



# Extrakce příznaků II

## Pokročilá analýza obrazu

Mgr. Markéta Trnečková, Ph.D.

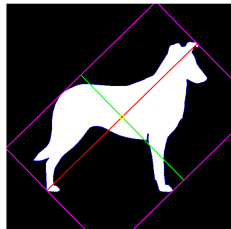
## Extrakce příznaků

- Segmentace = dělení obrazu na nějaké logické části
- Dalším krokem je popsání segmentovaných objektů kvůli dalšímu strojovému zpracování
- **Extrakce příznaků** = detekce příznaků a jejich popis
- Detekovaným příznakům přiřazujeme kvantitativní atributy
- Dělení podle toho, na jaké části je aplikujeme
  - hranice oblastí
  - oblasti
  - celé obrazy
- **Deskriptory příznaků** = způsob, jakým popisujeme jednotlivé příznaky
- Měly by být co nejméně citlivé na změny parametrů, jako jsou měřítko, translace, osvětlení, úhel, ...

## Deskriptory příznaků oblastí

- Některé jsou podobné, nebo podobně definované, jako příznaky hranic oblastí
- Některé známe z kurzu Úvod do analýzy obrazu
- **hlavní osa**, **vedlejší osa**, **ohraničující obdélník** – stejně definované jako u okrajů
- **obsah** ( $A$ ) – počet pixelů v oblasti
- **obvod** ( $p$ ) – délka okraje oblasti

- Obvod (počet pixelů hranice): 1728
- Obsah: 67302 pixelů



## Deskriptory příznaků oblastí

- **excentricita** – poměr hlavní a vedlejší osy

$$e = \sqrt{1 - \left(\frac{b}{a}\right)^2}$$

číslo mezi 0 a 1, které popisuje, jak „protažený“ je tvar

- **kompaktnost oblasti** (compactness) – míra toho, jak moc je oblast podobná kruhu

$$\text{kompaktnost} = \frac{P^2}{A}$$

bezrozměrná veličina

- **kruhovitost oblasti** (circularity) – míra toho, jak moc je oblast podobná kruhu

$$\text{kruhovitost} = \frac{4\pi A}{P^2}$$

bezrozměrná veličina

- **efektivní průměr** (effective diameter) – průměr kruhu, který má stejný obsah

$$D_{ef} = 2\sqrt{\frac{A}{\pi}}$$

- **centriod** (střed oblasti, těžiště) – představuje bod, ve kterém by byla oblast v rovnováze, pokud by měla rovnoměrnou hustotu

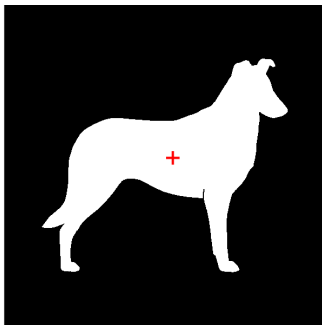
- v diskrétním obraze = průměrná poloha všech pixelů tvořících danou oblast

$$x_c = \frac{1}{N} \sum_{i=1}^N x_i, \quad y_c = \frac{1}{N} \sum_{i=1}^N y_i,$$

$(x_i, y_i)$  jsou všechny body oblasti,  $N$  je počet pixelů

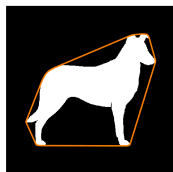
## Deskriptory příznaků oblastí

- Efektivní průměr: 292.73
- Excentricita: 0.85
- Kompaktnost: 52.13
- Kruhovitost: 0.24
- Centroid: 315.24, 284.97



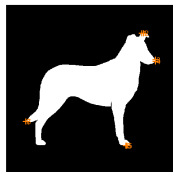
## Deskriptory příznaků oblastí

- **konvexní obal oblasti** (Convex hull) – nejmenší konvexní oblast, která obsahuje všechny body (pixely) dané oblasti
- **Solidita** (Solidity) – vyjadřuje, jak „vyplněný“ je konvexní obal oblasti.  
Je definována jako poměr plochy oblasti k ploše jejího konvexního obalu  
$$\text{solidita} = \frac{\text{plocha}}{\text{plocha\_konvexního\_obalu}}$$



Solidita: 0.52

- **extrémy oblasti** – souřadnice osmi „nejvzdálenějších“ bodů oblasti  
top-left, top-right, right-top, right-bottom,  
bottom-right, bottom-left, left-bottom, left-top



## Deskriptory příznaků oblastí

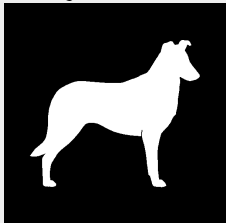
- **Eulerovo číslo**  $E$  – rozdíl mezi počtem souvislých komponent a počtem děr

$$E = C - H,$$

kde  $C$  je počet komponent a  $H$  počet děr.

### Příklad

Jaké je Eulerovo číslo následující oblasti?



# Textury

- Důležitým přístupem k popisu oblasti je kvantifikace **textury**
- Vyjadřujeme vlastnosti jako jsou:
  - Hladkost
  - Hrubost
  - Pravidelnost
- Dva přístupy:
  - Statistický – charakteristiky typu hladká, hrubá, zrnitá apod
  - Spektrální – založeny na vlastnostech Fourierova spektra, používají se především k detekci globální periodicity v obraze

# Textury

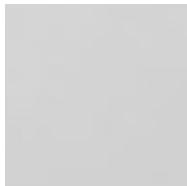
## Statistické přístupy

- Statistické momenty histogramu intenzit obrazu nebo oblasti
- $z$  je náhodná proměnná reprezentující intenzitu  
 $p(z_i)$ ,  $i = 0, 1, 2, \dots, L - 1$  je normalizovaný histogram  
 $L$  počet různých úrovní intenzity
- $n$ -tý moment vzhledem k průměru  
$$m_n(z) = \sum_{i=0}^{L-1} (z_i - m)^n p(z_i),$$
$$m = \sum_{i=0}^{L-1} z_i p(z_i)$$
- 2. moment (variance) je jeden z nejdůležitějších momentů popisující textury
- Dá se využít k vyjádření **relativní hladkosti**  
$$R(z) = 1 - \frac{1}{1 + \sigma^2(z)}$$
  
0 pro oblasti s konstantní intenzitou a blíží se 1 pro velké hodnoty rozptylu

# Textury

## Statistické přístupy

Oblast 1



$m = 210.8$   
 $m_2 = 1.6384$   
 $R = 0.6209$

Oblast 2

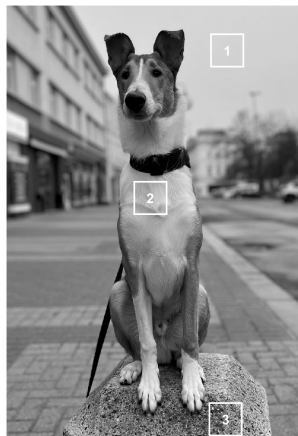


$m = 176.6$   
 $m_2 = 290.0816$   
 $R = 0.9966$

Oblast 3



$m = 116.0$   
 $m_2 = 2595.9514$   
 $R = 0.9996$

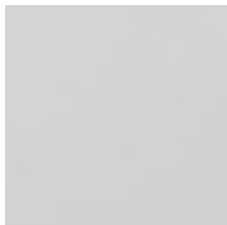


# Textury

## Statistické přístupy

- U šedotónových obrázků s intenzitami  $0, \dots, 255$  může být rozptyl velký
- Normalizujeme ho na interval  $[0, 1]$  vydělením  $(L - 1)^2$
- Nebo místo rozptylu používáme směrodatnou odchylku

Oblast 1



$$\begin{aligned}m &= 210.8 \\nm_2 &= 0.0000 \\nR &= 0.0000\end{aligned}$$

Oblast 2



$$\begin{aligned}m &= 176.6 \\nm_2 &= 0.0045 \\nR &= 0.0044\end{aligned}$$

Oblast 3



$$\begin{aligned}m &= 116.0 \\nm_2 &= 0.0399 \\nR &= 0.0384\end{aligned}$$

# Textury

## Statistické přístupy

- 3. moment vyjadřuje šikmost histogramu
- 4. moment jeho špičatost
- Vyšší momenty se hůře interpretují, ale poskytují další kvantitativní rozlišení textury

- **Uniformita:**

$$U(z) = \sum_{i=0}^{L-1} [p(z_i)]^2$$

je maximální pro obraz s konstantní intenzitou

- **Entropie** = míra variability:

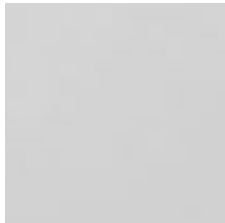
$$e(z) = - \sum_{i=0}^{L-1} p(z_i) \log_2 p(z_i)$$

pro konstantní obraz je nulová

# Textury

## Statistické přístupy

Oblast 1



$$U = 0.2169$$

$$e = 2.3117$$

Oblast 2



$$U = 0.0165$$

$$e = 6.1144$$

Oblast 3



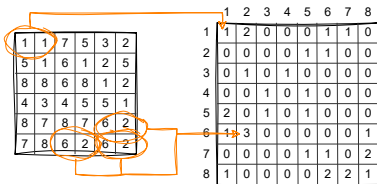
$$U = 0.0054$$

$$e = 7.6632$$

# Textury

## Statistické přístupy

- Míry založené na histogramu nenesou informaci o prostorových vztazích mezi pixely (u textur je to celkem důležité)
- Uvažujeme nejen rozdělení intenzit, ale také relativní polohu pixelů v obraze
- $Q$  je operátor, který určuje vzájemnou polohu dvou pixelů v obraze  $f$  s  $L$  možnými intenzitami
- $G$  je matice, kde každý prvek  $g_{ij}$  udává počet výskytů dvojic pixelů s intenzitami  $z_i$  a  $z_j$  v obraze  $f$  v poloze definované operátorem  $Q$
- Takto definovaná matice se nazývá **matice souvškytu úrovní šedi** (co-occurrence matrix)



$L = 8$ ,  $Q$  jeden pixel vpravo

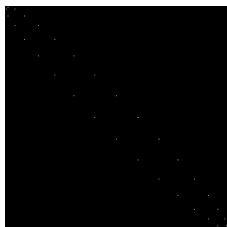
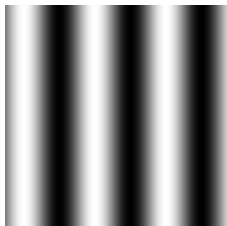
# Textury

## Statistické přístupy

- Velikost matice  $G$  závisí na počtu intenzit (pro 8-bit obraz  $256 \times 256$ )
- Většinou se barvy kvantují např. rozdělení na 8 intervalů (matice  $8 \times 8$ )
- Celkový počet dvojic pixelů splňujících podmínku  $Q$  je roven součtu všech prvků  $G$
- Pravděpodobnost výskytu dvojice  $(z_i, z_j)$   $p_{ij} = \frac{g_{ij}}{n}$
- $G$  závidí na operátoru  $Q$ , jeho vhodnou volbou lze detekovat různé vzory
- Matici  $G$  popisujeme pomocí deskriptorů:

# Textury

Matice souvýskytu



# Textury

## Statistické přístupy

### ■ Matici $G$ popisujeme pomocí deskriptorů:

- Maximální pravděpodobnost – největší hodnota matice  $G$

$$\max_{i,j}(p_{ij})$$

- Korelace – míra závislosti mezi pixelem a jeho sousedem

$$\sum_{i=1}^K \sum_{j=1}^K \frac{(i-m_r)(j-m_c)p_{ij}}{\sigma_r\sigma_c}, \quad \sigma_r \neq 0, \sigma_c \neq 0$$

$$m_r = \sum_{i=1}^K i \sum_{j=1}^K p_{ij}$$

$$m_c = \sum_{i=1}^K j \sum_{j=1}^K p_{ij}$$

- Kontrast – míra kontrastu intenzity mezi pixelem a jeho sousedem

$$\sum_{i=1}^K \sum_{j=1}^K (i-j)^2 p_{ij}$$

- Uniformita (energie)

$$\sum_{i=1}^K \sum_{j=1}^K p_{ij}^2$$

- Homogenita – prostorová blízkost prvků k diagonále matice  $G$

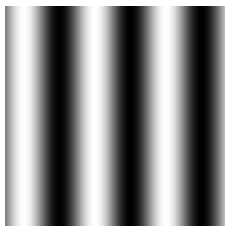
$$\sum_{i=1}^K \sum_{j=1}^K \frac{p_{ij}}{1+|i-j|}$$

- Entropie

$$-\sum_{i=1}^K \sum_{j=1}^K p_{ij} \log_2 p_{ij}$$

# Textury

## Matice souvýskytu



Max. pr.	0.040404	0.003168	0.000495
Korelace	0.979868	0.935480	0.682943
Kontrast	318.95	37.42	642.27
Uniformita	0.031527	0.000998	0.000141
Homogenita	0.093176	0.316456	0.086988
Entropie	4.994087	10.402654	12.920554

# Textury

## Spektrální přístupy

- Fourierovo spektrum je velmi vhodné pro popis periodických (nebo semiperiodických) 2D vzorů v obraze
- Tyto opakující se vzory se ve spektru projevují jako koncentrace vysokoenergetických špiček
- Vlastnosti FS (které využijeme při analýze textur)
  - výrazné peaky ve spektru určují hlavní směr texturových vzorů
  - poloha těchto peaků ve frekvenční rovině určuje základní prostorovou periodu vzorů
  - odstraněním periodických složek pomocí filtrace zůstávají neperiodické části obrazu, které lze následně popsat statistickými metodami
- Spektrum je symetrické podle počátku, takže stačí uvažovat pouze polovinu frekvenční roviny (v takovém případě by každý vzor byl reprezentován jedním peakem)

# Textury

## Spektrální přístupy

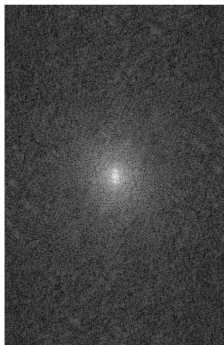
- Detekci a interpretaci těchto vlastností lze zjednodušit vyjádřením spektra v polárních souřadnicích jako funkce  $S(r, \theta)$ , kde  $r$  a  $\theta$  jsou polární souřadnice
- Pro pevný směr  $\theta$  lze  $S(r, \theta)$  chápat jako jednorozměrnou funkci  $S(r)$ , zatímco pro pevnou frekvenci  $r$  lze uvažovat funkci  $S(\theta)$
- Analýza  $S(r)$  při konstantním  $\theta$  poskytuje informace o chování spektra podél radiálního směru (např. výskyt peaků)
- Analýza  $S(\theta)$  při konstantním  $r$  popisuje chování po kružnici se středem v počátku
- Globálnější popis získáme integrací (v diskrétním případě sumací) funkcí
$$S(r) = \sum_{\theta=0}^{\pi} S(r, \theta)$$
$$S(\theta) = \sum_{r=0}^{R_0} S(r, \theta)$$
 $R_0$  je poloměr kružnice se středem v počátku
- Výsledkem jsou hodnoty  $S(r)$  a  $S(\theta)$  pro každou dvojici souřadnic  $(r, \theta)$
- Máme dvě jednorozměrné funkce, které představují spektrální energetický popis textury
- Dále můžeme použít deskriptory těchto funkcí (např. polohu maxima, střední hodnotu a rozptyl amplitud i směrových změn, a vzdálenost mezi střední hodnotou a maximem funkce a jiné)

# Textury

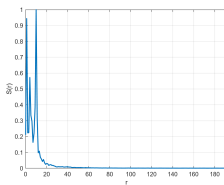
## Spektrální přístupy



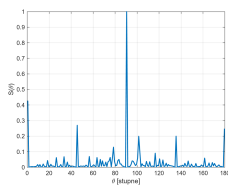
Obrázek



Spektrum



$S(r)$



$S(\theta)$

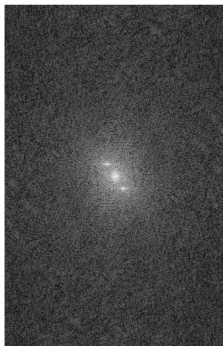
Výrazný vrchol v  $S(r)$  odpovídá přítomnosti pravidelného vzoru s určitou periodou.  
Výrazný vrchol v  $S(\theta)$  určuje dominantní směr frekvenčních složek, který je kolmý na orientaci samotného vzoru.

# Textury

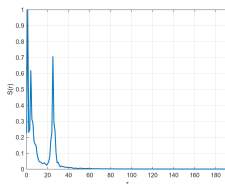
## Spektrální přístupy



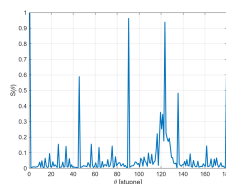
Obrázek



Spektrum



$S(r)$



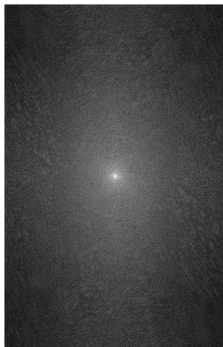
$S(\theta)$

# Textury

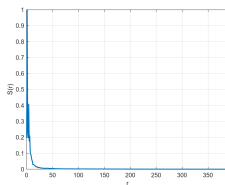
## Spektrální přístupy



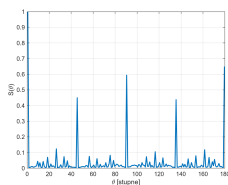
Obrázek



Spektrum



$S(r)$



$S(\theta)$

# Momentové invarianty

## Základní definice

- Momenty jsou klasické **globální deskriptory tvaru**, které shrnují rozložení intenzit v obraze nebo oblasti
- 2D moment řádu  $(p + q)$  diskrétního obrazu  $f(x, y)$  velikosti  $M \times N$  je definován vztahem
$$m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p y^q f(x, y)$$
- Tyto momenty závisejí na poloze objektu v obraze
- Proto se zavádějí **centrální momenty**, které se počítají vzhledem k těžišti objektu:
$$\mu_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (x - \bar{x})^p (y - \bar{y})^q f(x, y)$$
- Souřadnice těžiště jsou dány vztahy
$$\bar{x} = \frac{m_{10}}{m_{00}}, \quad \bar{y} = \frac{m_{01}}{m_{00}}$$
- Centrální momenty jsou invariantní vůči **translaci**, protože se vztahují k těžišti objektu

# Momentové invarianty

## Normalizace a invariance

- Abychom odstranili závislost na **měřítku**, používají se normalizované centrální momenty

- Ty jsou definovány vztahem

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}}$$

kde

$$\gamma = \frac{p+q}{2} + 1$$

- Tato normalizace zajišťuje invarianci vůči změně velikosti objektu
- Z normalizovaných centrálních momentů druhého a třetího řádu lze odvodit sedm klasických **momentových invariantů**
- Tyto invarianty jsou necitlivé na
  - translaci
  - změnu měřítka
  - rotaci
  - zrcadlení (až na případné znaménko)
- Díky tomu lze stejný objekt rozpoznat i tehdy, když je v obraze posunut, otočen nebo zvětšen

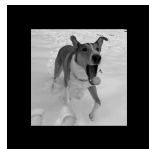
# Momentové invarianty

## Huovy invarianty

- $\phi_1 = \eta_{20} + \eta_{02}$
- $\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$
- $\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$
- $\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$
- $\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2]$   
 $+ (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$
- $\phi_6 = (\eta_{20} - \eta_{02}) [(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$
- $\phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12}) [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2]$   
 $- (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$

# Momentové invarianty

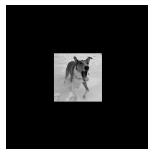
## Huovy invarianty



Originál



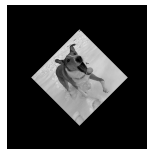
Posun



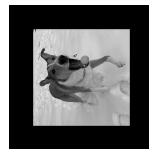
Měřítko



Zrcadlení



Otočení 45°



Otočení 90°

	$\phi_1$	$\phi_2$	$\phi_3$	$\phi_4$	$\phi_5$	$\phi_6$	$\phi_7$
Originál	0.20512	3.3534e-05	1.6835e-06	6.2149e-05	-5.2384e-10	-3.5977e-07	-3.6016e-10
Posun	0.20512	3.3534e-05	1.6835e-06	6.2149e-05	-5.2384e-10	-3.5977e-07	-3.6016e-10
Změna měřítka	0.20319	3.0896e-05	1.4156e-06	5.7327e-05	-4.3581e-10	-3.1850e-07	-2.7706e-10
Zrcadlení	0.20512	3.3534e-05	1.6835e-06	6.2149e-05	-5.2384e-10	-3.5977e-07	3.6016e-10
Otočení 45°	0.20322	3.0859e-05	1.4276e-06	5.7332e-05	-4.3591e-10	-3.1830e-07	-2.8110e-10
Otočení 90°	0.20512	3.3534e-05	1.6835e-06	6.2149e-05	-5.2384e-10	-3.5977e-07	-3.6016e-10

# Momentové invarianty

## Význam a praktické využití

- Momentové invarianty představují **kompaktní popis tvaru** objektu
- Některé invarianty nižších řádů mají i částečnou geometrickou interpretaci
  - např. první invariant souvisí s rozptylem dat kolem hlavních os
  - jiné invarianty mohou vyjadřovat míru protáhlosti objektu
- S rostoucím řádem momentů se jejich fyzikální interpretace ztrácí
- Jejich hlavní význam tedy nespočívá v intuitivní interpretaci, ale v **invarianci vůči geometrickým transformacím**
- Prakticky se používají zejména v klasických úlohách rozpoznávání tvarů
  - analýza binárních objektů
  - rozpoznávání znaků a symbolů
  - průmyslová kontrola
  - základní morfologický popis objektů
- V moderních aplikacích jsou často nahrazovány složitějšími deskriptory nebo metodami strojového učení, ale didakticky jsou velmi důležité

# Hlavní komponenty

## Příznakové deskriptory

- **Metoda hlavních komponent** (PCA) je využitelná jak pro hranice, tak regiony
- Na rozdíl od předchozích metod jsou příznaky odvozeny z více obrazů současně
- Např. barevný obraz lze chápat jako trojici obrazů, kde každý pixel tvoří vektor
- Pro  $n$  registrovaných obrazů tvoří odpovídající pixely vektor

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

- Vektory  $\mathbf{x}$  lze chápat jako náhodné veličiny
- Místo skalárních charakteristik (střední hodnota, rozptyl) používáme:
  - **střední vektor**
  - **kovarianční matici**

# Hlavní komponenty

## Statistický popis

### ■ Střední vektor populace:

$$\mathbf{m}_x = E\{\mathbf{x}\}$$

$E\{\mathbf{x}\}$  očekávaná hodnota  $\mathbf{x}$  (očekávaná hodnota vektoru (nebo matice) se počítá po složkách, tj. jako očekávaná hodnota každého prvku zvlášť)

$$E\{x\} = \sum_i x_i p(x_i)$$

### ■ Kovarianční matice:

$$C_x = E\{(\mathbf{x} - \mathbf{m}_x)(\mathbf{x} - \mathbf{m}_x)^T\}$$

### ■ Matice $C_x$ je symetrická, reálná a obsahuje:

- na diagonále rozptyly jednotlivých složek
- mimo diagonálu kovariance mezi složkami

### ■ Pokud jsou složky nekorelované, je $C_x$ diagonální

# Hlavní komponenty

## Hotellingova transformace

- Kovarianční matice má ortonormální vlastní vektory  $\mathbf{e}_i$  a vlastní čísla  $\lambda_i$
- Tyto vektory uspořádáme podle velikosti  $\lambda_i$  (od největšího)
- Sestavíme transformační matici  $A$  z těchto vektorů

- Transformace:

$$\mathbf{y} = A(\mathbf{x} - \mathbf{m}_x)$$

- Vlastnosti:

- $\mathbf{y}$  má nulovou střední hodnotu:

$$\mathbf{m}_y = 0$$

- Kovarianční matice  $C_y = AC_xA^T$  je diagonální

$$C_y = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

- Složky  $\mathbf{y}$  jsou tedy nekorelované

# Hlavní komponenty

## Rekonstrukce a redukce dimenze

- Původní vektor lze rekonstruovat:

$$\mathbf{x} = A^T \mathbf{y} + \mathbf{m}_x$$

- Pokud použijeme pouze  $k < n$  hlavních komponent dostaneme aproximaci  $\hat{\mathbf{x}}$ , tj. dochází k redukci dimenze

- Rekonstrukce:

$$\hat{\mathbf{x}} = A_k^T \mathbf{y} + \mathbf{m}_x$$

- Střední kvadratická chyba  $e_{ms} = \sum_{j=k+1}^n \lambda_j$

- Chyba je minimální při použití vlastních vektorů s největšími vlastními čísly

- Proto se metoda nazývá **transformace hlavních komponent (PCA)**

- Pro každou hlavní komponentu počítáme podíl její variability na celkové variabilitě dat

- Vlastní čísla  $\lambda_i$  reprezentují variabilitu zachycenou jednotlivými komponentami

- Procento vysvětlené variability je dáno vztahem:

$$\text{explained}_i = \frac{\lambda_i}{\sum_{j=1}^n \lambda_j} \cdot 100\%$$

- V praxi často sledujeme kumulativní variabilitu:  $\text{cumulative}_k = \sum_{i=1}^k \text{explained}_i$

## Hlavní komponenty



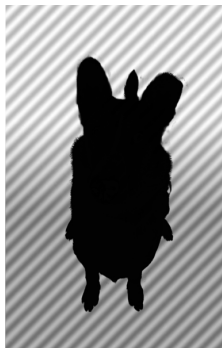
Očekávaná hodnota



1. komponenta



2. komponenta



3. komponenta

**Vlastní čísla komponent:** 0.1112, 0.0677, 0.0231, 0.0008,  $-0.0000$

PCA identifikuje hlavní směry variability mezi obrazy, přičemž konstantní struktury (např. objekt) jsou potlačeny

**Procenta vysvětlené variability:** 51.9663, 32.0230, 15.7365, 0.2742, 0.0000

# Hlavní komponenty

## Procento vysvětlené variability

- **Procento vysvětlené variability** udává, kolik z celkových rozdílů mezi obrazy zachycuje jednotlivá hlavní komponenta (tj. kolik informace nese)
- V tomto případě první tři komponenty vysvětlují téměř veškerou variabilitu ( $\approx 99.7\%$ ), přičemž první komponenta sama zachycuje více než polovinu ( $\approx 52\%$ )
- Zbývající komponenty mají zanedbatelný vliv, což znamená, že data lze dobře popsat pomocí malého počtu komponent

# Hlavní komponenty

Rozdíl od očekávané hodnoty



Světlé



Tmavé



Přechod



Sinus



Textura

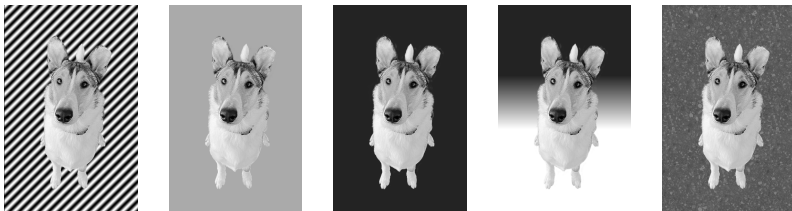
# Hlavní komponenty

## Rekonstrukce

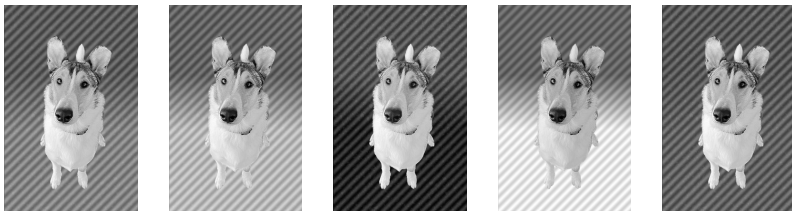
- Rekonstrukce z jedné hlavní komponenty používá pouze průměrný obraz a nejvýznamnější směr variability v datech
- Konstantní části obrazů se projeví zejména v průměrném obraze
- Proměnlivé části obrazů se projeví v hlavních komponentách
- Pokud rekonstrukce z 1 komponenty odpovídá původním obrazům dobře, znamená to, že dominantní část variability lze popsat jedinou komponentou

# Hlavní komponenty

## Originál



## Rekonstrukce z 1 komponenty



# Hlavní komponenty (PCA)

## Rekonstrukce z 1 komponenty

- Rekonstrukce obrazu pomocí jedné komponenty:

$$\hat{\mathbf{x}} = \mathbf{m}_x + a_1 \mathbf{u}_1$$

- Význam:

- použijeme pouze nejdůležitější směr variability
- silná redukce dimenze

- Výsledek:

- zachová hlavní změny (pozadí)
- ztratí detaily

- Chyba rekonstrukce:

$$\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}}$$

- ukazuje, co PCA nedokázalo popsat

# Hlavní komponenty

Chyba



Sinus



Světlé



Tmavé



Přejchod



Textura

## Příznaky celého obrazu

- Dříve zmíněné metody lze použít v případě, že je možné jednotlivé objekty dobře segmentovat
- V některých aplikacích to ale možné není
- Hledání příznaků, které jsou založené na celém obraze
- Dvě hlavní metody, které se v současnosti používají:
  - Založená na detekci rohů
  - Práce s celými oblastmi v obraze

# Příznaky celého obrazu

## Harrisův–Stephensův detektor rohů

- Intuitivně chápeme roh jako rychlou změnu směru křivky
- Rohy jsou vhodné příznaky, protože jsou výrazné a relativně invariantní vůči změně pohledu
- Používají se pro párování obrazových příznaků v různých aplikacích (trasování, registrace, ...)
- **Základní myšlenka** – roh detekujeme pomocí malého okna, které posouváme po obraze
- Okno je navrženo tak, aby měřilo změny intenzity. Zajímají nás tři situace:
  - oblasti s nulovou (nebo malou) změnou intenzity ve všech směrech (konstantní oblast),
  - oblasti se změnou v jednom směru, ale ne v ortogonálním (hrana),
  - oblasti s výraznou změnou ve všech směrech (roh).

# Příznaky celého obrazu

## Harrisův–Stephensův detektor rohů

- $f$  je obraz a  $f(s, t)$  jeho lokální výřez
- Vážený součet čtverců rozdílů je
$$C(x, y) = \sum_s \sum_t w(s, t) [f(s + x, t + y) - f(s, t)]^2$$
- Posunutý obraz lze aproximovat Taylorovým rozvojem:
$$f(s + x, t + y) \approx f(s, t) + xf_x(s, t) + yf_y(s, t)$$
- Dosazením dostaneme:

$$C(x, y) = \sum_s \sum_t w(s, t) [xf_x(s, t) + yf_y(s, t)]^2$$

nebo maticově

$$C(x, y) = \begin{bmatrix} x & y \end{bmatrix} M \begin{bmatrix} x \\ y \end{bmatrix}$$

$$M = \sum_s \sum_t w(s, t) \begin{bmatrix} f_x^2 & f_x f_y \\ f_x f_y & f_y^2 \end{bmatrix}$$

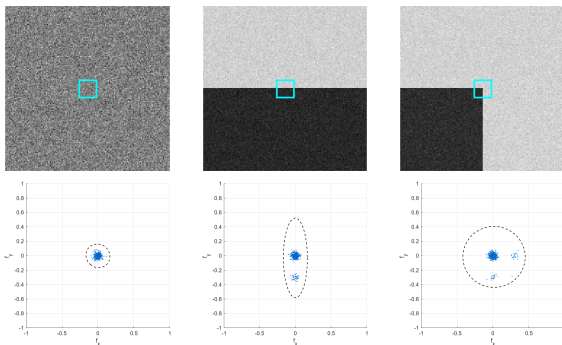
- Matice  $M$  se nazývá **Harrisova matice**
- Váhová funkce  $w(s, t)$  může být:
  - konstantní (box filtr),
  - nebo exponenciální  $w(s, t) = e^{-(s^2+t^2)/2\sigma^2}$

# Příznaky celého obrazu

## Harrisův–Stephensův detektor rohů

■ Vlastní čísla matice  $M$  umožňují rozlišit jednotlivé případy:

- dvě malá  $\rightarrow$  konstantní oblast,
- jedno velké a jedno malé  $\rightarrow$  hrana,
- dvě velká  $\rightarrow$  roh.



# Příznaky celého obrazu

## Harrisův–Stephensův detektor rohů

- Místo přímého výpočtu vlastních čísel se používá funkce odezvy

$$R = \det(M) - k (\text{trace}(M))^2$$

$\det(M)$  je determinant a  $\text{trace}(M)$  stopa matice (součet prvků na hlavní diagonále)

- Vyhodnocení:
  - $R$  velké kladné  $\rightarrow$  roh,
  - $R$  velké záporné  $\rightarrow$  hrana,
  - $R \approx 0 \rightarrow$  plochá oblast.
- Konstanta  $k$  se volí empiricky
- Detekce rohu nastane, pokud  $R > T$ , kde  $T$  je zvolený práh



# Příznaky celého obrazu

## MSER

- Detektor Harris–Stephens je vhodný pro ostré přechody intenzity (rohy)
- **MSER** jsou naopak orientované na složitější struktury
- Slouží pro získání příznaků celého obrazu a hledání korespondencí mezi obrazy
- Obraz ve stupních šedi lze chápat jako topografickou mapu
- Postupně prahujeme obraz po jednotlivých úrovních intenzity
- Výsledkem je sekvence binárních obrazů (Bílé pixely = intenzita  $\geq T$ , černé pixely = intenzita  $< T$ )
- Při zvyšování prahu vznikají černé oblasti odpovídající lokálním minimům
- Tyto oblasti mohou růst a splývat, ale nikdy se nezmenšují
- Každý prahovaný obraz obsahuje souvislé komponenty bílých pixelů
- Množina všech těchto komponent tvoří **extrémní oblasti**
- **Maximálně stabilní extrémní oblasti** = oblasti, jejichž velikost se výrazně nemění v určitém rozsahu prahů

# MSER

## Strom komponent

- Proces lze reprezentovat stromem komponent
- Každá úroveň odpovídá hodnotě prahu
- Každý uzel reprezentuje oblast  $R$
- Oblast je definována

$$\forall p \in R, \forall q \in \text{boundary}(R) : I(p) > I(q)$$

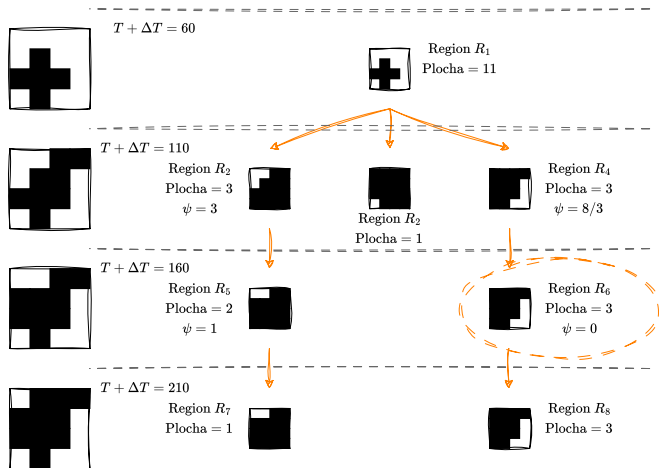
- Pro každou oblast se počítá **míra stability**

$$\psi(R_{T+(n)\Delta T}) = \frac{|R_{T+(n-1)\Delta T}| - |R_{T+(n+1)\Delta T}|}{|R_{T+n\Delta T}|}$$

- $|R|$  je velikost oblasti (počet pixelů)
- $T$  je práh,  $\Delta T$  jeho krok
- Platí  $\psi \geq 0$
- MSER jsou oblasti, kde má  $\psi$  lokální minimum
- Prakticky: velikost oblasti se výrazně nemění v rozsahu  $2\Delta T$

# MSER

225	175	90	90
125	5	90	225
5	5	5	225
125	5	225	225



# MSER

## Interpretace

- V jednoduchém případě může být stabilita nulová (oblast se nemění)
- U reálných obrazů bývá nenulová kvůli:
  - osvětlení
  - pohledu
  - šumu
- MSER jsou oblasti s minimální změnou velikosti vzhledem k okolí
- Počet oblastí lze ovlivnit:
  - volbou  $\Delta T$
  - omezením rozsahu velikostí
- MSER jsou afinně kovariantní
- Poměry ploch se zachovávají při afinních transformacích
- Detekují světlé oblasti na tmavém pozadí
- Aplikace na negativ obrazu detekuje tmavé oblasti na světlém pozadí
- Pro detekci obou typů použijeme sjednocení obou množin MSER

# MSER

## Princip



Stupně šedi



Snížená hloubka



$T_1$



$T_2$



$T_3$

# MSER

Výsledek



Všechny detekované MSER



nejsilnější MSER

# Příznaky celého obrazu

## SIFT

- **SIFT** (Scale-Invariant Feature Transform) = extrakce invariantních příznaků z obrazu
- SIFT je nejkompexnější metoda detekce a popisu příznaků, kterou si uvádíme
- Metoda obsahuje velké množství experimentálně určených parametrů
- Na rozdíl od předchozích metod je SIFT silně heuristický
- Důvodem je nedostatek znalostí, jak kombinovat jednotlivé metody
- Parametry jsou proto nastavovány experimentálně podle jejich vlivu na výkon
- **Motivace:**
  - Pokud jsou obrazy podobné (stejně měřítko, orientace), postačují rohy nebo MSER
  - Při změnách měřítka, rotace, osvětlení nebo pohledu je nutné použít metody jako SIFT
- SIFT příznaky (keypoints) jsou invariantní vůči změně měřítka a rotaci
- Jsou robustní vůči afinním deformacím, změnám 3D pohledu, šumu i změnám osvětlení
- Vstupem je obraz
- Výstupem je  $n$ -rozměrný vektor příznaků (obsahuje invariantní deskriptory)

# SIFT

## 1. Konstrukce škálového prostoru – octave

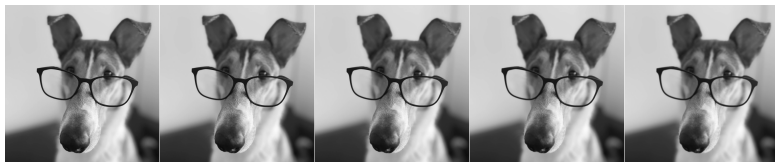
- Sestrojíme škálový prostor – víceměřítková reprezentace obrazu
- Scale space reprezentuje obraz jako množinu vyhlazených obrazů (ukládáme na zásobník)

$$L(x, y, s) = G(x, y, s) * f(x, y)$$

$$G(x, y, s) = \frac{1}{2\pi s^2} e^{-\frac{x^2+y^2}{2s^2}}$$

- Scale space je rozdělen do **octaves**
- Každá octave odpovídá zdvojnásobení parametru  $s$
- Každá octave je dále rozdělena na  $s$  intervalů
- Cílem je simulovat ztrátu detailů při zmenšování měřítka
- Je třeba zadat parametry  $\sigma$ ,  $s$ , ( $k$  se vypočítá ze  $s$ ) a počet oktáv
- Doporučené hodnoty:
  - $\sigma = 1.6$
  - $s = 2$
  - tři oktávy

# SIFT



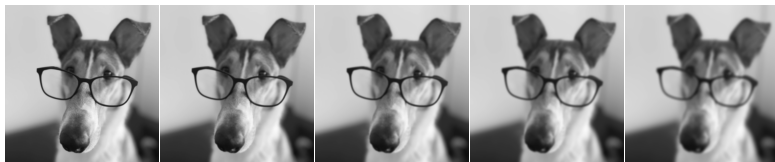
$\sigma = 0.707$

$k\sigma = 1.000$

$k^2\sigma = 1.414$

$k^3\sigma = 2.000$

$k^4\sigma = 2.828$



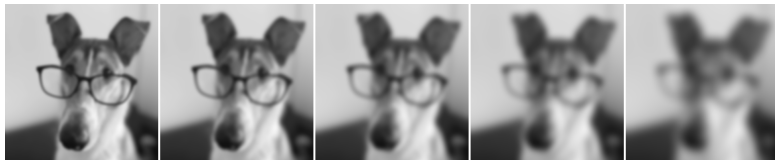
$2\sigma = 1.414$

$2k\sigma = 2.000$

$2k^2\sigma = 2.828$

$2k^3\sigma = 4.000$

$2k^4\sigma = 5.657$



$4\sigma = 2.828$

$4k\sigma = 4.000$

$4k^2\sigma = 5.657$

$4k^3\sigma = 8.000$

$4k^4\sigma = 11.314$

# SIFT

## Scale space – octaves

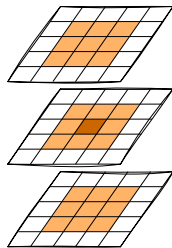
- Obrázky jsou na předchozím slidu zvětšeny na stejnou velikost, ale ve skutečnosti mají jednotlivé octave poloviční rozlišení



# SIFT

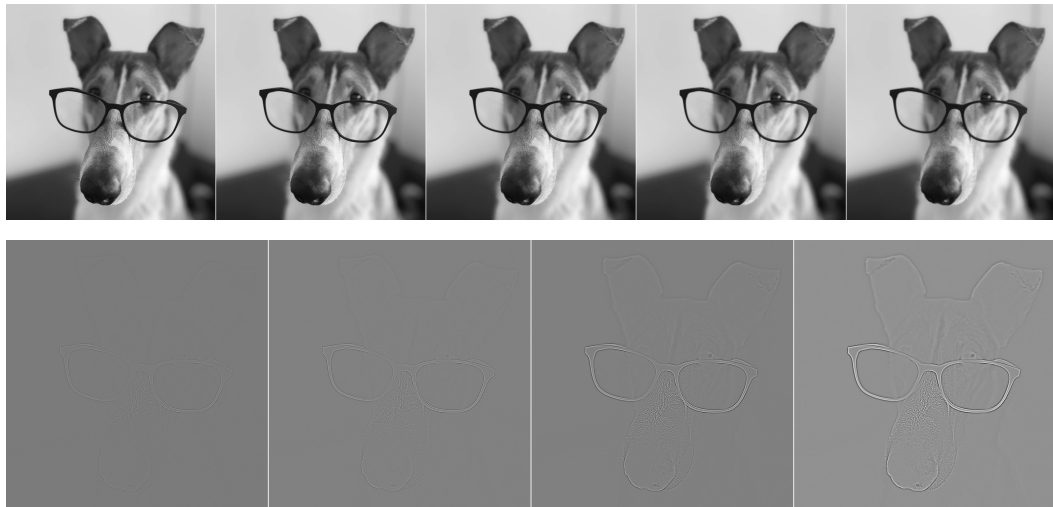
## 2. Získání počátečních klíčových bodů

- Vypočítáme rozdíl Gaussiánů  $D(x, y, \sigma)$
- Použijeme vyhlazené obrazy ve škálovém prostoru
- Nalezneme extrémy v každém obrazu  $D(x, y, \sigma)$
- Bod je extrém (maximum nebo minimum), pokud je větší nebo menší než všechny jeho sousedi
- Takto získáme **počáteční klíčové body**



# SIFT

Jedna octave a rozdíly Gaussů



# SIFT

## 3. Zpřesnění polohy klíčových bodů

- Při vzorkování spojitě funkce může její skutečné maximum nebo minimum ležet mezi vzorky
- Hodnoty  $D(x, y, s)$  interpolujeme pomocí Taylorova rozvoje
- Poloha extrému  $\hat{\mathbf{x}}$

$$\hat{\mathbf{x}} = -H^{-1} \frac{\partial D}{\partial \mathbf{x}}$$

$$\frac{\partial D}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial D}{\partial x} \\ \frac{\partial D}{\partial y} \\ \frac{\partial D}{\partial \sigma} \end{bmatrix}$$

$$H = \begin{bmatrix} \frac{\partial^2 D}{\partial x^2} & \frac{\partial^2 D}{\partial x \partial y} & \frac{\partial^2 D}{\partial x \partial \sigma} \\ \frac{\partial^2 D}{\partial y \partial x} & \frac{\partial^2 D}{\partial y^2} & \frac{\partial^2 D}{\partial y \partial \sigma} \\ \frac{\partial^2 D}{\partial \sigma \partial x} & \frac{\partial^2 D}{\partial \sigma \partial y} & \frac{\partial^2 D}{\partial \sigma^2} \end{bmatrix}$$

# SIFT

## 4. Odstranění nevhodných klíčových bodů

- Odstraníme klíčové body s nízkým kontrastem a/nebo špatnou lokalizací
- To provedeme vyhodnocením  $D$  z kroku 3 ve zpřesněných polohách
- Všechny klíčové body, jejichž hodnota  $D$  je menší než práh, se odstraní
- Doporučená prahová hodnota je 0.03
- Odstraňují se také klíčové body odpovídající hranám
- Pro  $r$  je doporučena hodnota 10

# SIFT

## 5. Výpočet orientací klíčových bodů

- Pro každý klíčový bod určujeme **velikost a orientaci gradientu**

$$m(x, y) = \sqrt{(D_x)^2 + (D_y)^2}$$

$$\theta(x, y) = \arctan\left(\frac{D_y}{D_x}\right)$$

- Sestaví se **histogram orientací** (typicky 36 binů,  $10^\circ$  na bin)
- Každý příspěvek do histogramu je vážen:
  - velikostí gradientu  $m(x, y)$
  - Gaussovským váhováním
- Dominantní orientace = maximum histogramu
- Pokud existují další významná maxima:
  - vytvoří se více klíčových bodů se stejnou pozicí, ale různou orientací
- Výsledkem je **rotačně invariantní reprezentace**

# SIFT

## 6. Výpočet deskriptorů klíčových bodů

- Pro každý klíčový bod vytvoříme **deskriptor** z jeho okolí
- Okolí je zarovnáno podle dominantní orientace (rotační invariance)
- Použije se oblast velikosti  $16 \times 16$
- Oblast se rozdělí na  $4 \times 4$  podbloky
- V každém podbloku:
  - se počítá histogram orientací gradientu (typicky 8 binů)
  - příspěvky jsou váženy velikostí gradientu a Gaussovským oknem
- Celkový deskriptor:
  - $4 \times 4 \times 8 = 128$  složek
- Deskriptor se **normalizuje** (invariance vůči změně jasu)
- Velké hodnoty se omezují (robustnost vůči kontrastu) a znovu normalizují
- Výsledkem je **128-rozměrný příznakový vektor**