

# Počítačové vidění - Otázky

Jarek Kusák

May 2025

## 1 Těžké otázky (podrobně vysvětlené)

### 1.1 1. Popište Viola-Jones metódu detekcie tváří. Ktoré 3 základné myšlienky využíva?

Viola-Jones algoritmus je prvni efektivni metoda pro detekci oblicejů v reálném čase. Jeho úspěch spočívá v kombinaci rychlých výpočtů a jednoduchých příznaků. Využívá tři klíčové myšlenky:

**1. Haarovy příznaky:** Základem detekce jsou tzv. Haarovy příznaky – obdélníkové vzory zachycující kontrastní oblasti, jako jsou světlé čelo vs. tmavé oči. Každý příznak je spočítán jako rozdíl součtu pixelů ve světlé a tmavé části:

$$f = \sum_{\text{světlá oblast}} I - \sum_{\text{tmavá oblast}} I$$

**2. Integrální obraz:** Aby bylo možné Haarovy příznaky spočítat velmi rychle, používá se tzv. integrální obraz, kde hodnota v bodě  $(x, y)$  reprezentuje součet všech pixelů nad a vlevo:

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y')$$

To umožňuje spočítat součet pixelů v libovolném obdélníku pomocí pouhých čtyř hodnot.

**3. Kaskáda klasifikátorů:** Namísto jednoho komplexního modelu je použita kaskáda mnoha jednoduchých klasifikátorů (např. slabé klasifikátory vybrané pomocí AdaBoost). První úroveň rychle eliminuje většinu negativních oblastí. Detaily jsou testovány jen v "nadějných" regionech.

**Intuitivně:** Metoda funguje jako vícestupňový filtr – v každé fázi propustí jen části obrazu, které připomínají obličej, čímž se drasticky snižuje výpočetní náročnost.

### 1.2 2. Popište Ittiho model vizuálnej pozornosti

Ittiho model simuluje přirozenou lidskou pozornost založenou na tzv. bottom-up principech – tedy podnětech, které přitahují pozornost bez vědomého řízení (např. kontrasty, pohyb, barvy).

1. Obraz je rozdělen do tří kanálů: intenzita (I), barva (R-G, B-Y) a orientace (typicky  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ ,  $135^\circ$  pomocí Gabor filtrů).
2. Pro každý kanál se vytvoří Gaussovská pyramida a spočítají se center-surround rozdíly (analogické receptivním polím sítnice):

$$c(s) = |F(c) - F(s)|$$

kde  $F(c)$  je výstup z "centra" a  $F(s)$  z "okolí" (větší měřítko).

3. Výsledné mapy (např. intenzita, barva, orientace) se normalizují, aby nebyly zvýhodněny příliš silné podněty.
4. Nakonec se všechny mapy sečtou do finální saliency mapy:

$$S = \sum_k N(M_k)$$

**Intuitivně:** Ittiho model napodobuje, jak si všimneme červeného objektu na zeleném pozadí – hledá nečekané nebo kontrastní jevy v různých vizuálních doménách.

### 1.3 3. Čo je navigačný prístup k výberu obrazov z databázy (CBIR)? Detailne popíšte projekčné metódy.

Navigační přístup v CBIR (Content-Based Image Retrieval) označuje vizuální prohlížení databáze obrázků, kde pozice obrázků v prostoru odpovídá jejich vizuální podobnosti. Místo vyhledávání podle klíčových slov si uživatel vybírá další snímky podle blízkosti v projekci.

Tento přístup je založen na zmenšení dimenzionality příznakového prostoru (typicky desítky až tisíce dimenzí) do 2D/3D prostoru:

- **PCA (Principal Component Analysis):** Lineární metoda, která hledá ortogonální osy s největším rozptylem dat. Projekce zachová co nejvíce informací o variabilitě dat.
- **MDS (Multidimensional Scaling):** Snaží se zachovat původní vzdálenosti mezi všemi páry obrázků:

$$\min \sum_{i < j} (d_{ij}^{\text{orig}} - d_{ij}^{\text{proj}})^2$$

- **Sammon mapping:** Upřednostňuje zachování malých vzdáleností, čímž zvyšuje lokální přesnost.
- **Nemetrické MDS:** Zachovává pouze pořadí vzdáleností (ranking), ne konkrétní hodnoty.

**Intuitivně:** Jako bys měl mapu fotek – ty s podobným obsahem jsou blízko sebe. Můžeš "cestovat" skrze obrázky podle vizuální podobnosti.

#### 1.4 4. Ako funguje hľadanie objektov pomocou lokálnych príznakov? Popíšte jednotlivé kroky. Popíšte zvolený detektor a deskriptor.

Lokální příznaky umožňují rozpoznávat objekty i při změnách měřítka, natočení, perspektivy nebo osvětlení. Proces má tři hlavní fáze:

##### 1. Detekce příznaků: Detektor Harris:

- Využívá gradientní matici:

$$M = \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}$$

- Vlastní čísla matice určují typ bodu (rovina, hrana, roh).
- Odezva:

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

##### 2. Popis příznaku: SIFT (Scale-Invariant Feature Transform):

- Okolí bodu se rozděluje na  $4 \times 4$  buňky, v každé se počítá histogram gradientů.
- Histogramy se spojí do vektoru délky 128.
- Normalizace vektoru zvyšuje odolnost vůči osvětlení.

##### 3. Přiřazení příznaků:

- Příznaky z dotazového a cílového obrazu se porovnávají (např. pomocí L2 vzdálenosti).
- Pro přesnější shodu se používá geometrická verifikace (např. RANSAC).

**Intuitivně:** Každý roh nebo významný bod má svůj "otisk prstu" – vektor, který popisuje jeho okolí. Pokud tento otisk najdeme i jinde, víme, že jde o stejný objekt.

#### 1.5 5. Popíšte Lucas-Kanade metódu a jej iteratívne vylepšenie na výpočet optického toku.

Lucas-Kanade metoda odhaduje pohyb pixelů mezi dvěma snímky. Předpokládá, že intenzita pixelu se v čase nemění (jasová konzistence) a že pohyb v malém okolí je konstantní.

**Základní rovnice optického toku:**

$$I_x u + I_y v + I_t = 0$$

$u, v$  jsou složky pohybu,  $I_x, I_y$  prostorové derivace,  $I_t$  časová.

**Postup:**

- V okolí bodu (např.  $3 \times 3$  okno) sestavíme soustavu lineárních rovnic.
- V maticovém tvaru:  $A\vec{d} = \vec{b}$ , kde:

$$A = \begin{bmatrix} I_x & I_y \end{bmatrix}, \quad \vec{b} = -I_t$$

- Řešení pomocí metody nejmenších čtverců:

$$\vec{d} = (A^T A)^{-1} A^T \vec{b}$$

#### Iterativní vylepšení:

- Použití Gaussovske pyramid – výpočet se provádí nejprve na zmenšeném obraze.
- Odhad se postupně zpřesňuje směrem k původnímu rozlišení.

**Intuitivně:** Sleduješ, jak se malý vzorek obrazu (např. roh) pohnul – hledáš nejlepší posun, který "sedí" mezi dvěma snímky.

### 1.6 6. Popište Kalmanov filter.

Kalmanův filtr je optimalizační algoritmus pro odhad stavu dynamického systému, který kombinuje predikci podle modelu a korekci pomocí měření.

Používá se tam, kde je pohyb předvídatelný, ale měření je zatíženo šumem – např. sledování objektu v obraze.

#### Fáze 1 – predikce:

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_k$$

$$P_k^- = AP_{k-1}A^T + Q$$

$\hat{x}_k^-$  je predikovaný stav,  $P_k^-$  jeho kovarianční matice.

#### Fáze 2 – korekce:

$$K_k = P_k^- H^T (H P_k^- H^T + R)^{-1}$$

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-)$$

$$P_k = (I - K_k H) P_k^-$$

**Intuitivně:** Kalman filtr je jako chytrý odhadce – má očekávání (model pohybu) a pak dostane pozorování (např. kamera). Věří oběma částečně, podle jejich spolehlivosti, a vypočítá nejlepší kombinaci.

## 2 Středně těžké otázky

### 2.1 1. Popíšte Harrisov detektor rohů. Voči ktorým transformáciám ja a nie je invariantný? Prečo?

Harrisův detektor rohů detekuje body, ve kterých se intenzita obrazu výrazně mění ve všech směrech – tedy rohy. Využívá matici gradientních součinů a analyzuje její vlastní čísla.

**Je invariantní vůči:**

- Rotaci – změna orientace neovlivní výpočet gradientu.
- Osvětlení (částečně) – pracuje s rozdíly intenzit.

**Není invariantní vůči:**

- Změně měřítka – okno zůstává stejně velké, což způsobí problémy u různých velikých objektů.
- Projekčním transformacím – mění se geometrie obrazu.

**Intuitivně:** Najdeš roh v malé fotce, ale při zvětšení se už neobjeví, protože se změnila velikost struktury, na kterou je senzor citlivý.

### 2.2 2. Popíšte Houghovu transformáciu pre priamky ľubovoľného smeru.

Houghova transformace převádí detekci přímek z obrazového prostoru do parametrického prostoru  $(\rho, \theta)$ .

- Každý bod v obraze odpovídá sinusové křivce v parametrovém prostoru.
- Přímka je definována jako  $\rho = x \cos \theta + y \sin \theta$ .
- Pokud více bodů leží na stejné přímce, jejich křivky se protínají – toto maximum odpovídá přímce v obraze.

**Intuitivně:** Každý bod v obraze "hlasuje" pro všechny přímky, které by jím mohly procházet. Přímka s nejvíce hlasy je ta správná.

### 2.3 3. Popíšte metódu SIFT (detektor a deskriptor).

SIFT (Scale-Invariant Feature Transform) je robustní metoda pro detekci a popis lokálních příznaků.

**Detektor:**

- Detekuje extrémy v prostoru měřítek (DoG – Difference of Gaussians)
- Každý bod má určenou polohu, měřítko a orientaci

**Deskriptor:**

- $16 \times 16$  okolí se rozdělí na  $4 \times 4$  bloky
- V každém bloku se vytvoří histogram orientací (8 směrů)
- Celkem 128-dimenzionální vektor

**Výhody:** Odolný vůči změně měřítka, rotaci, částečně i osvětlení.

**Intuitivně:** SIFT hledá výrazné body a "otiskne" jejich okolí do číselného vektoru.

## 2.4 4. Popište metódu SURF (detektor a deskriptor).

SURF (Speeded-Up Robust Features) je rychlejší alternativa k SIFT.

**Detektor:**

- Používá Hessovu matici a aproximační filtry založené na Haarových vlnkách
- Výpočet je urychlen pomocí integrálního obrazu

**Deskriptor:**

- Okolí bodu se rozdělí na  $4 \times 4$  bloky
- V každém bloku se počítají hodnoty Haarových filtrů v horizontálním a vertikálním směru
- Výsledkem je 64-dimenzionální vektor

**Intuitivně:** SURF funguje podobně jako SIFT, ale výpočty jsou jednodušší a rychlejší, což ho činí vhodným pro realtime aplikace.

## 2.5 5. Popište postup detekcie tváří pomocou Gaussiánov. Akým spôsobom môžeme pri nej použiť boosting?

Při detekci obličejů pomocí Gaussiánů se využívají klasifikátory, které kombinují jednoduché Gaussovské modely (např. Gaborovy filtry) pro popis textury obličeje.

**Boosting:**

- Každý slabý klasifikátor sám o sobě není moc přesný
- Boosting (např. AdaBoost) kombinuje více slabých klasifikátorů do silného
- Nejznámější použití: Viola-Jones detektor používá Haar příznaky a boosting pro výběr nejlepších z nich

**Intuitivně:** Každý slabý detektor se dívá na jiný detail – jeden pozná oči, jiný obrys. Boosting je spojí dohromady, aby rozpoznali celý obličej.

## 2.6 6. Popíšte metódy sledovania pohybov oka a analýzu získaných dát.

Eyetracking sleduje směr pohledu pomocí specializovaných kamer nebo senzorů.

### Způsoby měření:

- Videometoda – sledování zornice a odrazu infračerveného světla
- Mechanické – kontakt. čočky s magnetickým senzorem

### Typy dat:

- Fixace – stabilní pohled (informace je vnímána)
- Sakkády – rychlé přeskoky mezi fixacemi

### Analýza:

- Heatmapy (intenzita pohledu)
- Trajektorie pohledu (cesta očí)
- Srovnání různých skupin (např. expert vs laik)

**Intuitivně:** Eyetracking ukáže, co člověk opravdu sleduje – nejen co je na scéně, ale kde je jeho pozornost.

## 2.7 7. Popíšte možnosti vizualizácie dát z eyetrackingu a metódy ich porovnávania.

Vizualizace dat z eyetrackingu slouží k pochopení toho, kam a jak dlouho se člověk dívá. Existuje několik způsobů, jak tato data zobrazit:

- **Heatmapa:** Barevné překrytí obrazu, kde teplejší barvy (červená) ukazují oblasti s delšími nebo častějšími fixacemi.
- **Trajektorie pohledu:** Sekvence kruhů (fixace) spojených čarami (sakkády). Ukazuje sled pohledu.
- **AOI – Area of Interest:** Specifické oblasti analyzované zvlášť (např. oči, titulky).

### Porovnávání:

- **Statistické metody:** průměrná délka fixace, počet přechodů mezi AOI, počet fixací
- **Korelace heatmap:** podobnost rozložení pohledu mezi skupinami (např. Pearsonova korelace)
- **Scanpath porovnání:** jak moc se podobají trajektorie (např. Levenshteinova vzdálenost)

## 2.8 8. Popíšte kvantovanie farebného obrazu. Typy palet, ich vytváranie, chyba kvantovania...

Kvantování barev znamená převod obrazu z velkého množství barev (např. 16 milionů RGB) na omezený počet (např. 256) tak, aby byl vizuálně podobný.

**Typy palet:**

- **Pevná paleta:** předem definované barvy (např.  $6 \times 6 \times 6$  mřížka v RGB)
- **Adaptivní paleta:** vypočítaná z daného obrázku (např. median cut, k-means)

**Metody kvantování:**

- Histogramová kvantizace
- K-means clustering v barevném prostoru
- Median cut – dělí prostor podle největší odchylky intenzity

**Chyba kvantování:**

- Rozdíl mezi originální a kvantovanou barvou
- Měří se např. průměrnou kvadratickou chybou (MSE),  $\Delta E$  v Lab prostoru

## 2.9 9. Popíšte Moravcov a Harrisov detektor rohů.

**Moravec:** hledá rohy jako body, kde se intenzita hodně mění při posunu ve více směrech.

- Porovnává čtvercové okno s posunutými verzemi ve čtyřech směrech
- Citlivý na rotaci, šum a změnu měřítka

**Harris:** vylepšuje Moravce – využívá matici gradientních součinů a analyzuje její vlastní čísla.

- Odolnější vůči šumu
- Funguje lépe při rotaci

**Intuitivně:** Moravec se dívá na rozdíly při posunu. Harris se dívá na zakřivení intenzitní funkce – roh je místo, kde intenzita rychle mění ve všech směrech.



## 2.10 10. Popíšte hľadanie objektov pomocou prikladania šablón.

Šablonové hledání (template matching) porovnává části obrazu s danou šablonou (např. obličej, číslo).

### Metody porovnání:

- Korelace
- SSD (Sum of Squared Differences)
- Normalizovaná korelace (odolnější vůči osvětlení)

### Postup:

- Posouvání šablony po obraze
- Výpočet podobnosti v každém místě
- Detekce maxima = nejpravděpodobnější výskyt objektu

### Nevýhody:

- Citlivost na rotaci, měřítko a deformace
- Pomalejší při více velikostech a orientacích

## 2.11 11. Popíšte možnosti vizualizácie dát z eyetrackingu a ich porovnávania. Popíšte vybraný spôsob hľadania významných oblastí v obraze.

Vizualizace eyetrackingu (viz otázka 7) zahrnuje heatmapy a trajektorie pohledu.

### Významné oblasti v obraze (salientní regiony):

- Používají se saliency modely, jako např. Ittiho model
- Detekují oblasti, které kontrastují s okolím (barvou, intenzitou, orientací)

### Příklad metody:

- Vytvoření několika kanálů (intenzita, barva, orientace)
- Výpočet center-surround rozdílů v různých měřítkách
- Normalizace a sloučení do saliency mapy

**Intuitivně:** Hledáme místa, která by přirozeně přitáhla pozornost oka.

## 2.12 12. Popíšte výpočet optického toku.

Optický tok představuje vektorový popis pohybu pixelů mezi dvěma snímky. Využívá se např. pro sledování objektů nebo analýzu pohybu.

**Základní předpoklad:** Intenzita pixelu se v čase nemění – jen se posouvá:

$$f(x, y, t) = f(x + \Delta x, y + \Delta y, t + 1)$$

**Po derivaci:** získáme tzv. rovnice optického toku:

$$f_x u + f_y v + f_t = 0$$

Kde  $f_x, f_y$  jsou prostorové derivace a  $f_t$  je časová.  $u, v$  jsou složky toku.

**Problém:** Jedna rovnice, dvě neznámé  $\Rightarrow$  nutná další omezení (např. hladkost okolí).

**Řešení:**

- Lucas-Kanade – předpoklad konstantního pohybu v okolí
- Horn-Schunck – přidává hladkost jako regulační člen

## 2.13 13. Popíšte metódu K-means. Na čo sa používa?

K-means je metoda shlukování, která rozděluje data do  $K$  skupin podle podobnosti.

**Postup:**

1. Náhodně zvolíme  $K$  počátečních středů
2. Každý bod přiřadíme ke středu, ke kterému má nejbližší vzdálenost
3. Aktualizujeme středy jako průměr bodů ve skupině
4. Opakujeme, dokud se středy nemění

**Použití:**

- Kvantování barev
- Rozdělení obrázků do podobných skupin
- Shlukování příznaků v BoVW

**Nevýhody:** citlivost na výběr  $K$ , na počáteční hodnoty, neřeší nelineární tvary shluků.

## 2.14 14. Aký je rozdiel medzi vzorkovaním a kvantovaním obrazovej funkcie? Popíšte princíp oboch. Popíšte metódy kvantovania farebných obrazov. Ako sa hodnotí kvalita kvantovania?

**Vzorkovanie (sampling):** Zmena spojitého obrazu na diskretní mřížku pixelů. Určuje prostorové rozlišení obrazu.

**Kvantování (quantization):** Převod spojitých intenzit na diskretní úrovně (např. 0–255).

**Příklad:** Mějme obraz s intenzitou mezi 0–1. Vzorkování znamená, že ho reprezentujeme jen v určitých bodech (pixels). Kvantování znamená, že každému pixelu přiřadíme např. celé číslo 0–255.

**Metody kvantování barev:** Viz otázka 8 – median cut, K-means, histogramové metody.

**Hodnocení kvality:**

- MSE (Mean Squared Error)
- PSNR (Peak Signal-to-Noise Ratio)
- Perceptuální metriky ( $\Delta E$  v Lab)

## 2.15 15. Čo sú Gaussovská a Laplaceovská pyramída obrazov, uveďte príklad využitia v PV.

Pyramidy jsou vícerozměrné reprezentace obrazu v různém měřítku.

**Gaussovská pyramida:**

- Každá vyšší úroveň je rozmazaná a zmenšená verze předchozí
- Používá se pro zpracování v různém měřítku (např. SIFT)

**Laplaceovská pyramida:**

- Ukládá rozdíl mezi dvěma Gaussovskými úrovněmi (*high-pass*)
- Vhodná pro kompresi, rekonstrukci a detekci hran

**Příklad použití:** SIFT využívá rozdíl Gaussiánů (DoG), který se získává z Gaussovské pyramidy. Laplaceovská pyramida se používá v obrazové kompresi a blendingu.

## 2.16 16. Popíšte metódu mean-shift. Na čo sa používa?

Mean-shift je nelineární metoda shlukování a detekce mod v hustotě dat.

**Postup:**

1. Zvolíme bod a okno (např. kruh v prostoru příznaků)
2. Spočítáme vážený průměr bodů v okně

3. Posuneme okno do nového průměru a opakujeme
4. Konečný bod je lokální maximum hustoty

**Použití:**

- Segmentace obrazu (barevný prostor + souřadnice)
- Sledování objektů
- Shlukování bez nutnosti určovat počet shluků

**Intuitivně:** Bod se pohybuje do oblasti s vyšší hustotou dat – jako by lezl do kopce hustoty.

## 2.17 17. Popíšte 4 základné morfológické operácie $\oplus$ , $\ominus$ , $\circ$ , $\bullet$ .

Morfologické operace pracují s tvarem objektů v binárních (nebo šedotónových) obrazech. Základem je strukturální prvek (SE), který slouží jako "srovnávací maska".

- **Dilatace** ( $A \oplus B$ ): rozšiřuje objekt – přidává pixely tam, kde se SE překrývá s objektem. "Nafukuje" objekty.
- **Eroze** ( $A \ominus B$ ): zmenšuje objekt – ponechá jen pixely, kde celý SE leží uvnitř objektu. "Ohlodává" objekty.
- **Otevření** ( $A \circ B = (A \ominus B) \oplus B$ ): odstraní malé objekty a šum, zachová větší tvary.
- **Uzavření** ( $A \bullet B = (A \oplus B) \ominus B$ ): vyplní malé díry a spojí blízké části objektů.

**Intuitivně:** Eroze čistí, dilatace doplňuje, otevření odstraní rušení, uzavření zacílí mezery.

## 2.18 18. Vysvetlite pojmy hĺbka ostrosti, circle of confusion, ohnisková vzdialenosť. Ilustrujte na obrázku. Čo ovplyvňuje hĺbku ostrosti?

**Hlubka ostrosti** (DOF – Depth of Field) je rozsah vzdáleností, ve kterém se objekty jeví ostré.

**Circle of confusion (CoC):** Pokud bod není přesně zaostřený, zobrazí se jako rozmazaný kotouček. Maximální přípustná velikost CoC určuje hranici ostrosti.

**Ohnisková vzdálenost (f):** Vzdálenost mezi čočkou a obrazovým bodem při zaostření na nekonečno. Určuje zorný úhel.

**Hlubku ostrosti ovlivňuje:**

- Clona (menší clona  $\rightarrow$  větší DOF)
- Vzdálenost objektu (vzdálenější objekty  $\rightarrow$  větší DOF)
- Ohnisková vzdálenost (delší objektiv  $\rightarrow$  menší DOF)

**Intuitivně:** Pokud fotíš portrét na teleobjektiv se světelným objektivem, pozadí bude rozmazané (malá DOF). Pokud fotíš krajinářsky se širokým objektivem, všechno bude ostré (velká DOF).

## 2.19 19. Aké chyby môžu mať šošovky? Čo je príčinou jednotlivých chýb?

Hlavní optické vady čoček:

- **Sférická aberace:** Paprsky vzdálené od optické osy se lámou jinak než ty blízké. Výsledkem je rozmazaný obraz.
- **Chromatická aberace:** Různé vlnové délky (barvy) se lámou pod jiným úhlem  $\rightarrow$  barevné kontury.
- **Koma:** Bod mimo osu se zobrazí jako "kometa" – zvláště při nízké cloně.
- **Astigmatismus:** Vodorovné a svislé čáry nejsou ostře zároveň.
- **Zkreslení (distortion):** Rovné čáry se ohýbají – soudkovité nebo poduškovité.
- **Vinětace:** Ztmavení rohů obrazu – způsobeno čočkou nebo clonou.

## 2.20 20. Popíšte segmentáciu pomocou prahovania. Detailne popíšte Otsuho prahovanie.

Prahování je metoda, kdy se obraz převede na binární podle prahu  $T$ :

$$g(x, y) = \begin{cases} 1, & f(x, y) \geq T \\ 0, & \text{jinak} \end{cases}$$

**Typy:**

- Globální – jedno  $T$  pro celý obraz
- Lokální –  $T$  závisí na okolí bodu

**Otsuho metoda:**

- Předpokládá bimodální histogram (dvě třídy: objekt a pozadí)
- Pro každý možný  $T$  spočte vážený rozptyl mezi třídami
- Vybere  $T$ , které maximalizuje tento rozptyl (tj. největší separace)

**Intuitivně:** Hledáš takový práh, který nejlépe oddělí tmavé a světlé části histogramu.

## 2.21 21. Popíšte segmentáciu založenú na regiónoch – narastanie, delenie, kombinácia

Segmentace založená na regiónech předpokládá, že podobné oblasti (např. podle barvy) by měly tvořit jeden segment.

**Region growing (narastanie):**

- Začíná se zvolenými "semínky"
- Přidávají se sousední pixely, pokud splňují homogenitu (např. podobná barva)

**Region splitting (delenie):**

- Začíná se celým obrazem
- Pokud oblast není homogenní, rozdělí se (např. na čtvrtiny)

**Split and merge (kombinace):**

- Nejprve rozdělíme obraz (splitting)
- Poté spojíme podobné sousední regiony (merging)

**Intuitivně:** Rosteš, kde se to hodí, dělíš, kde to nepasuje – a na konci spojíš, co patří k sobě.

## 2.22 22. Popíšte metódu RANSAC.

RANSAC (Random Sample Consensus) je metoda pro robustní odhad parametrů modelu z dat obsahujících odlehle hodnoty (outliery).

**Postup:**

1. Náhodně vybereme minimální počet bodů nutný k určení modelu (např. 2 body pro přímku).
2. Spočítáme parametry modelu (např. rovnice přímky).
3. Ověříme, kolik dalších bodů je "v souladu" s tímto modelem (tj. mají chybu menší než  $\varepsilon$ ).
4. Opakujeme  $N$  krát a vybereme model s největším počtem inlierů.
5. (Volitelně) přepočítáme model přes všechny inliery.

**Použití:** odhad homografie, přímky, epipolární geometrie, přiřazování bodů z různých obrázků.

**Intuitivně:** Hledáš model, který nejlépe vysvětlí většinu dat, i když část dat je chybná nebo náhodná.

## 2.23 23. Popíšte deskriptor HOG.

HOG (Histogram of Oriented Gradients) popisuje tvar objektu pomocí směrových histogramů gradientů.

### Postup:

1. Spočítáme gradient obrazu (např. Sobelovým operátorem).
2. Rozdělíme obraz na malé buňky (např.  $8 \times 8$  px).
3. V každé buňce vytvoříme histogram směrů gradientů (např. 9 směrů).
4. Bloky několika buněk (např.  $2 \times 2$ ) se normalizují (např. L2 normou).
5. Všechny bloky se spojí do jednoho vektorového popisu.

**Použití:** detekce postav, rozpoznávání objektů, sledování. Slavným použitím je detektor chodců od Dalala a Triggse.

**Intuitivně:** Místo práce s jasem sleduješ, kde jsou hrany a kam směřují – a to popisuješ čísly.

## 2.24 24. Popíšte metódu BoVW.

BoVW (Bag of Visual Words) je metoda inspirovaná zpracováním textu (bag-of-words). Používá se k reprezentaci obrázku jako histogramu výskytu příznaků.

### Postup:

1. Detekce lokálních příznaků (např. SIFT) ve velké sadě obrázků.
2. Pomocí K-means se příznaky rozdělí do  $K$  "vizuálních slov" – vytvoří se slovník.
3. Každý obrázek se převede na histogram výskytu těchto slov.
4. Histogram se použije jako vstup do klasifikátoru (např. SVM).

**Výhoda:** obrázek má pevně daný počet čísel (bez ohledu na počet příznaků). Možno použít pro klasifikaci nebo vyhledávání.

**Intuitivně:** Místo abys obrázek popisoval pixel po pixelu, shrneš ho jako "sáček vizuálních slov" – jako kdybys počítal výskyty motivů.

## 2.25 25. Ako sa dá využiť metóda PCA na rozpoznávanie tvárí? Popíšte princíp.

PCA (Principal Component Analysis) je metoda redukce dimenze – zachová co nejvíce informace s co nejméně proměnnými.

### Použití v rozpoznávání tváří:

- Každá tvář se reprezentuje jako vektor (např.  $100 \times 100$  obrázek  $\rightarrow$  10 000 hodnot).

- PCA najde hlavní směry rozptylu (eigenfaces) a zredukuje rozměr (např. na 50).
- Nové tváře se promítnou do PCA prostoru a porovnají se s databází (např. Eukleidovskou vzdáleností).

**Výhoda:** kompaktní reprezentace, rychlé porovnání.

**Nevýhoda:** citlivost na osvětlení a výrazy.

**Intuitivně:** Hledáš hlavní rysy obličejů – jako bys tvořil "průměrný obličej" a popisoval odchylky od něj.

## 2.26 26. Popište Lucas-Kanade metódu na výpočet optického toku.

Lucas-Kanade metoda předpokládá, že pohyb je v malém okolí (např.  $3 \times 3$  okno) konstantní.

**Základ:** optický tok z rovnice:

$$f_x u + f_y v + f_t = 0$$

**Postup:**

1. V každém okně se sestaví soustava rovnic
2. Řeší se metoda nejmenších čtverců:  $A^T A \vec{d} = A^T \vec{b}$
3.  $\vec{d} = [u, v]^T$  = posun v čase

**Vylepšení:**

- Iterativní zpřesnění (např. warping obrazu podle odhadu)
- Pyramidový přístup pro větší pohyby

**Intuitivně:** Sleduješ malé okno a hledáš, kam se pohnulo, tak aby nejvíc sedělo s dalším snímkem.

## 3 Lehké otázky

### 3.1 1. Gaussovská pyramída obrazov a jej využitie v PV

Gaussovská pyramída je hierarchická struktura tvořená opakovaným rozmazáním a zmenšováním obrazu pomocí Gaussova filtru. Umožňuje efektivní analýzu obrazu v různých měřítkách, např. pro detekci objektů v různých velikostech (SIFT, optický tok).

### 3.2 2. Laplacovská pyramída obrazov a jej využitie v PV

Laplacovská pyramída ukládá rozdílové obrazy mezi hladšími verzemi z Gaussovské pyramídy. Využívá se pro zvýraznění hran, blending (hladké spojování obrázků) a kompresi.



### **3.3 3. Popíšte spôsob použitia spektrálnych reziduí pri hľadaní významných oblastí v obraze.**

Spektrální rezidua se získají odečtením spektrální průměrné složky od původní Fourierovy transformace obrazu. Výsledná mapa zdůrazní neobvyklé oblasti (salientní body), které výrazně vybočují od globálního pozadí.

### **3.4 4. Popíšte spôsob použitia významných frekvencií pri hľadaní významných oblastí v obraze (2 frequency-tuned metódy).**

Tyto metody využívají frekvenční charakteristiky obrazu – například výpočet barevného kontrastu v Lab prostoru nebo rozdílů mezi rozostřenými obrazy různých měřítek. Výsledkem je saliency mapa ukazující vizuálně výrazná místa.

### **3.5 5. Ako hľadáme hrany v zašumených obrazoch? Čo je LoG?**

Obraz se nejprve rozmaže Gaussovským filtrem a poté se použije Laplaceova operace (druhá derivace). LoG (Laplacian of Gaussian) zvýrazní oblasti s rychlou změnou intenzity, zatímco potlačí šum.

### **3.6 6. Popíšte Moravcov detektor rohov.**

Moravec posouvá čtvercové okno v různých směrech a měří změnu intenzity. Roh je místo, kde je změna intenzity při posunu ve všech směrech vysoká.

### **3.7 7. Popíšte detektor rohov SUSAN.**

SUSAN (Smallest Univalued Segment Assimilating Nucleus) detekuje rohy podle velikosti oblasti, kde mají sousední pixely podobnou intenzitu. Rohy mají nejmenší podobné okolí (malé USAN).

### **3.8 8. Akým spôsobom hľadáme škvrny pomocou laplaciánu Gausiánu?**

Použijeme LoG filtr ve vícero měřítkách (scale space) a sledujeme, při jaké velikosti je odezva filtru maximální. Výsledkem je detekce blobů, tedy oblastí s homogenní texturou nebo barvou.

### **3.9 9. Popíšte detektor FAST a deskriptor BRIEF.**

FAST detekuje rohy podle toho, zda je 12 ze 16 sousedních pixelů jasně odlišných od středu. BRIEF pak popisuje okolí bodu pomocí binárních testů mezi dvojicemi pixelů – velmi rychlý a kompaktní popis.

### **3.10 10. Čo je Bayerova mriežka? Ako sa používa?**

Bayerova maska je uspořádání barevných filtrů nad senzorem kamery (RGGB). Každý pixel měří jen jednu barvu a zbytek se dopočítá interpolací (demosaiicing).

### **3.11 11. Čo vieme určiť z CIE chromatického diagramu?**

CIE diagram ukazuje viditeľné barvy a umožňuje porovnať barevné priestory. Pomáha definovať rozdiely medzi farbami a nájsť doplnkové farby.

### **3.12 12. Popíšte distance transform. Na čo sa používa a ako súvisí s morfológiou?**

Distance transform priradí každému pixelu vzdialenosť k najbližšiemu pozadiu. V morfológii sa využíva pre nájdenie skeletu, centrum tvaru alebo segmentáciu.

### **3.13 13. Na čo sa používajú White top hat a Black top hat transformácie?**

White top-hat zdôrazňuje svetlé objekty menšie než SE na tmavom pozadí, black top-hat zdôrazňuje tmavé diery vo svetlom pozadí. Pomáha to odstrániť nerovnosti alebo lokalizovať defekty.

### **3.14 14. Ako nájdete hrany v binárnych a gradient v šedotónových obrazoch pomocou morfologických operácií?**

V binárnom obraze lze hrany nájsť jako rozdíl mezi dilatací a původním obrazem. U šedotónového se morfologický gradient spočítá jako rozdíl mezi dilatací a erozí.

### **3.15 15. Ako skombinujeme dva obrazy, aby nebolo vidieť prechod medzi nimi?**

Použije se Laplacovská pyramida – obrazy se rozloží do více úrovní, kombinují se v každé úrovni a nakonec se rekonstruují zpět. Výsledkem je hladké spojení.

### **3.16 16. Aký je rozdiel medzi saliency map a task map?**

Saliency map vychází z vlastností scény (barvy, kontrast) a je nezáporná. Task map závisí na úkolu – zvýrazní oblasti důležité pro daný cíl.

### **3.17 17. Čo je heatmap a ako je dostaneme z dát eyetrackera?**

Heatmapa je vizualizace frekvence pohledu – čím více se pozorovatel dívá na oblast, tím je barva teplejší. Vzniká agregací fixací z více uživatelů nebo časových

snímků.

**3.18 18. Ako by ste riešili výpočet optického toku pri veľkom pohybe?**

Použije se Gaussovská pyramida – nejprve se pohyb odhadne v nejnižším rozlišení a pak se postupně zpřesňuje v jemnějších úrovních. To řeší problém s velkými posuny.

**3.19 19. Čo je rozdielový obraz a kumulatívny rozdielový obraz?**

Rozdílový obraz je výsledek odečtu dvou snímků. Kumulativní rozdílový obraz (CDI) je součet rozdílů více po sobě jdoucích snímků – zvýrazní opakovaný pohyb.

**3.20 20. Ako nájdeme pozadie scény pre určenie pohybu?**

Pozadí lze odhadnout jako statistický model – např. medián pixelů v čase nebo adaptivní Gaussovský model (GMM), který aktualizuje rozložení pro každý pixel.