe odlišit aplikací diferenciálních operátorů na optický tok.

## 7.3 Detekce významných bodů

Zpřístupněním metody optického toku i pro obrazy snímání v časových intervalech, které nelze považovat za velmi malé, je analýza pohybu na základě detekce významných bodů pohybujících se objektů. Jejím základem je opět vyřešit problém vzájemné korespondence sobě odpovídajících částí pohybujících se objektů v různých časových okamžicích. Známe-li vzájemnou korespondenci bodů v obrazech snímání v různých časových okamžicích, lze již snadno vytvořit obraz rychlostního pole.

Prvním krokem je nalézt v posloupnosti taková místa, která jsou významná – jsou co nejméně podobná svému okolí, vrcholy, hranice atd. Pro potřebu analýzy pohybu jsou podstatné jen významné body jejichž poloha se v čase mění. K jejich detekci lze úspěšně použít rozdílové metody. Druhým krokem je zjištění korespondence významných bodů mezi po sobě následujícími obrazy, a tak postupně vytvořit obraz rychlostního pole. Pro určení potencionálních korespondencí je výhodné užít předpoklad maximální

rychlosti pohybu, který výrazně omezí množinu potencionálních korespondujících bodů.

## 8 Analýza 3D objektů

Cílem základní úlohy počítačového vidění je porozumět objektům v 3D scéně. Řídící strategie vedoucí k porozumění obrazu lze rozdělit na dvě skupiny: řízení zdola nahoru a řízení shora dolů. Oba řídicí mechanismy se neliší typem operací, nýbrž pořadím jejich vykonávání a skutečností, zda jsou aplikována na všechna data nebo jen na data vybraná.

### 8.1 Řízení obrazovými daty – řízení zdola nahoru.

Zpracování postupuje od rastrového obrazu k obrazu segmentovanému, k popisu a rozpoznání objektů a oblastí obrazu. Vývoj strategie zdola nahoru ovlivnila práce D. Marra v 80. letech, inspirovaná principy biologického vidění. Pro řešení úlohy jsou použity čtyři úrovně reprezentace scény:

**Intenzitní obraz** – obrazová fce.

**Prvotní náčrtek** uchovává informaci podstatnou pro porozumění tvaru – hrany.

**2.5D rozměrný náčrtek** vyjadřuje vzdálenost každého bodu obrazu od pozorovatele a lokální orientaci povrchu.

**Plná 3D reprezentace objektu** zachycuje objekt ve formě vhodné k rozpoznávání. Reprezentace se opírá o geometrické vlastnosti objektu vyjádřené v souřadném systému objektu. Tuto reprezentaci se dosud nedaří uspokojivě získat.

#### Získání 2.5D náčrtku

K získání 2.5D náčrtku se používají metody „tvar z X“, kde X je jedna z několika možných technik – stereovidění, informace o pohybu, textuře, jasů, stínování apod.

**Tvar ze stera** umožňuje získat hloubkovou mapu ze dvou (někdy i více, robustnost) obrazů. Mějme dvě kamery (obrázek 6) s rovnoběžnými osami a známou vzdáleností. Najdeme-li k bodu levé kamery odpovídající bod pravé kamery, lze získat hloubku tohoto bodu od pozorovatele:

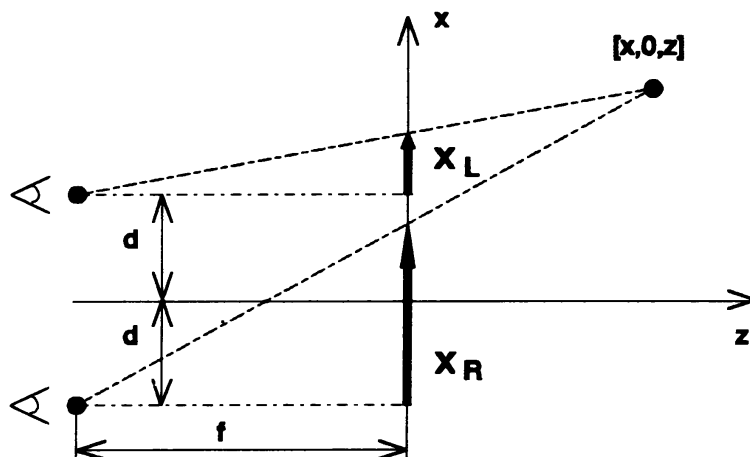
$$z = \frac{-2df}{x_L - x_R} - f,$$

$2d$  je vzdálenost kamer,  $f$  jejich ohnisková vzdálenost a  $x_L$  a  $x_R$  souřadnice zobrazených bodů.

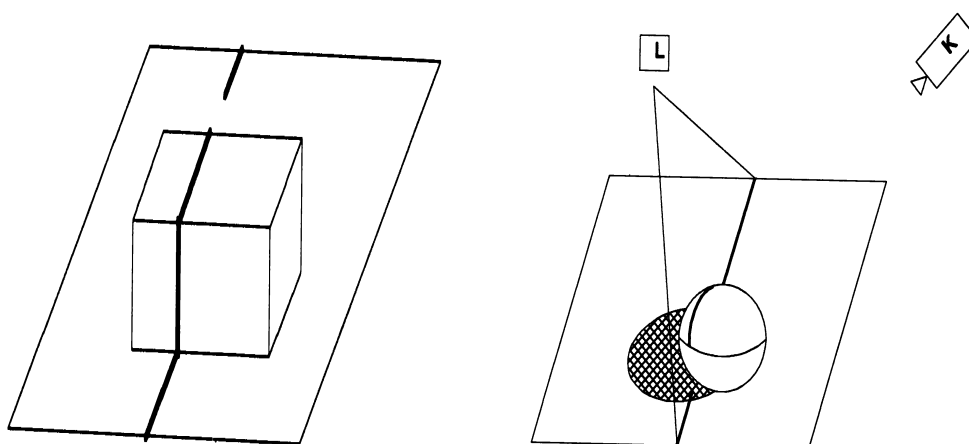
Stereo není jednoduchou úlohou, potíže způsobuje korespondenční problém – je obtížné automaticky najít odpovídající body. Pro hledání korespondencí se často užívá informace o hranách. Podobnou úlohou jako tvar ze stera je tvar z pohybu.

**Hloubkoměry** (range finders) se dělí na aktivní (vysílají energii na měřenou scénu) a pasivní. Hloubkoměr může pracovat na principu radaru (elmg. vlny nebo laser) nebo sonaru. Dalším používaným principem je triangulace – hloubkoměr na principu proužkového osvětlení (obrázek 7). Předmět je osvětlen pruhem světla a pozorován kamerou z jiného místa. Vzájemná poloha kamery a zdroje světla





Obrázek 6: Geometrie stereovidění.



Obrázek 7: Hloubkoměr na principu proužkového osvětlení.

(laser) je pevná a známá a celá sestava se pohybuje přes celý předmět. Hloubkovou mapu lze vypočítat z deformace zobrazovaného proužku na povrchu předmětu.

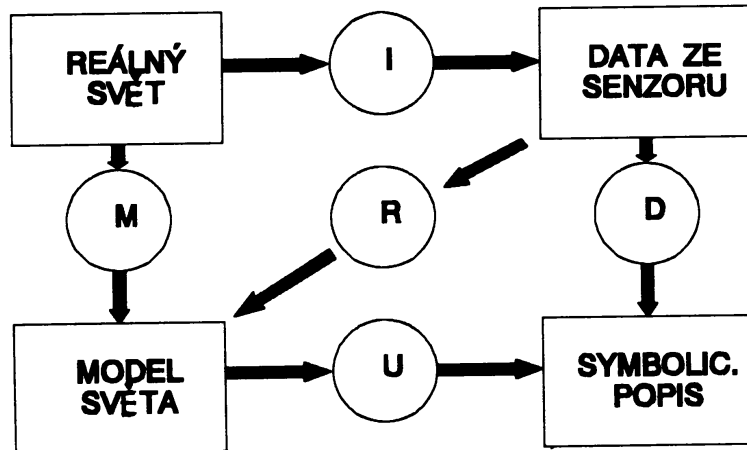
**Moiré proužky.** Scéna se osvětlí přes pravidelnou mřížku. Díky interferenci se na povrchu objektu objeví střídající se tmavé a světlé proužky s významem vrstevnic.

**Tvar ze stínování, tvar z jasu.** Jas závisí na orientaci povrchu a poloze světelných zdrojů. Za zjednodušujících předpokladů (známa poloha bodových zdrojů světla, Lambertovský povrch) lze informaci o jasu použít k zjištění 3D tvaru.

**Tvar z textury** využívá perspektivního zkreslení pravidelných opakujících se částí povrchu. Orientace povrchu se hledá z gradientu textury.

## 8.2 Řízení podle vnitřního modelu – řízení shora dolů.

Na základě znalostí vyšší úrovně je vytvořen souhrn předpokladů a očekávaných vlastností, jejichž pravdivost je ověřována v obrazových reprezentacích různých úrovní směrem shora dolů až po původní obrazová data. Jedná se o verifikaci vytvořeného modelu.



Obrázek 8: Interakce mezi částmi vidění na základě 3D modelu.

Proces získání sensorových dat  $I$  poskytuje obraz (hloubkovou mapu). Proces popisů objektů  $D$  poskytuje příznaky. Proces  $M$  slouží k získání plného 3D modelu. Proces rozpoznávání  $U$  hledá shodu mezi daty a modelem. Model v nalezené pozici se může promítnout procesem  $R$  technikami počítačové grafiky.

## 9 Geometrie pro počítačové vidění

### 9.1 Jedna kamera

Pro popis projekce kamery je výhodné používat projektivní geometrii (homogenní souřadnice). V homogenních souřadnicích lze vyjádřit posun i rotaci jednotně maticovým násobením. Je-li  $\mathbf{x} = (x_1, x_2, x_3)$  bod v 3D euklidovských souřadnicích, potom jeho homogenní souřadnice jsou  $\mathbf{X} = \alpha(X_1, X_2, X_3, 1)$ , kde  $\alpha$  je libovolné reálné nenulové číslo. Přímka je duální entita k bodu a zapisuje se stejným způsobem,  $\lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)$  (je to vlastně normálový vektor přímky v projektivním prostoru).

Projekce bodu  $\mathbf{X}$  scény na bod  $\mathbf{u}$  obrazu kamerou je lineární projekce z projektivního prostoru dimenze 3 do projektivního prostoru dimenze 2 (obrázek 9):

$$\alpha \mathbf{u} = \mathbf{P}\mathbf{X},$$

kde  $\mathbf{P}$  je projekční matice  $3 \times 4$ . Ta může být rozložena:

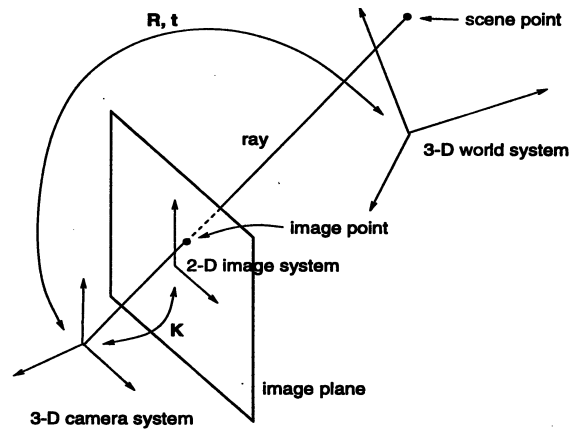
$$\mathbf{P} = \mathbf{K}(\mathbf{R} | - \mathbf{R}\mathbf{t}),$$

$\mathbf{R}$  je rotace a  $\mathbf{t}$  je posun – transformace se souřadného systému scény do souřadného systému kamery,  $\mathbf{K}$  obsahuje vnitřní parametry kamery (ohniskovou vzdálenost, rozlišení, tvar stínítka) a reprezentuje transformaci ze soustavy kamery do soustavy obrazu.

Projekční matici  $\mathbf{P}$  lze získat analyticky nebo měřením – kalibrací. Je třeba znát 3D souřadnice několika kalibračních bodů a jejich projekci kamerou. Kalibrace pak představuje řešení (přeuročené) maticové rovnice a řeší se numerickými metodami (SVD a spol).

### 9.2 Systém více kamer

Nechť bod  $\mathbf{X}$  scény je pozorován  $K$  kamerami s projekčními maticemi  $\mathbf{P}^{(1)}$  až  $\mathbf{P}^{(K)}$ :



Obrázek 9: Geometrie lineární perspektivní kamery.

$$\alpha^{(i)} \mathbf{u}^{(i)} = \mathbf{P}^{(i)} \mathbf{X}$$

Jestliže označíme

$$\mathbf{P} = \begin{pmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \mathbf{p}_3 \end{pmatrix}$$

a eliminujeme-li  $\alpha$ , můžeme rovnice psát ve tvaru

$$\begin{pmatrix} u_1 \mathbf{p}_2 - u_2 \mathbf{p}_1 \\ u_2 \mathbf{p}_3 - u_3 \mathbf{p}_2 \\ u_3 \mathbf{p}_1 - u_1 \mathbf{p}_3 \end{pmatrix} \mathbf{X} = \begin{pmatrix} \mathbf{r} \\ \mathbf{s} \\ \mathbf{t} \end{pmatrix} \mathbf{X} = \mathbf{0}$$

pro každou kameru. Rovnice jsou závislé, volíme pouze dvě pro každou kameru a sestavíme je do jedné matice:

$$\begin{pmatrix} \dots \\ \mathbf{r}^{(i)} \\ \mathbf{s}^{(i)} \\ \dots \end{pmatrix} \mathbf{X} = \mathbf{A} \mathbf{X} = \mathbf{0}.$$

Tato rovnice musí platit pro nenulové  $\mathbf{X}$ , tj  $\text{hod}(\mathbf{A}) \leq 3$  a determinanty sub-matic řádu 4 vybrané z  $\mathbf{A}$  musí být nulové. Označme determinant matice ze čtyř řádků jako  $[\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}]$ . Podle počtu vázaných pohledů existují tři skupiny vazebních podmínek pro systém více kamer:

1. *Dva pohledy.* Do sub-matice řádu 4 vybereme dva řádky z jedné a dva řádky z druhé kamery:

$$[\mathbf{r}^{(1)}, \mathbf{s}^{(1)}, \mathbf{r}^{(2)}, \mathbf{s}^{(2)}] = 0.$$

Tuto podmínku zapíšeme (po několika úpravách) jako:

$$\mathbf{u}^{(1)T} \mathbf{F} \mathbf{u}^{(2)} = 0,$$

kde  $\mathbf{F}$  je *fundamentální matice*, obsahující veškerou informaci o vzájemné poloze dvou kamer.

2. *Tři pohledy.* Dva řádky jsou vybrány z jedné kamery, třetí řádek z jiné kamery a čtvrtý rovněž z jiné kamery. To dává 12 skalárních podmínek:

$$\begin{pmatrix} [\mathbf{r}^{(1)}, \mathbf{s}^{(1)}, \mathbf{r}^{(2)}, \mathbf{r}^{(3)}] & [\mathbf{r}^{(2)}, \mathbf{s}^{(2)}, \mathbf{r}^{(3)}, \mathbf{r}^{(1)}] & [\mathbf{r}^{(3)}, \mathbf{s}^{(3)}, \mathbf{r}^{(1)}, \mathbf{r}^{(2)}] \\ \vdots & \vdots & \vdots \end{pmatrix} = \mathbf{0}.$$

Každý determinant může být přepsán jako trilineární funkce. Pouze čtyři z těchto trilinearit jsou nezávislé. Koeficienty lze sdružit do *trifokálního tenzoru*, což je analogie fundamentální matice pro tři pohledy. Výhodnější než dávat do vztahu body je dávat do vztahu bod  $u^{(1)}$  a přímky  $\lambda^{(2)}$  a  $\lambda^{(3)}$ :

$$T_i^{jk} \lambda_j^{(2)} \lambda_k^{(3)} u_i^{(1)} = 0$$

3. *Čtyři pohledy.* Každý řádek sub-matice je brán z jiné kamery. To dává  $2^4 = 16$  podmínek. Může být zapsáno jako kvadrilineární funkce s 81 nezávislými koeficienty.

## Literatura

- [1] V. Hlaváč, M. Šonka. *Počítačové vidění*. Grada, Praha, 1992.
- [2] V. Mařík, O. Štěpánková, J. Lažanský a kolektiv. *Umělá inteligence (2)*. Academia, Praha, 1997.
- [3] Milan Šonka, Václav Hlaváč, Roger D. Boyle. *Image Processing, Analysis and Machine Vision*. Chapman and Hall, London, UK, first edition, 1993.