

# Počítačové videnie

Povedka z prednášok 1 až 6

## [1. prednáška]

Počítačové videnie (CV, computer vision) je:

- automatická extrakcia,
- analýza,
- a chápanie

užitočných informácií z:

- obrázka,
- sekvencie obrázkov,
- alebo z iných vizuálnych dát.

Zahŕňa teóriu a vývoj algoritmov pre automatické chápanie vizuálnych dát.

Počítačová grafika - syntéza

**Počítačové videnie - analýza**

**Machine vision - počítačové videnie, image processing -> použité v priemysle**

Nízkoúrovňové operácie - spracovanie obrazu (*vstup aj výstup je obraz*):

1. predspracovanie (preprocessing)
2. redukcia šumu
3. úprava kontrastu
4. detekcia hrán
5. morfológické spracovanie
6. zaostrovanie (sharpening)
7. spracovanie farieb
8. transformácie (Fourier, cosine, wavelet)
9. kompresia

Vysokoúrovňové operácie - počítačové videnie(*vstup je obrázok, výstupom sú získané informácie z obrázka*):

1. segmentácia
2. detekcia objektu
3. rozpoznanie objektu
4. analýza/interpretácia scény
5. hľadanie vzorov
6. sledovanie (tracking) objektov
7. detekcia pohybu
8. detekcia pózy (Madaras intensifies)

Od **nízkej (image processing)** po **vysokú (computer vision)** úroveň: **reprezentácia, prevod obrazu << vylepšenie obrazu << extrakcia hrán << hľadanie vzorov << chápanie, rozpoznávanie obrazu**

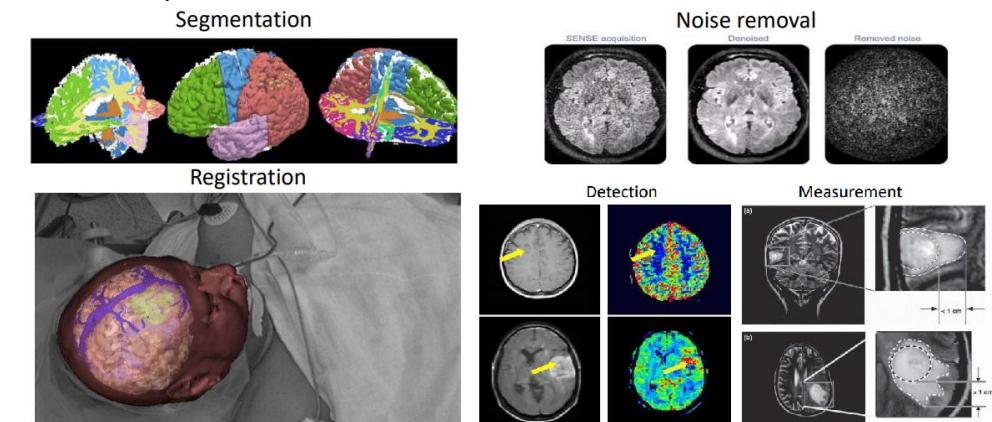
Príklady využití CV:

- rozšírená realita

- autonómne vozidlá ([Elon Musk](#))
- rozpoznanie tváre
- medicína
- robotika
- bezpečnosť ([nie L. Hudec](#))
- detekcia, sledovanie, počítanie chodcov
- snímanie odtlačkov prstov

Bližší popis:

- priemysel:
  - inšpekcia kvality (tvar, poloha, povrch, OCR, merania)
- doprava:
  - autonómne autá
  - detekcia áut/chodcov
  - štatistiky
  - meranie rýchlosťí
  - detekcia nehôd
  - autonómne riadenie - stereo vision, sensor fusion, mapping, deep learning tools...
- **medicína** (dôležitá, samostatná oblasť)
  - vytváranie 3D obrazu z 2D snímkov
  - požadovaná vysoká spoľahlivosť (**mnohokrát ide o život**)
  - rôzna interpretácia farieb



- dôvody:
  - šetrenie času doktorov
  - predchádzanie chybám
  - upozornenie na anomálie
  - zobrazovanie len tých dôležitých častí

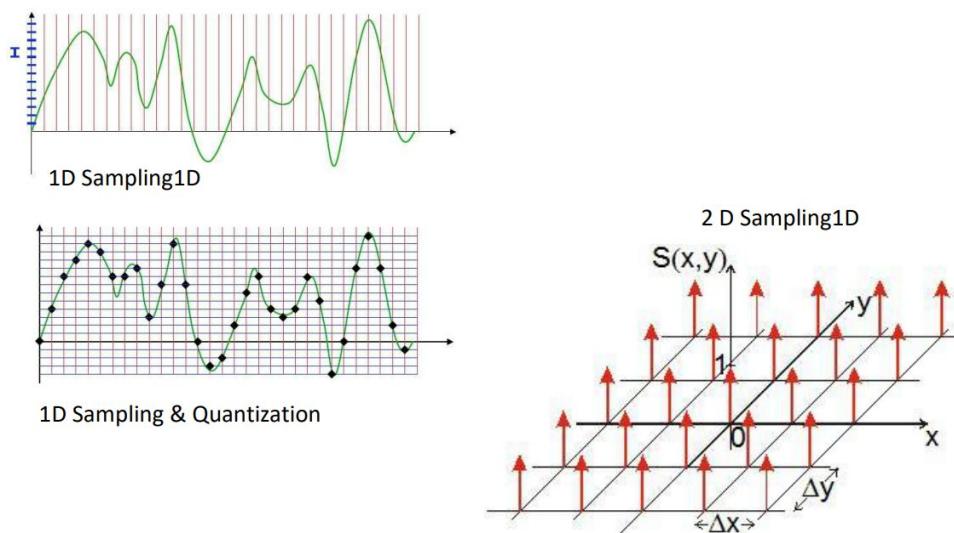
## Získavanie vizuálnych dát

Obrázok je dvoj-rozmerná funkcia  $f(x,y)$ , kde  $x$  a  $y$  sú súradnice. Amplitúda  $f(x,y)$  je jas.

Konverzia analógového signálu na digitálny (vľavo 1D, vpravo 2D):

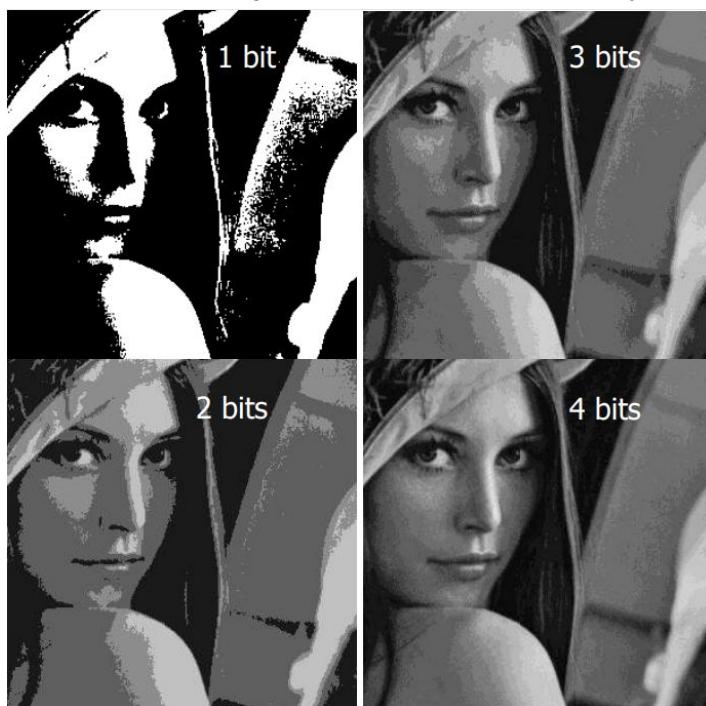
- sampling - vzorkovanie -> konverzia spojitého signálu v čase na diskrétny signál v čase. (os  $x$ )

Signál po vzorkovaní je diskrétny v čase ale nie v hodnote => kvantizácia (os y)

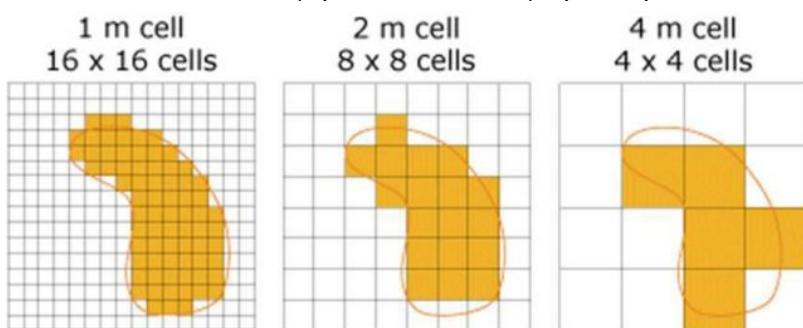


**Dynamický rozsah:** popisuje pomer medzi maximálnou a minimálnou merateľnou intenzitou svetla (biela a čierna). Meria sa v bitoch/decibeloch.

**Počet úrovní šedej** (v rámci daného rozsahu) býva mocnina dvojkys -> vieme vyjadriť bitmi.



Priestorové rozlíšenie (Spatial resolution) - počet pixelov v oboch rozmeroch (x a y):



**Nyquist sampling theorem:** vzorkovacia frekvencia by mala byť aspoň 2x najväčšia frekvencia signálu. (Napr. na širku fotky sa zmestí 8 tyčiek plota, takže rozlíšenie by malo byť aspoň 16 pixelov, aby sa dali jednotlivé tyčky rozlíšiť.)

Ked' túto frekvenciu nedodržíme (Sub-Nyquist sampled image), vzniká **Moiré pattern**.

Obrázok je matica čísel. Rozmernosť obrázku (v závislosti od počtu kanálov):

- $m.n.1$  - stupne šedej,
- $m.n.3$  - farebný obrázok (RGB),
- $m.n.4$  - RGB + alpha (priehľadnosť),
- $m.n.x$  - multi-spectral image.

Video - o jednu dimenziu navyše - **čas**; frekvencia obrázkov (FPS), pre človeka min 10 FPS.

Vizuálne dátá:

- viditeľné svetlo (360 - 770 nm)
- infračervené svetlo
- ultrafialové svetlo
- ultrazvuk
- radar
- CT, MRI, Röntgen
- hĺbková mapa

## Fotáky/kamery

Hĺbka ostrosci - čím širší je priezor, tým viac svetla dopadá (vzniká menej zašumený obraz), ale zároveň je tým menší hĺbkový rozptyl, v ktorom sú objekty ostré.

Snímač (čip) - zariadenie citlivé na svetlo, delenie:

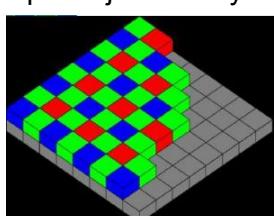
- Line (1D, napr. v priemysle nad výrobnými pásmi)
- Array (2D, klasika) senzory

Optické senzory:

- CCD - postupné posúvanie informácií po riadkoch obrázku
- CMOS - každý pixel je prístupný zvlášť

3CCD - na ihlane sa svetlo rozdelí na R, G, B zložky, kde každá zložka dopadá na zvlášť CCD senzor.

RGB Bayerova matica - každý pixel má len jeden farebný kanál (R/G/B), ostatné sa dopočítajú z okolitých.



Infračervená kamera - sníma teplotu na základe vlnovej dĺžky (teplejšie telesá vyžarujú žiarenie s menšou vlnovou dĺžkou) pre každý pixel zvlášť.

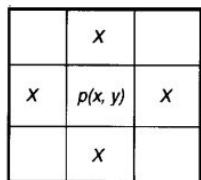
## 2,5D kamery

- 3D (resp. 2,5D) kamery
  - dva snímače vzdialené od seba určitú pevnú vzdialenosť (ako oči).
- TOF (TimeOfFlight) 2,5D kamera
  - vysiela vysokofrekvenčne (MHz) modulované svetlo
  - podľa fázového posunu sa zistí vzdialenosť bodu pre každý pixel.
- IR vzor
  - vyžarovanie vzoru (bodky) v infračervenom spektre
  - z relatívnych pozícii bodov sa vypočíta hĺbka

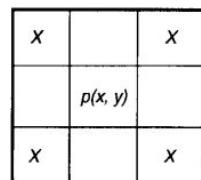
## [2. prednáška]

**Susedné pixely** (susednosť):

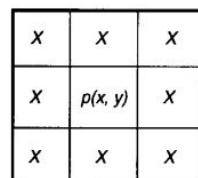
- 4 susedné pixely
- 4 diagonálne susedné pixely
- 8 susedných pixelov



(a)



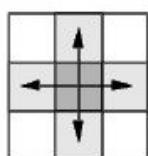
(b)



(c)

**Prepojenosť** (súvislosť):

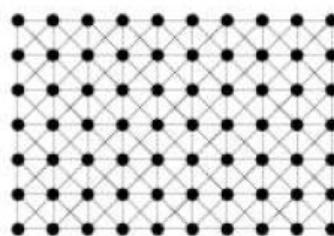
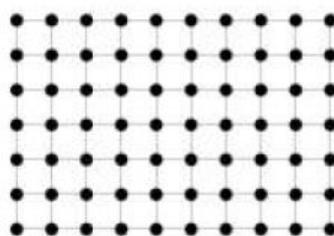
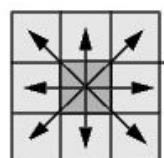
- 4-súvislé - pixely p a q sú 4-súvislé, ak sú 4 susedné a p eV and q eV
- 8-súvislé - pixely p a q sú 8-súvislé, ak sú 8 susedné a p eV and q eV



4-connected

8-connected

:



## Vzdialenosťné metriky

Všeobecné pravidlá:

1.  $D(p,q) \geq 0$  non-negativity
2.  $D(p,q)=0 \Leftrightarrow p=q$  identity
3.  $D(p,q)=D(q,p)$  symmetry
4.  $D(p,z) \leq D(p,q)+D(q,z)$  triangle inequality

Konkrétnie metriky (vzdialenosťné funkcie):

- *Euclidean distance*:  $\sqrt{dx^2 + dy^2}$
- *City-block distance*:  $\text{abs}(dx) + \text{abs}(dy)$
- *Chess-board distance*:  $\max(\text{abs}(dx), \text{abs}(dy))$

Image	Distance transform

Image	Distance transform

Image	Distance transform

## Morfologické spracovanie obrazu

Morfologická operácia: mení tvar objektov na obraze

Operátor: structuring element (kernel)

Vstupy: binárny, gray-scale, farebný obraz

Flat structuring element: matica (kernel) jednotiek a nul ľubovoľného tvaru a veľkosti.

### Základné morfologické operácie:

- **rozšírenie** (dilation) - pridáva biele pixely na okraje objektov
- **erózia** (erosion) - odoberá biele pixely z okrajov objektov

Počet ovplyvnených pixelov závisí od veľkosti matice (structuring element).

Morfologické operácie sú typicky aplikované na binárne obrázky, ale používajú sa aj na gray-scale.

**Rozšírenie**: výstupom je maximum z hodnôt pixelov, ktoré sa prekrývajú s jednotkami matice.  $F \oplus SE$

Input binary image



Image after dilation



**Erózia**: výstupom je minimum z hodnôt pixelov, ktoré sa prekrývajú s jednotkami matice.  $F \ominus SE$

Input binary image



Image after erosion

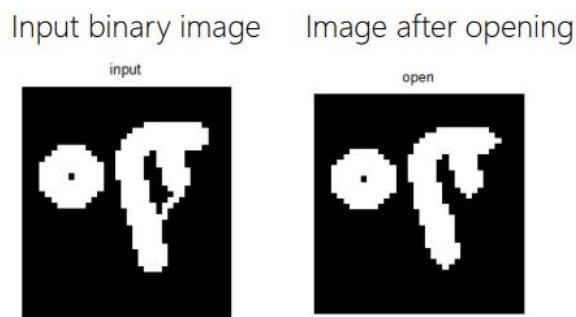


Rozšírenie a erózia NIE SÚ navzájom inverzné operácie, takže postupným aplikovaním jednej a druhej operácie NEDOSTANEME pôvodný obraz. Toto sa dá však využiť na nové operácie.

Kombinovaním rozšírenia a erózie vznikajú **nové morfologické operácie**:

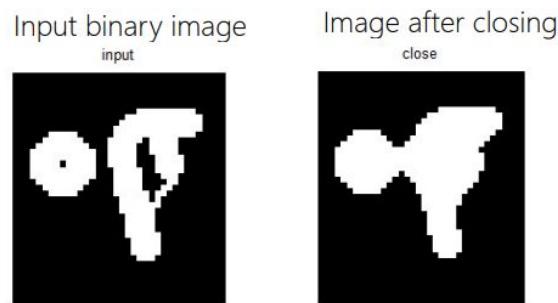
- **otváranie** (opening): erózia nasledovaná rozšírením
- **zatváranie** (closing): rozšírenie nasledované eróziou

**Otváranie:** tenké čiary a malé bodky (tenšie/menšie ako matica) zanikajú, tvary sa "otvárajú". 1. Erózia, 2. diletácia ( $F \ominus S$ )  $\oplus S$



**Zatváranie:** malé diery zanikajú, objekty blízko seba sa spájajú, tvary sa "zatvárajú".

1. Diletácia, 2. erózia ( $F \oplus S$ )  $\ominus S$



### Distance transform

- každý nenulový pixel sa nahradí hodnotou vzdialosti k najbližšiemu nulovému pixelu



### Skeletonization:

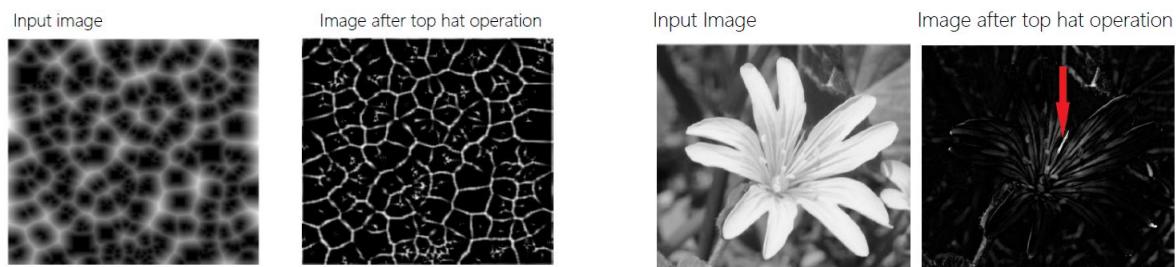
- extrakcia kostry z obrazu
- redukcia objektov na čiary bez zmeny štruktúry



- algoritmy:
  - maximal circles
  - sekvenčné stenšovanie
  - distance transform

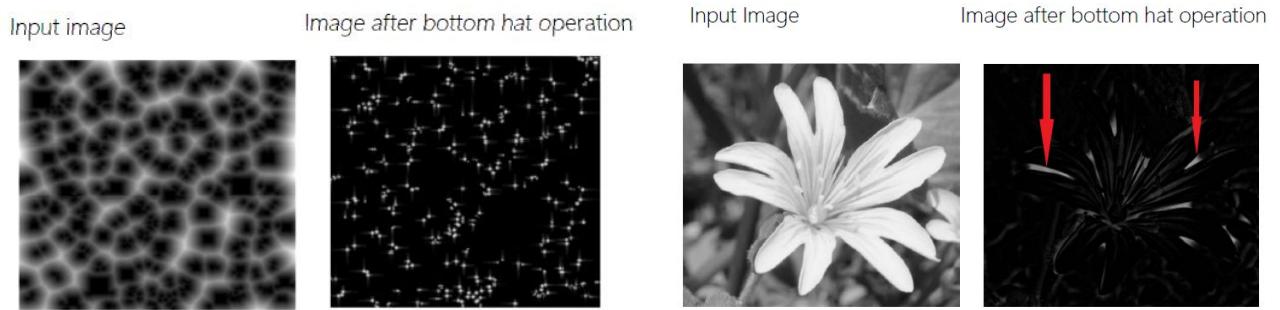
### Top hat operation: original image MINUS its opening

$$F - ((F \ominus S) \oplus S)$$



### Bottom hat operation: image closing MINUS original image

$$((F \oplus S) \ominus S) - F$$

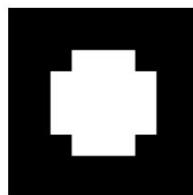


**Morfologický gradient:** rozšírenie MÍNUS erózia (s tým istým štruktúrnym elementom)

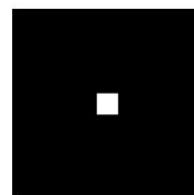
- použitie: detekcia hrán



Input image



Dilation



Erosion

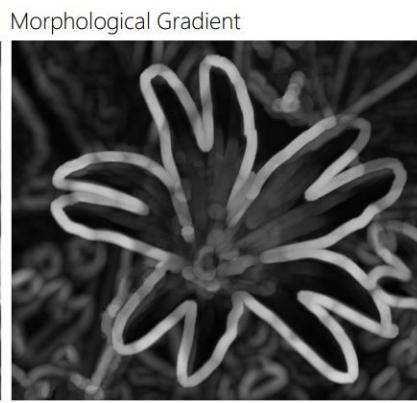
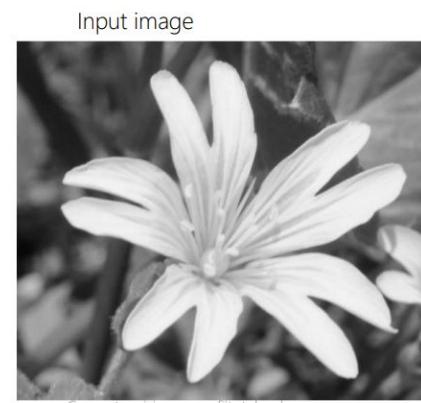


Subtraction

Str. element:  

$$\begin{matrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{matrix}$$

$$g = (f \oplus b) - (f \ominus b)$$



### Detekcia hrán (boundary extraction):

- rozdiel originálneho obrázku a jeho erózie (alebo rozšírenia)
- hrúbka hrán závisí od veľkosti matice pre eróziu

### Hit-or-Miss operation:

- máme dve rôzne matice (structuring elements - SE), ktoré sa neprekryvajú
- zachováva tie pixely, ktorého susedstvá zodpovedajú tvaru SE1 a nezodpovedajú tvaru SE2
- výstupom je **priek** erózie obrázku s jednou maticou a erózie invertovaného obrázku s druhou maticou
- vhodné na hľadanie/označenie znakov v texte

$$HM = (F \ominus SEhm1) \cap (F^c \ominus SEhm2).$$

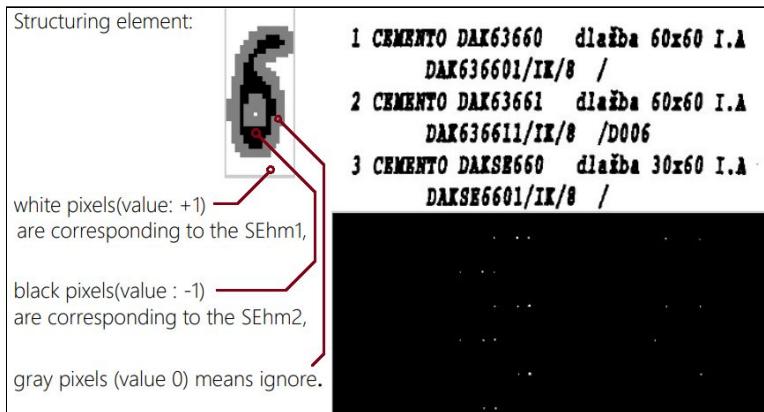
We will develop morphological image processing techniques that use combinations and sequences of operations. These techniques are useful in a wide range of applications.

Input image



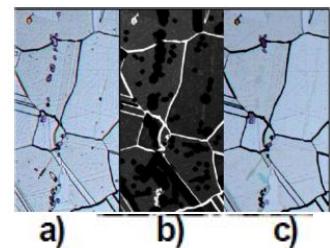
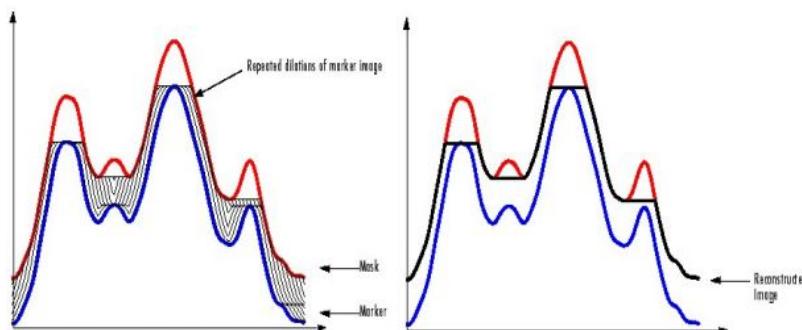
Result image

Structuring elements:



### Morfologická rekonštrukcia:

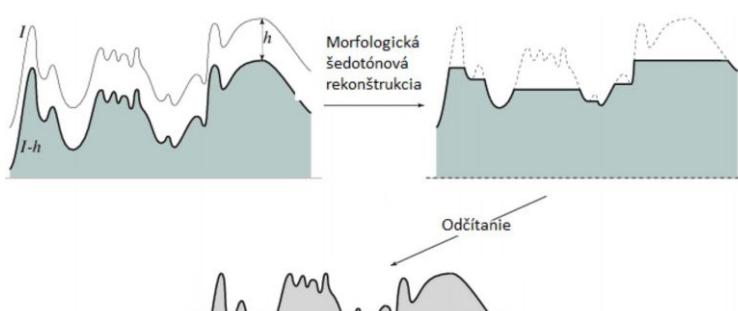
- opakované rozširovanie (dilation) obrázku, kým nedosiahneme požadovaný obraz (masku)
- špičky sa tým odstránia



a) original image,  
b) marker image  
c) reconstructed image

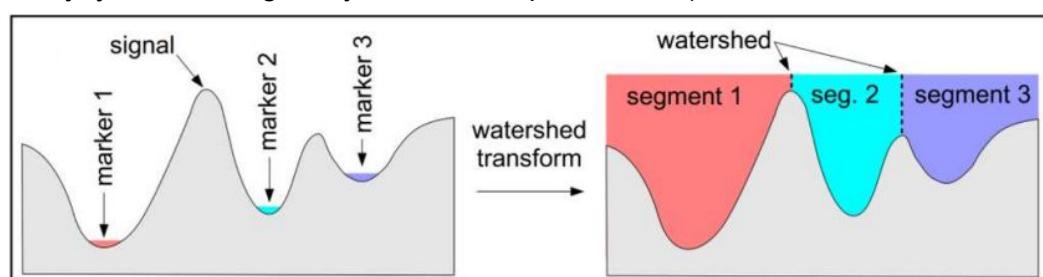
### Extrakcia lokálnych maxím:

- od obrazu odčítame jeho morfologickú rekonštrukciu

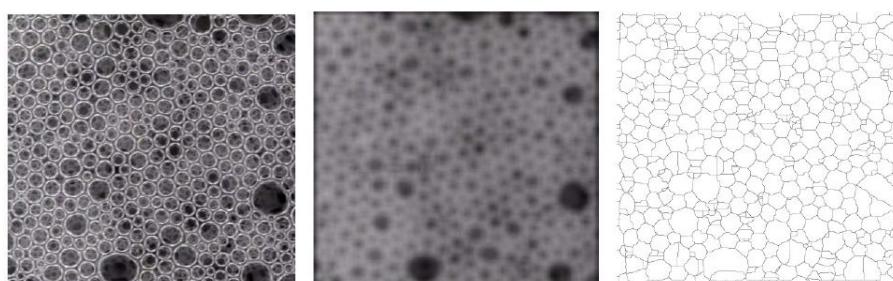


### Watershed transformation:

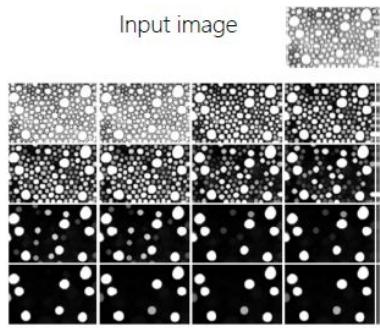
- chápame gray-scale obrazu ako výškovej mapy (najsvetlejšie pixely - vyššie body oddelujú jednotlivé segmenty, ktoré sa "naplnia vodou")



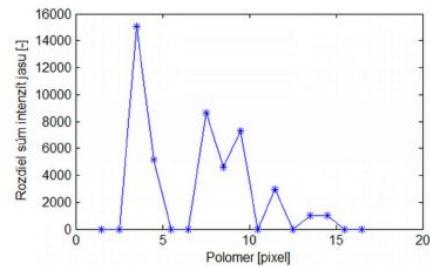
- pri zašumených obrazoch treba najskôr spraviť Gaussian filter



## Granulometria:



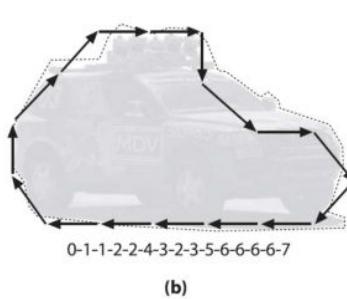
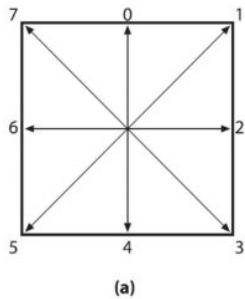
Morphological opening with gradually varying size of structural elements.



Dependence of the difference of brightness before and after the opening operation by using the morphological structural element of circular shape for varying radius.

## Hľadanie obrysov:

- vhodné pre analýzu tvarov a detekciu a rozpoznanie objektov
- obrys** (kontúra) je zoznam bodov reprezentujúcich krvky v obrázku
- opencv *findContours()* môže vytvoriť zoznam alebo strom
- Freemanov reťazový kód:



a: Freeman chain moves are numbered 0–7;

b: contour converted to a Freeman chain-code representation starting from the back bumper

## [3. prednáška]

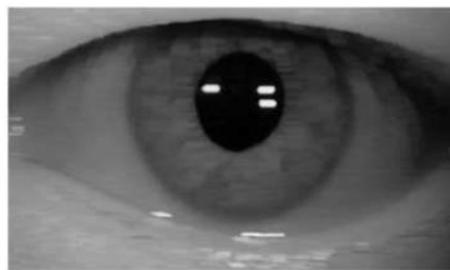
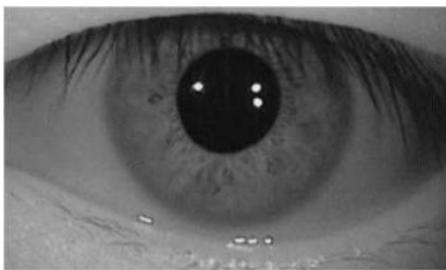
### Filtrácia:

- Lineárna
  - v priestorovej doméne - konvolúcia (prechod obrázka kernelom), manipulujeme pixely obrázka
  - vo frekvenčnej doméne - transformácia (kosínová/Fourierová), modifikujeme samotnú transformáciu obrázka (prečo? - býva rýchlejšia ako filtrácia v priestorovej doméne. Je jednoduchšie zostaviť vhodný filter; napr. ideálny HP/LP filter)
- Iné filtračné metódy

## Iné filtračné metódy:

### 2-D order-statistic filtering

- je zoradujúci filter - zoradí hodnoty susedstva podľa numerického poradia -> nahradí každý prvok vo výstupnom obrazu prvkom podľa poradia.
- (napr. odstránenie mihalníc z obrázka)



- špeciálne prípady - minimum/maximum/medián filter.

*Medián* - nelineárna digitálna technika na odstraňovanie šumu z obrázka/signálu.  
Ako novú hodnotu pixelu vyberie medián zo zoradených hodnot susedstva -> dobrý na odstránenie Salt & Pepper šumu.



## Lineárna filtrácia:

**Diskrétna konvolúcia** - zapisuje sa ako \*

- suma všetkých násobení susedstva kernelom (násobenie po prvkoch)
- komutatívna  $f^*g = g^*f$  a asociatívna operácia  $f^*(g_1^*g_2) = (f^*g_1)^*g_2$

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

- vysporiadanie sa s okrajmi obrázka:
  - nula
  - konštanta
  - zrkadlovo

- zmenšenie obrázka - obrázok po konvolúcii bude menší ako pôvodný

### Konvolúcia - smooth filters

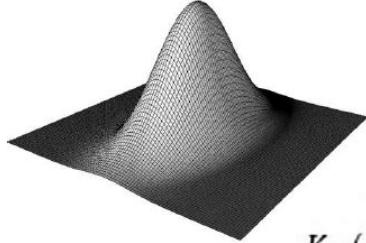
- Mean (blur) filter** – priemer (strácajú detaily)
- Gauss filter** – rozmazáva

Convolution kernel 3x3:

1	1	1
1	1	1
1	1	1

1/9      sa

Gauss convolution kernel 5x5 (integer):  $\frac{1}{273}$

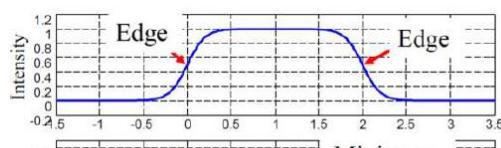


$$K_\sigma(x, y) := \frac{1}{2\pi\sigma^2} \cdot \exp\left(\frac{-x^2 - y^2}{2\sigma^2}\right)$$

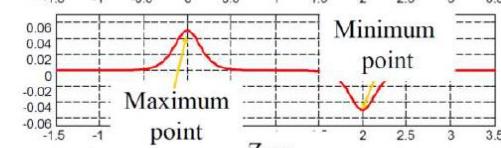


### Edge filters

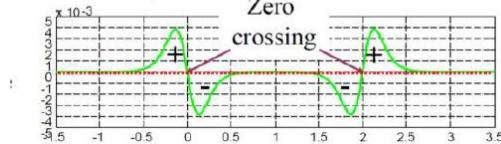
Gray level profile line



The 1-st derivative



The 2-nd derivative



## Prewitt filter

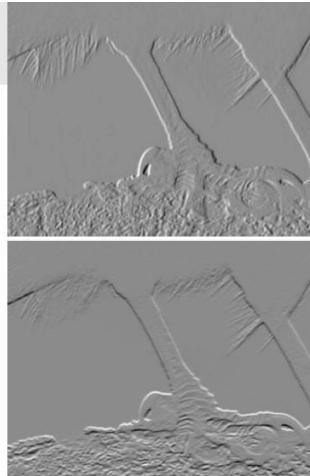
Gradient edge filter (1-st deriv.)

Vertical convolution kernel:

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$$

Horizontal convolution kernel:

$$\begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}$$



## Sobel filter

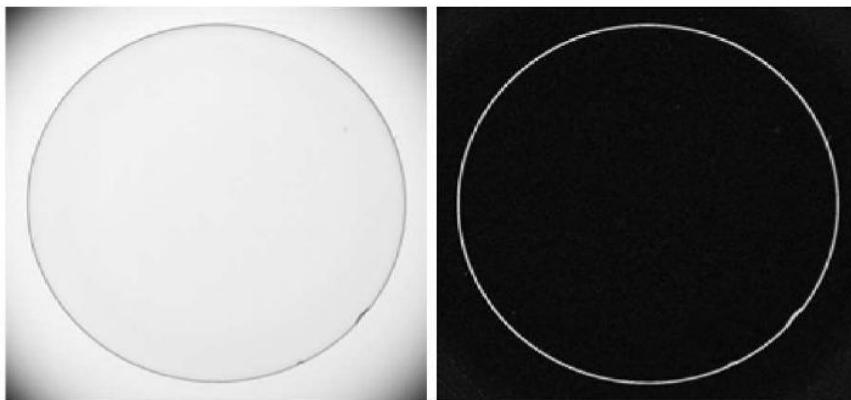
Gradient edge filter (1-st deriv.)



Sobel convolution kernels:

$$\mathbf{G}_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * \mathbf{A} \quad \text{and} \quad \mathbf{G}_x = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} * \mathbf{A}$$

- **Sobel gradient** – vylepšenie hrán použitím prvých derivátov (odstránenie drobných chýb)



## Roberts filter

Gradient edge filter (1-st deriv.)

Roberts convolution kernels:

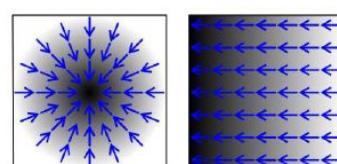
$$h_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad h_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

- **Image gradient** – ráta vektor gradientu intenzity

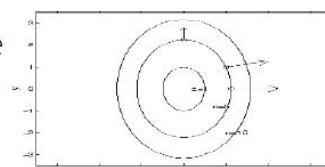
v každom bode obrázka pomocou Prewitt (Sobel) filteru

$$\nabla I = \left( \frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right)$$

image gradient  
(example):



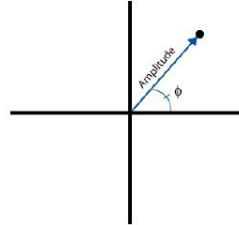
edge filtered image  
(example):



Magnitude       $\mathbf{G} = \sqrt{\mathbf{G}_x^2 + \mathbf{G}_y^2},$

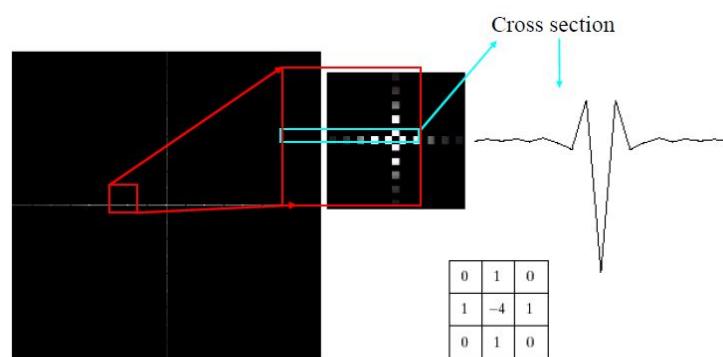
(amplitude)

angle             $\Theta = \arctan(\mathbf{G}_y, \mathbf{G}_x)$



- **Laplacian filter** – detekcia hrán (2. Derivácia). Pre obrázok  $f(x,y)$  je definovaný:

$$\nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$



Laplacian mask in Chapter 3

- kernel:

Mask includes/or not includes also the diagonal neighbors.

0	1	0	1	1	1
1	-4	1	1	-8	1
0	1	0	1	1	1
0	-1	0	-1	-1	-1
-1	4	-1	-1	8	-1
0	-1	0	-1	-1	-1

convolutional kernels 3x3:

convolutional kernel 5x5:

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 1 & 2 & -16 & 2 & 1 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

1. Zaostrenie
2. Obnovenie obrázka (Image restoration)
3. Redukcia šumu

• Vylepšenie  
obrázka (Image  
Enhancement)

Image  
preprocessing

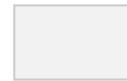
Computer vision  
application

- cieľom je vylepšiť pôvodný obrázok :D

- Zaostrenie
  - použitie konvolúcie (priestorová doména)
  - zvýraznenie jemných detailov/hrán

- zaostrenie rozmazaného obrázka
  1. spôsob: *originál – rozmazaný originál = maska*, ktorá sa pridá do originálneho obrázka
  2. spôsob: použitie Laplacovho obrázka – *sčítaj/odčítaj Laplaca od originálneho obrázka.*

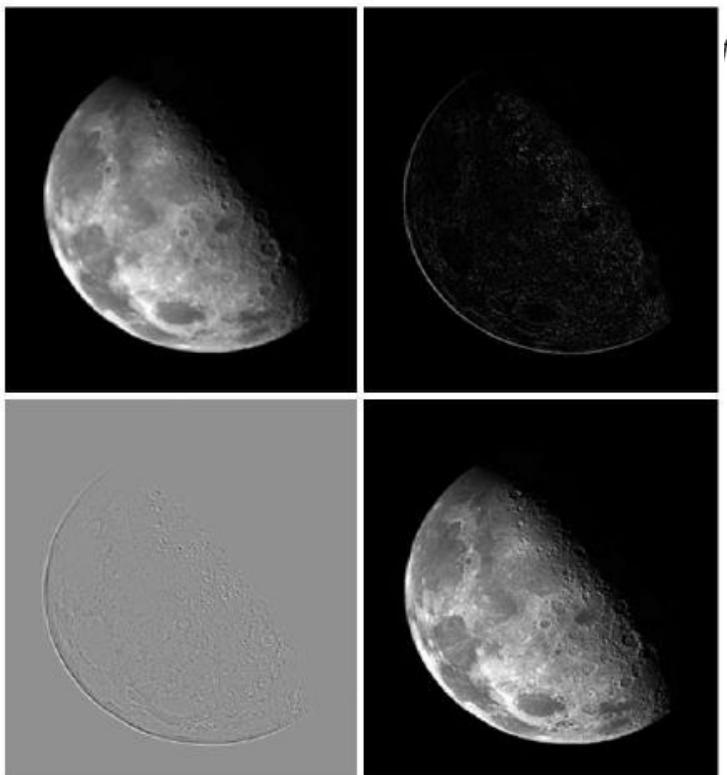
Sčítanie – ak má kernel kladný centrálny koeficient.  
 Odčítanie – ak má kernel záporný centrálny koeficient.



0	-1	0
-1	4	-1
0	-1	0

$$\left( f(x, y) - \nabla^2 f(x, y) \right)$$

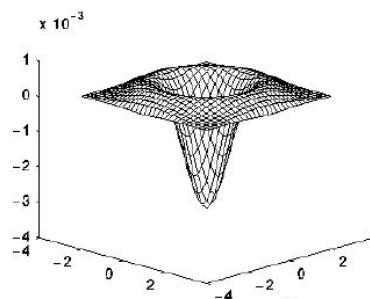
0	1	0
1	-4	1
0	1	0



### Edge detection

- **Laplacian of Gaussian – LoG**
  - Na nájdenie hrán
  - LoG = Gaussian smoothing + Laplacian detection
  - Laplacian kernaly aproximujú druhú deriváciu merania -> sensitívne na šum, preto sa najprv rozmaže obrázok Gaussom
  - Konvolúcia je asociatívna -> vykonáme najprv konvolúciu Gaussového filtrov Laplacovým filtrom a výsledným hybridným filtrom konvolujeme obrázok. (jeden prechod obrázka namiesto dvoch)

0	-1	0	-1	-1	-1
-1	4	-1	-1	8	-1
0	-1	0	-1	-1	-1



The 2-D Laplacian of Gaussian (LoG) function.

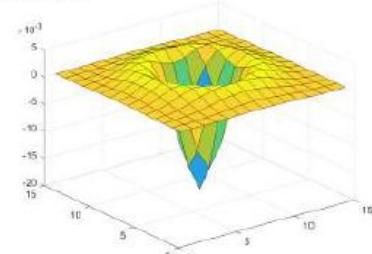
0	1	1	2	2	2	1	1	0
1	2	4	5	5	5	4	2	1
1	4	5	3	0	3	5	4	1
2	5	3	-12	-24	-12	3	5	2
2	5	0	-24	-40	-24	0	5	2
2	5	3	-12	-24	-12	3	5	2
1	4	5	3	0	3	5	4	1
1	2	4	5	5	5	4	2	1
0	1	1	2	2	2	1	1	0

discrete kernel that approximates  
this function  
(for a Gaussian  $\sigma = 1.4$ )



Input image

LoG Filter Kernel  
Size = 15,  
Sigma=2



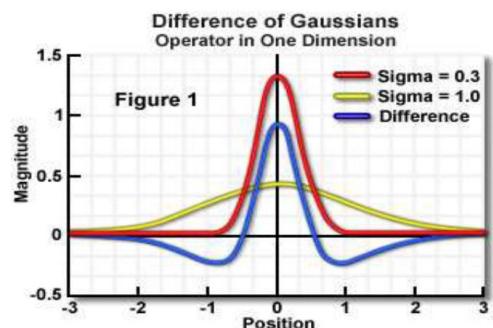
LoG filtered Image

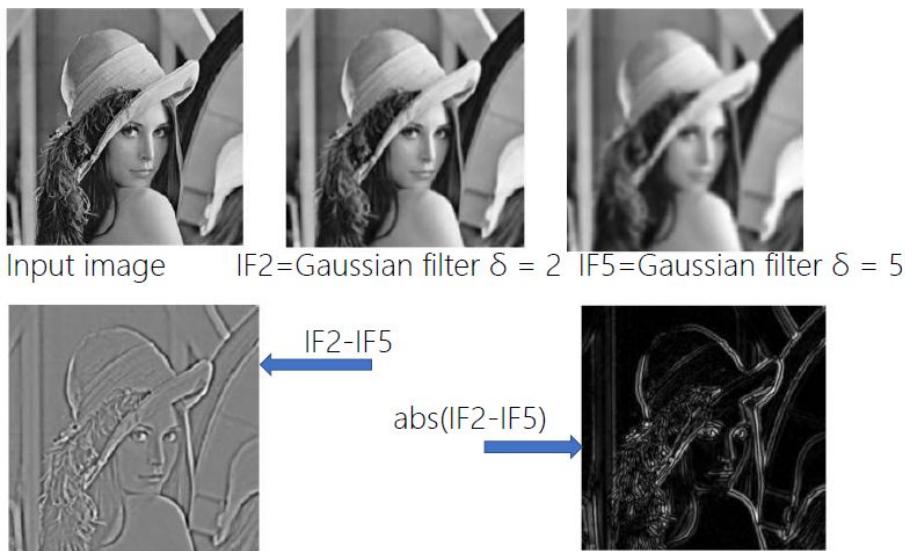


abs(LoG filtered Image)

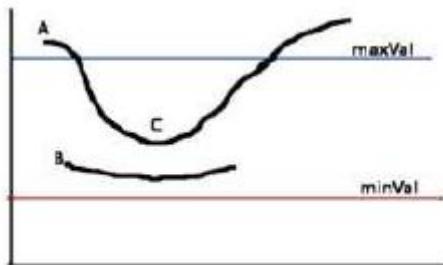
- **Difference of Gaussians – DoG**

- Rozmazanie pôvodného obrázka Gausom s dvoma rôznymi sigmami  $\rightarrow$  absolútne hodnoty odčítania





- [Canny edge detection - algoritmus](#)
  1. Redukcia šumu Gaussom
  2. Nájdenie gradientu intenzity obrázka
    - vyhladený obraz sa filtruje pomocou Sobelovho kernela v horizontálnom aj vertikálnom smere (vznikne prvá derivácia v oboch smeroch) -> gradient hrán a smer pre každý pixel.
  3. Non-maximum Suppression – eliminácia bodov neležiacich na podstatných hranách. -> Vyrávanie magnitúdy hrany.
  4. Hysteresis Thresholding – odstránenie hrán na základe prahu –  $\text{minVal}$ ,  $\text{maxVal}$ 
    - ak je gradient intenzity pod  $\text{minVal}$  – nie je to hrana -> odstrániť
    - ak je nad  $\text{maxVal}$  – je to hrana nechat'
    - v intervale ( $\text{minVal}$ ,  $\text{maxVal}$ ) sa posudzuje podľa konektivity. Ak sú spojené s „istými hranami“ je to hrana a sú ponechané, inak sú zahodené



5. opakuj kroky 1-4 pre vzostupné hodnoty štandardnej derivácie sigmy  $\sigma$
6. zhrň informácie o hránach použitím ‚feature synthesis‘ prístupu.

**Nekorelovaný šum** – náhodná variácia graylevelu v obraze.

Klasifikujeme podľa :

- tvaru funkcie hustoty pravdepodobnosti ([pdf](#)) – analogický prípad šumu
- tvaru histogramovej funkcie – diskrétny prípad šumu

## Noise models:

- uniformný
  - opisujeme histogramom
  - kvantizačný šum má skoro uniformné rozdelenie
  - Použitie: generovanie iného šumu, dogabanie obrázka pre vyhodnotenie algoritmov obnovenia obrazu.

Original image  
and histogram:

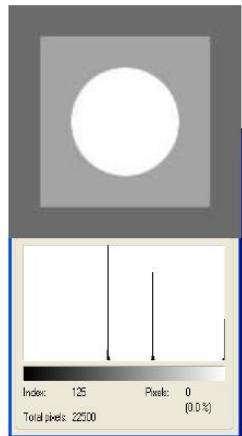
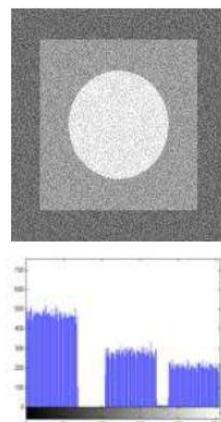


Image disturbed by uniform  
noise and histogram:



- gaussov (normálny)

- má pdf s normálnym rozložením
- je obrovskou časťou konštantného levelu šumu v tmavých oblastiach obrázka

Original image  
and histogram:

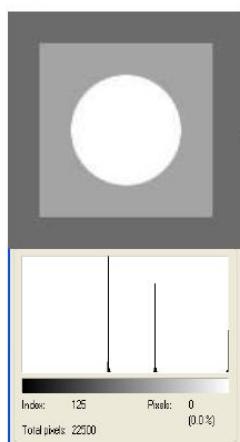
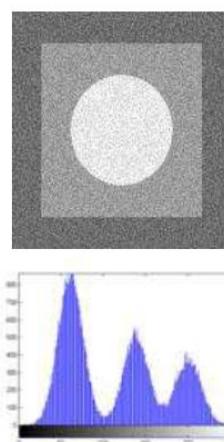


Image disturbed by Gaussian  
noise and histogram:



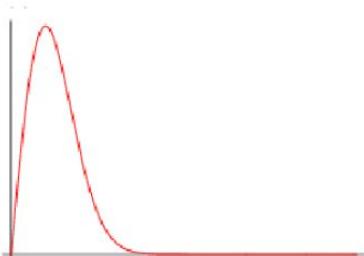
- salt & pepper

- hodnoty 0 a 255 (pre 8bit obrázok)
- impulse/spike noise
- zapríčinujú ho nefunkčné pixely v kamerových senzoroch, chybné miesta v pamäti, chyby v procese digitalizácie

- rayleigh distribúcia – šum v radarových snímkach

PDF (Probability density function)

Rayleigh distribution is defined as:



$$HISTOGRAM_{Rayleigh} = \frac{2g}{\alpha} e^{-\frac{g^2}{\alpha}}$$

$$\text{where : mean} = \sqrt{\frac{\pi \alpha}{4}}$$

$$\text{variance} = \frac{\alpha(4 - \pi)}{4}$$

- gamma noise

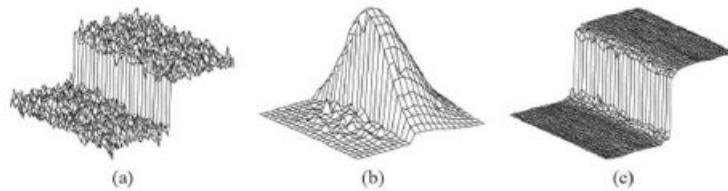
Šum – aditívna zložka:

Model degradovaného obrazu pomocou aditívneho šumu

degraded image = original image + additive noise function

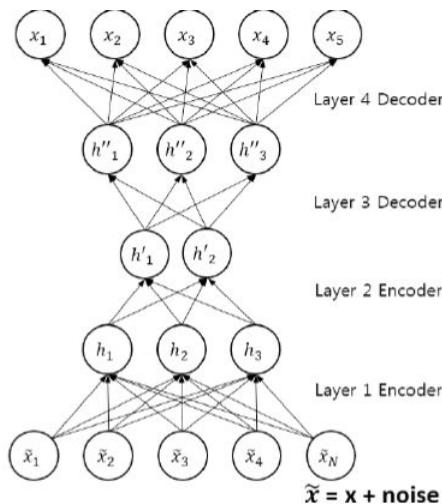
Redukcia šumu:

- filtrácia v priestorovej/frekvenčnej doméne
- na niektoré konkrétné šumy sú konkrétné filtre (salt&pepper -> median)
- dve kategórie filtrov na redukciu šumu v priestorovej doméne
  1. **Zoraďujúce filtre** - Order filters (median) – výber vhodnej hodnoty pixela z usporiadанého susedstva pixelov. Dobré na salt-and-pepper, negative exponential, a Rayleigh šum.
  2. **Lineárne filtre** (LP filter) – Dobré na Gaussian/uniform šum. Nevýhodou je rozmazenie hrán a detailov pri použití low-pass filtra.
- V noiseless obrázkoch -> energia priestorovej frekvencie je koncentrovaná v nízkych frekvenciach. Pri zašumenom obrázku je teda väčšina vysokofrekvenčného obsahu obrázka = šum.
- Periodický šum – spôsobený napr. Vibráciami (elektrické/mechanické systémy). Vo frekvenčnej oblasti sa javí ako sinusovka. => riešenie band reject & notch filters
- Odhad šumu – chceme osamostatniť šum a z histogramu vytvoriť model šumu.
- Bilaterálny filter – tvar kernelu závisí od obsahu obrázka.
  - Nerobíme priemer cez hrany = bilaterálne hrany
  - Bilaterálne filtrovanie = kombinácia doménového a rozsahového filtrovania. Uvažujeme len pixely blízke v priestore a rozsahu.



- (a) A 100-gray-level step perturbed by Gaussian noise with  $\sigma=10$  gray levels.
- (b) Combined **similarity weights** for a  $c(\xi, x)$ ,  $s(f(\xi), f(x))$  for a  $23 \times 23$  neighbourhood centered two pixels to the right of the step in (a). The range component effectively suppresses the pixels on the dark side.
- (c) The step in (a) after bilateral filtering

- Zbavenie sa šumu využitím hlbokých neurónových sietí

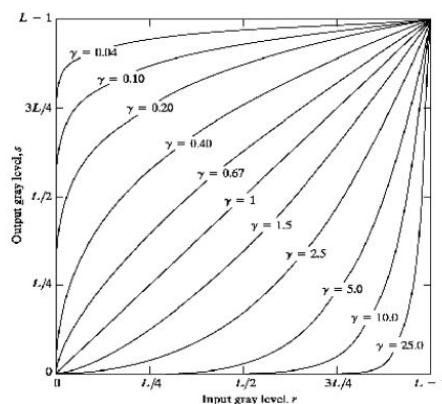


## Priama manipulácia intenzity pixelov

- Power-Law Transformations

- $c$  a gamma sú kladné konštanty

$$s = cr^\gamma$$



Plots of the equation

$$s = cr^\gamma$$

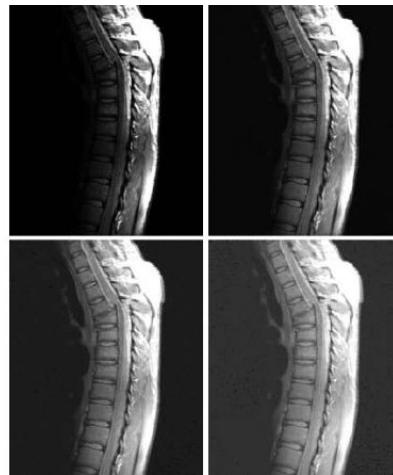
various values of  $\gamma$

Signal range: 0-1

- použitie Power-Law T. Sa nazýva gamma korekcia

(a) Magnetic resonance (MR) image of a fractured human spine.

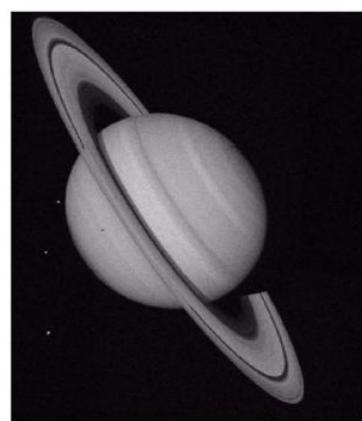
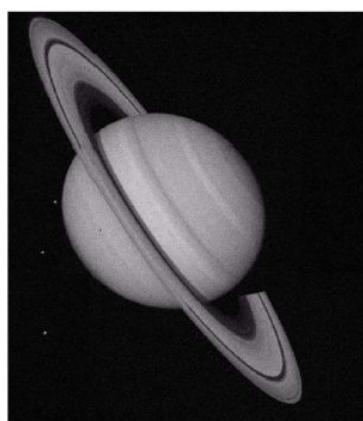
(b)–(d) Results of applying the power-law transformation with  $c=1$  and  $g=0.6, 0.4,$  and  $0.3,$  respectively.



Adaptívna filtrácia – aplikuje lineárny filter adaptívne -> prispôsobuje sa vlastnostiam obrazu.

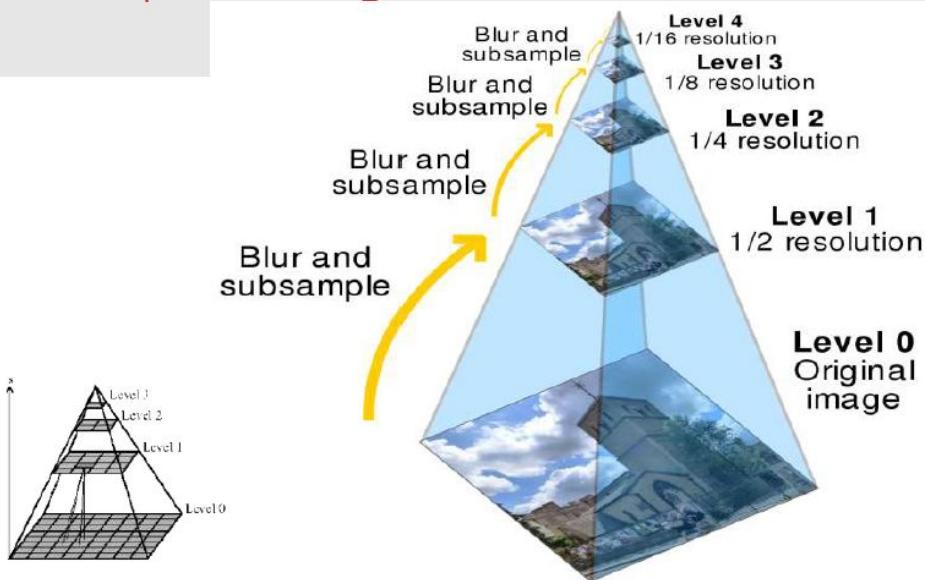
- Lepšie výsledky ako lineárne filtrovanie
- Zachováva hrany a detaľy (vysoké frekvencie)

The example below applies adaptive filtration using the Wiener filter to an image of Saturn that has had Gaussian noise added.



Vylepšovacie priestorové metódy možno kombinovať a dosiahnuť výborné výsledky! Viď slajdy 89-92

# Pyramid processing



## Zaostrenie obrazu

- rozmazaný - ak bol obrázok zachytený pri pohybe kamery/subjektu, nezaostrená optika...

A blurred or degraded image can be approximately described by this equation

$$g = H*f + n, \text{ where}$$

g The blurred image

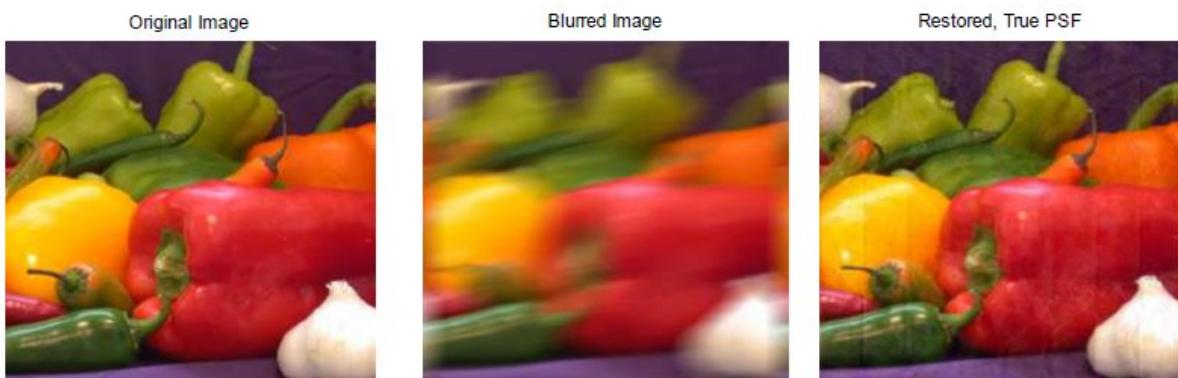
H The distortion operator- point spread function (PSF)

The distortion operator, when convolved with the image, creates the distortion.

f The original true image

n Additive noise, introduced during image acquisition, that corrupts the image

- Zaostrenie pomocou dekonvolúcie (reversing effect of convolution) – hlavná úloha -> dekonvolúcia rozmazaného obrázka s PSF, ktorá presne opisuje skreslenie.
  - Kvalita výsledného obrázka závisí od PSF!!



## [4. prednáška]

**Meranie farieb** - dôležité pre mnohé aplikácie / odvetvia:

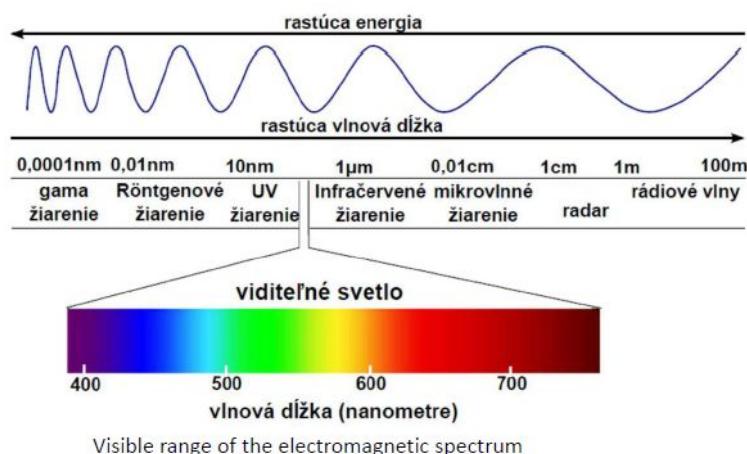
- textilný priemysel
- papierový priemysel
- kožený priemysel...

**Rádiometria** - Rádiometria je veda a technológia merania elektromagnetického žiarenia v celom spektri elektromagnetických vln.

**Fotometria** - Fotometria je veda o meraní svetla z hľadiska jasu vnímaného ľudským okom.

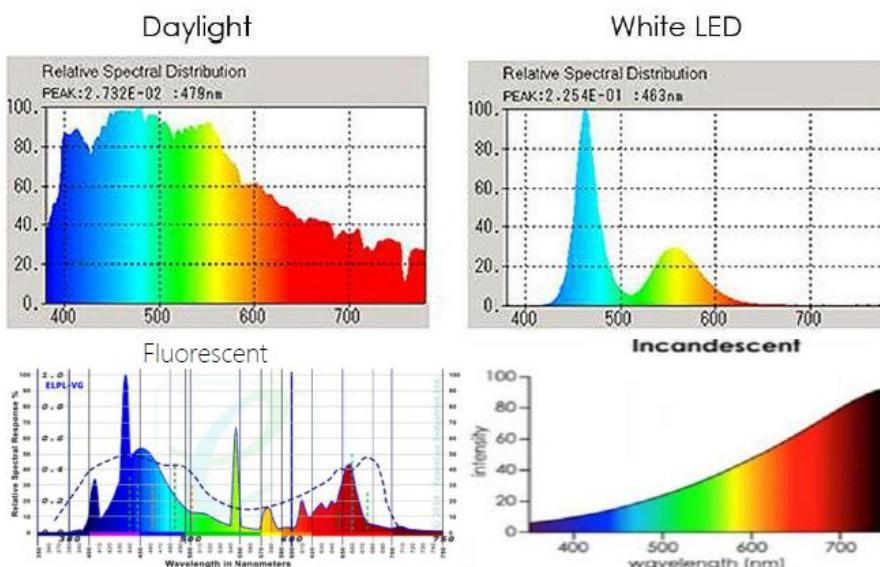
### Elektromagnetické spektrum

- s rastúcou vlnovou dĺžkou klesá energia
- viditeľné svetlo - 360-770nm

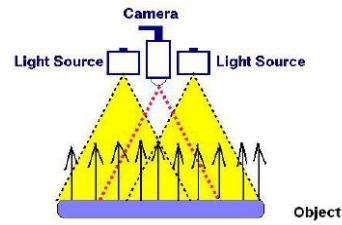


### Distribúcia spektrálneho výkonu (SPD)

- reprezentuje silu žiarenia vyžarovanú svetelným zdrojom pri každej vlnovej dĺžke.
- opisuje ako je distribuovaný výkon signálu podľa frekvencie.
- $S(\lambda)$
- SPD pre rôzne svetelné zdroje:



**Vnímanie a získavanie obrazu** - v závislosti od povahy zdroja sa energia osvetlenia odráža alebo prenáša cez objekty.



**Odrážanie sa od povrchu** - vidíme konkrétnu farbu (absorpcia vlnovej dĺžky pigmentovými časticami)

- farba povrchu závisí od spektra odrazivosti (reflectance spectrum  $r(\lambda)$ ) a spektra osvetlenia (illuminance spectrum  $i(\lambda)$ )
- Odraz svetla od povrchu  $R(\lambda)$  - násobenie odrazu s osvetlením  
 $R(\lambda) = \text{illumination } i(\lambda) * \text{reflectance } r(\lambda)$   
 $R(\lambda) = i(\lambda) * r(\lambda)$
- (napr. ak červeným svetlom zasvetime na modrý objekt, neodráži sa nič, lebo spektrá sa neprekryvajú)
- Reakcia fotoreceptora  $q_i$ : integrál súčinu 3 spektier podľa vlnovej dĺžky, spektrá:

  - $I(\lambda)$ : spektrum osvetlenia
  - $R_i(\lambda)$ : spektrálna citlivosť senzora
  - $S(\lambda)$ : spektrum odrazivosti snímanej plochy

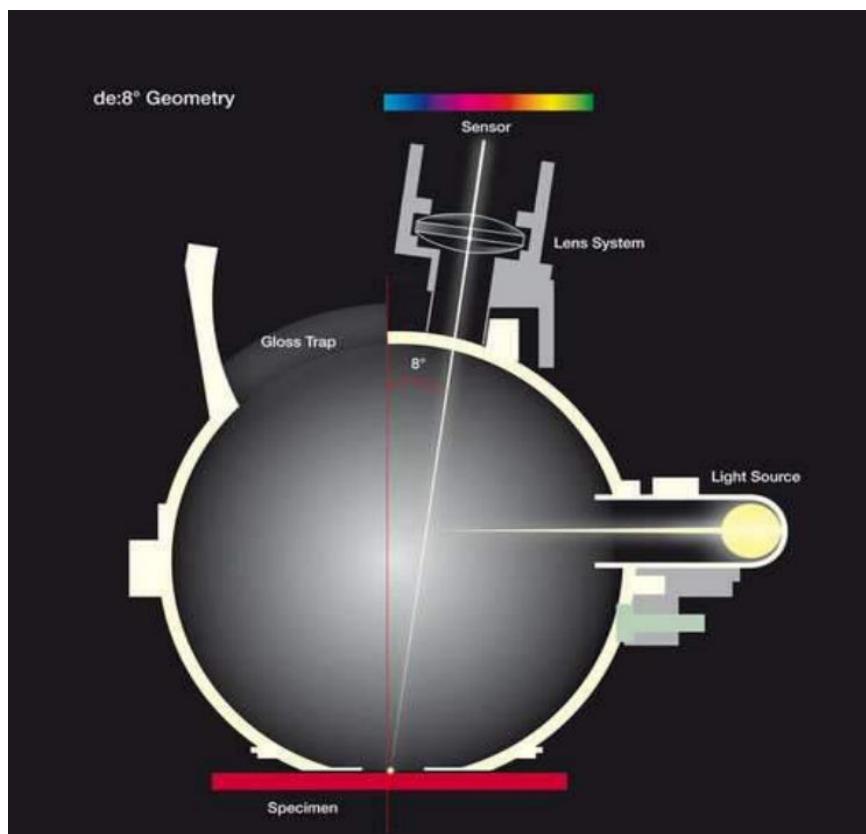
$$q_i = \int_{\lambda_1}^{\lambda_2} I(\lambda) R_i(\lambda) S(\lambda) d\lambda$$

#### Multispektrálne obrázky:

- väčšina senzorov (fotákov) nemá presnú informáciu o farebnom spektri
- výnimkou je **spektrometer**:
  - každé spektrálne pásmo (často 5 alebo 10 nm) je digitalizované nezávisle a uložené ako monochromatický obrázok
- využitie multispektrálnych snímok:
  - satelia
  - senzory
  - priemysel

#### Spektrometre:

- svetlo napr. z xenónovej lampy je premietané veľmi lesklú sféru (Ulbricht sphere), ktorá vytvorí takmer dokonalé difúzne svetlo
- senzor je hore, vychýlený napr. o  $8^\circ$  (pre spektrometer Minolta 2600d)
- zrkadlovo k senzoru môžu byť "dvierka" (gloss trap), ktorými sa dá zvoliť, či chceme do výsledku zahrnúť "specular component", teda lesk meraného materiálu

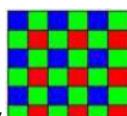


**CIE odporúčané osvetlenia** (uhol svetelného zdroja/uhol senzora,  $d=90^\circ + \text{sphere}$ ):

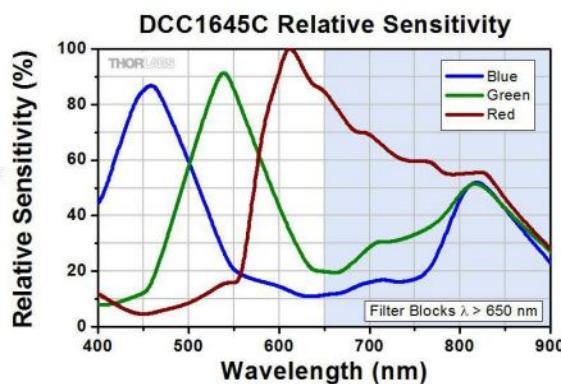
- 45/0
- 0/45
- d/0
- 0/d

**Bežné farebné fotoaparáty/kamery:**

- majú 3 snímače (RGB)

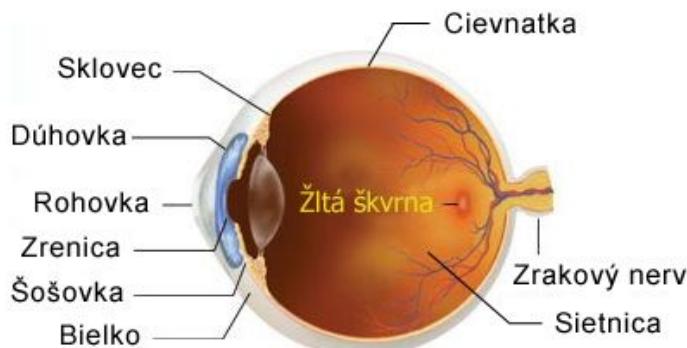


- bayerova matica pixelov
- alebo 3 senzory
- z dôvodu ceny sa preferuje bayerova matica
- spektrálna citlosť RGB kanálov:



## Vnímanie farieb človekom:

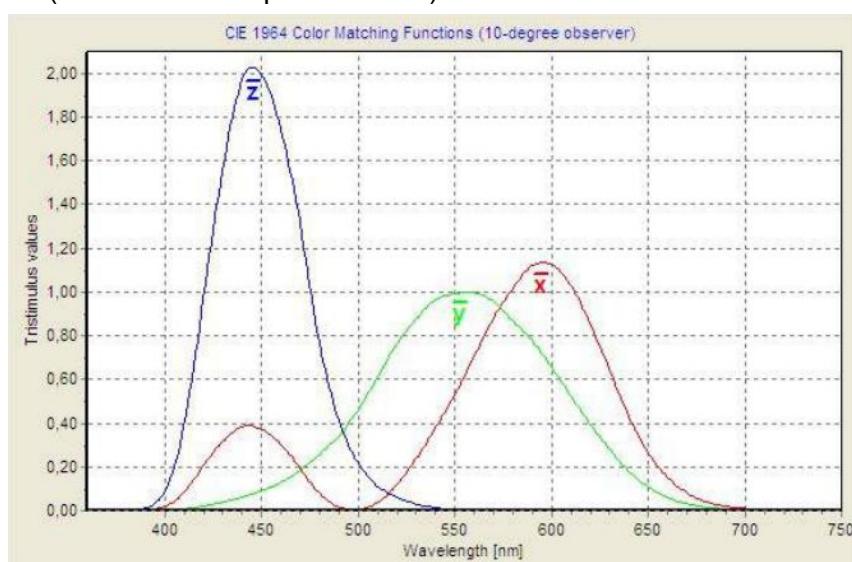
- okrem zdroja svetla a materiálu vnímaná farba závisí aj od:
  - receptorov,
  - spracovania mozgom,
  - a od prostredia.
- Človekom vnímaná farba má od fyzikálnych vlastností navyše subjektívnu vrstvu.
- Oko:



- čapíky sietnice ("color") - farebné videnie
- tyčinky sietnice - veľmi citlivé, čierno-biele videnie (v tme čapíky "nefungujú")
- tri-chromatickost':
  - bolo experimentálne dokázané, že stačí kombinovať 3 správne farby na vytvorenie ľubovoľnej výslednej farby

## CIE štandardy:

- CIE 1931: prvá definícia prepojenia viditeľného spektra s fyziologickým vnímaním farieb človekom
- Commission Internationale de l'Eclairage
- CIE kolorimetrický systém:
  - 1931 ( $2^\circ$ )
  - 1964 ( $10^\circ$  zo zorného poľa človeka)



- X, Y, Z zložky farby:

$$X_{10} = k_{10} \int_{\lambda} P_{\lambda} x_{10}(\lambda) d\lambda$$

$$Y_{10} = k_{10} \int_{\lambda} P_{\lambda} y_{10}(\lambda) d\lambda$$

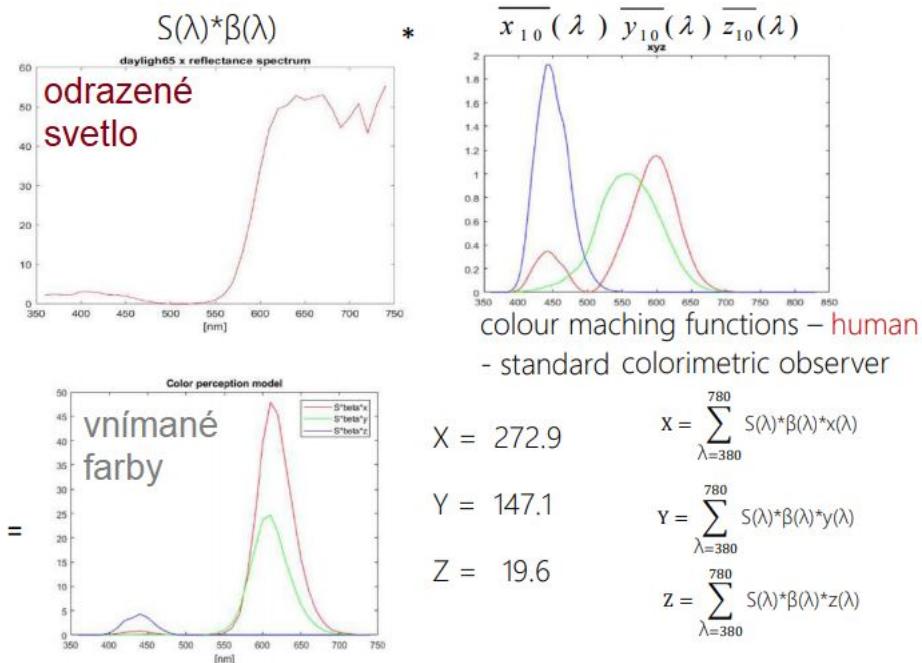
$$Z_{10} = k_{10} \int_{\lambda} P_{\lambda} z_{10}(\lambda) d\lambda$$

$$k_{10} = \frac{100}{\int_{\lambda} P_{\lambda} y_{10}(\lambda) d\lambda}$$

$P_{\lambda}$  denotes the monochromatic component of given color stimulus with radiant power distribution  $P_{\lambda} d\lambda$

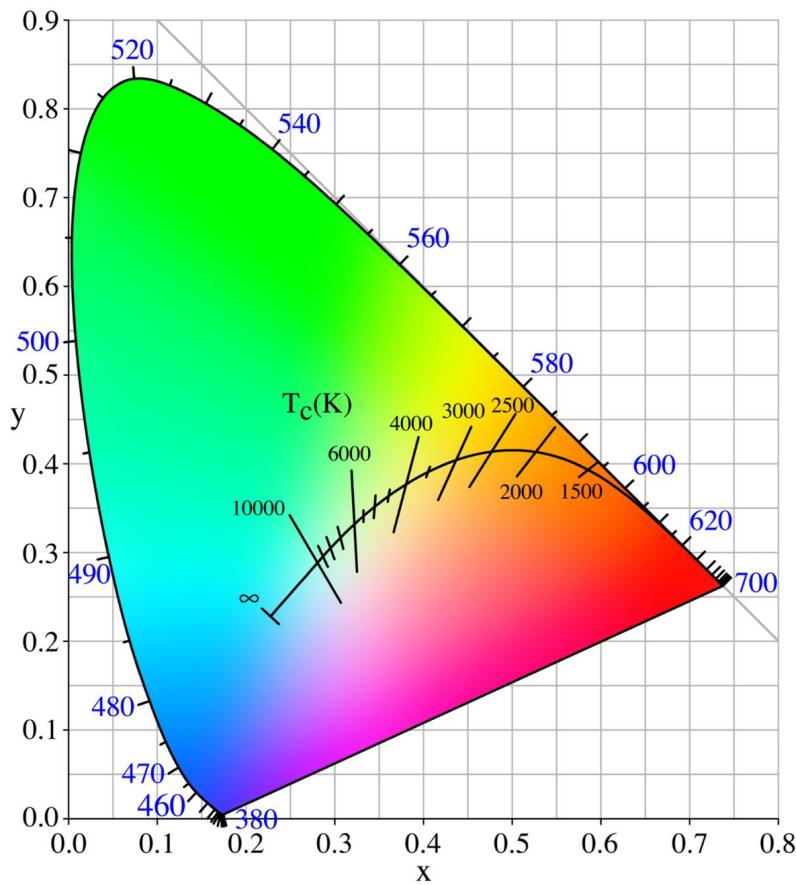
$x_{10}(\lambda)$ ,  $y_{10}(\lambda)$ ,  $z_{10}(\lambda)$  are color matching functions that define the CIE 1964 standard colorimetric observer

- vnímanie jednotlivých zložiek človekom:



- $S(\lambda)$ : spektrum zdroja svetla
- $\beta(\lambda)$ : spektrum odrazivosti sledovaného objektu

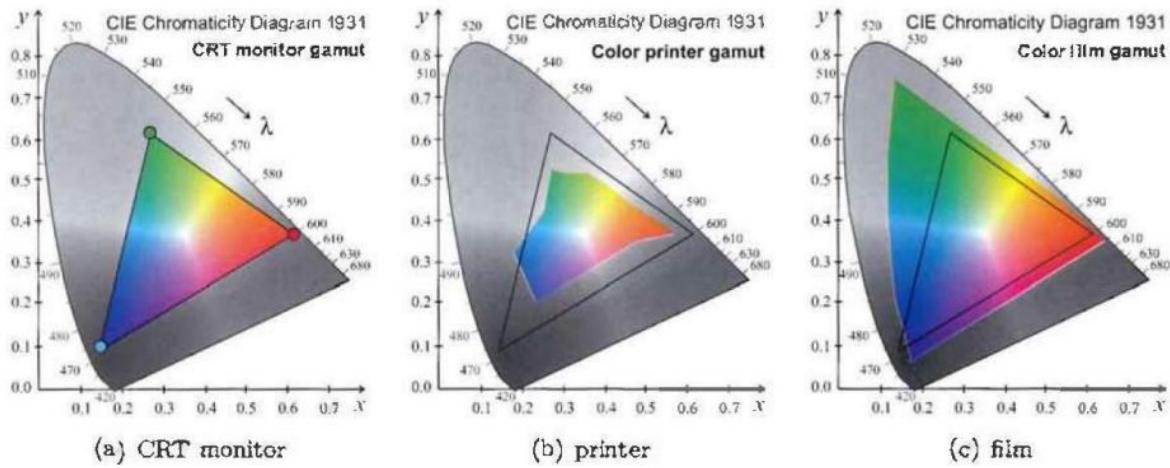
## CIE chromatický diagram



$$x_{10} = \frac{X_{10}}{X_{10} + Y_{10} + Z_{10}}$$

$$y_{10} = \frac{Y_{10}}{X_{10} + Y_{10} + Z_{10}}$$

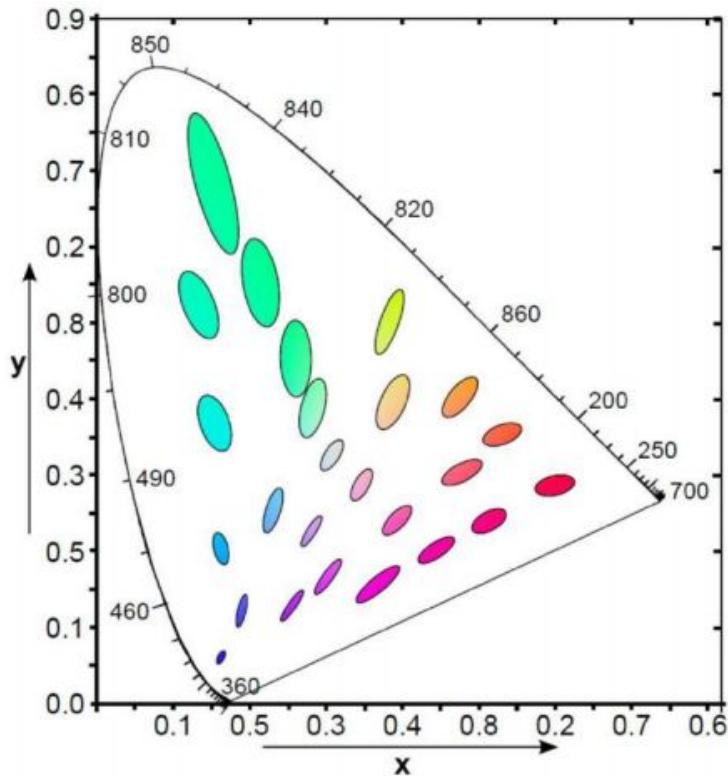
Kamery, obrazovky a tlačiarne dokážu zachytiť/zobraziť len určitú časť plochy CIE diagramu (napr. trojuholník v diagrame pre CRT monitor):



## Porovnanie farieb

MacAdamove elipsy:

- poukazujú na to, že citlivosť človeka na rôzne farby v rámci CIE diagramu nie je rovnomerná



- človek nedokáže rozlíšiť farbu v strede elipsy s farbou na kraji elipsy
- elipsy sú na diagrame 10x zväčšené kvôli prehľadnosti

### Rovnomerný farebný priestor

- CIE L\*a\*b
- prepočet z XYZ:

$$L^* = 116 \left( \frac{Y}{Y_n} \right)^{\frac{1}{3}} - 16 \quad \text{if} \quad \frac{Y}{Y_n} > 0.008856 \quad a^* = 500 \left[ \left( \frac{X}{X_n} \right)^{\frac{1}{3}} - \left( \frac{Y}{Y_n} \right)^{\frac{1}{3}} \right]$$

$$L^* = 903.3 \left( \frac{Y}{Y_n} \right)^{\frac{1}{3}} \quad \text{if} \quad \frac{Y}{Y_n} \leq 0.008856 \quad b^* = 200 \left[ \left( \frac{Y}{Y_n} \right)^{\frac{1}{3}} - \left( \frac{Z}{Z_n} \right)^{\frac{1}{3}} \right]$$

where  $[X_n, Y_n, Z_n]$  is the white reference point.

- rozdiel farby:

$$\Delta E^*_{ab} = \sqrt{[(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2]}$$

### Metamerizmus

2 farby s rôznymi spektrami môže človek vnímať ako rovnaké farby, pretože celkový vnem je zložený len z 3 farieb. Takáto zhoda nemusí zostať platná, keď sa zmení spektrum svetelného zdroja.

## Zdroje svetla

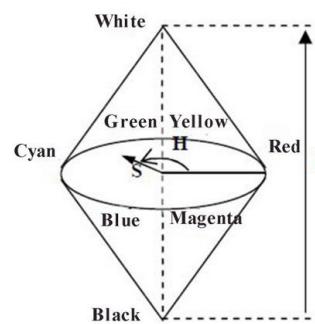
- A - typické volfrámové žiarovky (2856 K)
- D65 - denné svetlo (6500 K)
- F1-F6 - fluorescentné lampy

**Perfect reflecting diffuser** - CIE 1971 odporúča pre meranie, ale žiadny taký materiál neexistuje :D

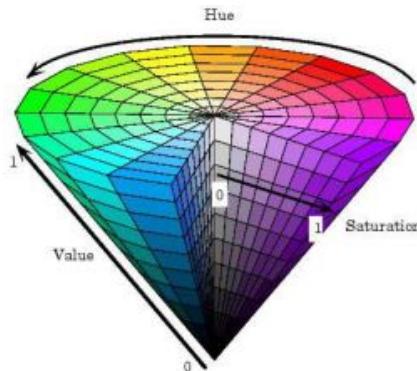
## Farebné priestory

- Aditívne miešanie (svetlo, **RGB**)
- Subtraktívne miešanie (pigment, **CMY**)
- Prevod medzi XYZ a RGB v podobe vektorov je možné urobiť prenásobením daného vektoru maticou  $3 \times 3$ . Jednotlivé RGB štandardy definujú hodnoty v tejto matici.
- YCbCr:
  - efektívnejšie uloženie farby
  - pre video encoding
  - $Y = 0.299(R-G) + G + 0.114(B-G)$  luma
  - $Cb = 0.564(B-Y)$  blue chroma
  - $Cr = 0.713(R-Y)$  red chroma

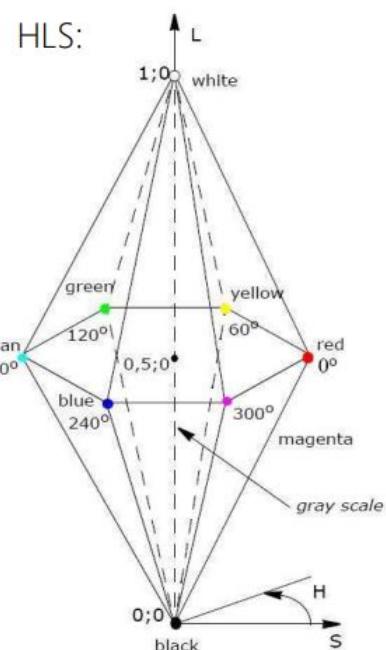
- HSI:
  - hue
  - saturation
  - intensity



- HSV:
  - hue
  - saturation
  - value



- HLS
  - hue
  - lightness
  - value



## CCT - correlated colour temperature

- akú teplotu v Kelvinoch by muselo mať teleso, aby vyžarovalo rovnakú farbu

Colour Temperatures in the Kelvin Scale

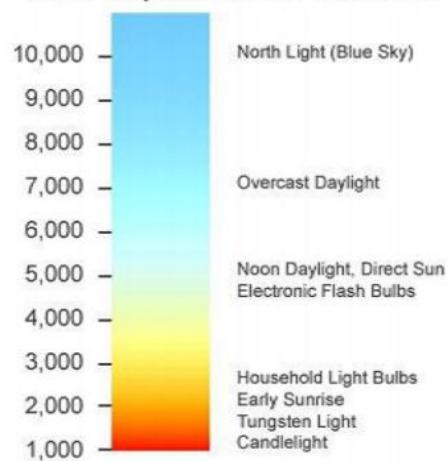
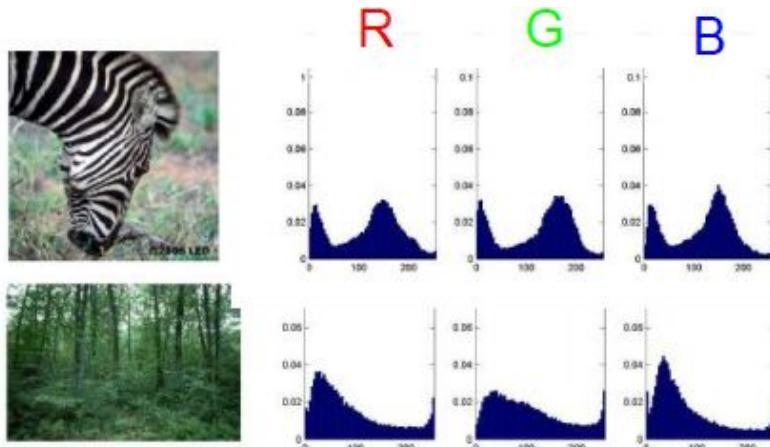


Image courtesy of [www.medicolege.com](http://www.medicolege.com)

## Farebné histogramy



- zachytávajú rozdelenie farieb
- sú nezávislé od umiestenia, veľkosti a otočenia objektov na obraze
- vhodné pre porovnávanie obrázkov:
  - majme databázu obrázkov s histogramom pre každý obrázok (stačí 1 farba)
  - extrahujme histogram pre nový obrázok
  - zoradíme obrázky z db podľa podobnosti histogramu

## Vizuálna pozornosť'

- výber: človek sa sústredí na určité objekty a ignoruje ostatné
  - hľadáme charakteristické črty, dôležité objekty (salient parts)
- limitácie: rýchlosť vnímania a spracovania informácií mozgom je obmedzená
- správanie sa oka:
  - fixácia (prijímanie informácií, 200 - 300 ms)
  - sakáda (rýchly pohyb, 30 - 80 ms)
  - sledovanie pohybujúceho sa objektu

## Pozornosť zdola nahor (bottom-up)

- mimovoľná nevedomá pozornosť podľa charakteriky scény
- automaticky pritahuje pozornosť

- saliency (charakteristický črt)
- bežne nás pritahuje farba, kontrast, orientácia, textúra a pohyb
- dáva nám len základné informácie o danom objekte

### **Pozornosť zhora nadol**

- cieľom riadená
- na základe predchádzajúcich znalostí
- pomalšia, dobrovoľná a vedomá (na rozdiel od poz. zdola nahor)
- objekt je rozpoznaný vďaka predošlým znalostiam

### **Eye tracking**

- sledovanie pohybu očí a kreslenie bodov/čiar do obrázku na miesta, kam sa pozorovateľ pozeral

### **Attention model**

- druhy:
  - hierarchická dekompozícia (Itti)
  - Bayesian: kombinácia salientnosti a predch. znalostí
  - teoretické rozhodovanie
  - maximalizácia informácie
  - založené na grafe
  - spektrálna analýza
  - klasifikácia vzorov
  - posilnenie učenia
  - hlboké učenie
- predikcia salientnej mapy (saliency map, heat map):
  - nájdu sa časti obrázku kam sa človek pravdepodobne najčastejšie pozrie

## [5. prednáška]

### Segmentácia

Typy úloh:



Object detection

Object segmentation

Object detection and recognition

Segmentácia:

- kompletnej
- parciálnej

Segmentačné metódy:

- založené na hranách
  - snake model
  - graph-cut
- založené na regiónoch
  - thresholding
  - growing regions
  - clustering
  - split/merge
- morfológické metódy - watershed
- založené na znalostach
- porovnávanie so vzormi
- štatistické metódy
- umelá inteligencia, hlboké siete

Segmentácia je komplexná úloha

- rôzne konkrétné úlohy - rôzne metódy
- neexistuje univerzálna metóda
- prečo je to také ťažké?
  - kvalita obrázku
  - komplexné tvary
  - prekrývajúce sa objekty
  - tiene
- často môže byť nutné predspracovanie
- dodatočné znalosti sú dôležité
  - vždy keď máme nejaké známe znalosti, treba ich využiť

### Segmentácia podľa hrán

- určenie hraníc objektu podľa zdetektovaných hrán na obrázku
- hrany sa prekrývajú, bývajú prerušované

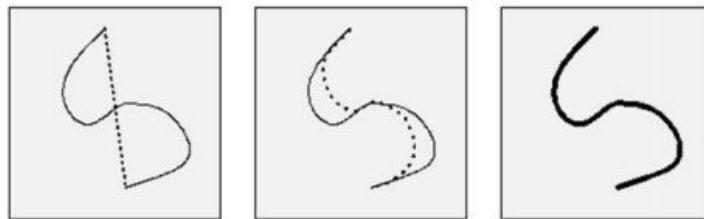
- nerelevantné hrany musíme ignorovať
- často nutný pre- a post-processing
- Sobel, Laplace, Canny

Riešenie problému prerušovaných hrán v 3 krokoch:

1. detekcia hrán,
2. distance transform,
3. watershed transform.

Active contour model - had:

- optimalizačný algoritmus
- "had" sa snaží podobať sa hranám na obrázku každou iteráciou viac a viac



- celková energia

$$E_{total} = E_{internal} + E_{external}$$

- interná en. vyjadruje hladkosť krivky, vypočítava sa ako štvorec dĺžky celej krivky (čím viac kľukatá bude, tým bude dlhšia):

$$E_{in\_elasticity} = \alpha \sum_{i=0}^{n-1} L_i^2$$

$$= \alpha \sum_{i=0}^{n-1} (x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2$$

- externá en. vyjadruje podobnosť krivky hranám, vypočítava sa ako gradient:

$$E_{ext\_grad} = -\lambda_l \sum_{i=0}^{n-1} (\left|G_x(x_i, y_i)\right|^2 + \left|G_y(x_i, y_i)\right|^2)$$

- kroky:
  - inicializácia (náhodne alebo ručne)
  - voľba vzorcov pre energie
  - jednotlivé iterácie minimalizácie celkovej energie
  - ak bola dosiahnutá stanovená energia, aktuálny stav krivky je výsledok

## Heuristic postprocessing

Heuristické hľadanie hrán:

1. nájdi zrázy najviac vyčnievajúcich hrán
2. hľadaj dopredu body hrany až kým neprídeš na koniec hrany (napr. prerušenie)
3. hľadaj dozadu podobne
4. ak takto nájdeš 3 alebo viac zrázov, nájdená cesta je správna, ulož ju a pokračuj na 1.

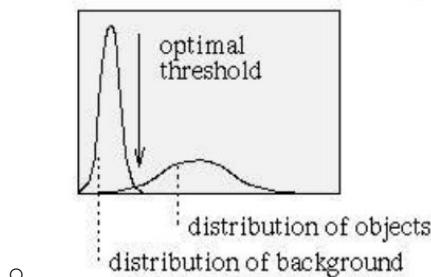
## Odstraňovanie hrán podľa kontextu:

- hrany sú veľmi závislé od šumu
- vlastnosti obrázku sú iteratívne evaluané až kým kontext hrán nie je úplne jasný
- slabá hrana medzi silnými sa považuje za správnu
- ak je osamote jedna hrana (aj keď je silná), tak sa považuje za nesprávnu

## Metódy založené na regiónoch

### Thresholding

- vhodné pre kontrastné objekty na pozadí s rovnomernou farbou
- ľahké úlohy:
  - segmentácia krvných buniek
  - vytlačený text
- Global thresholding:
  - $g(x,y) = 0$  ak  $f(x,y) < \text{threshold}$ ,  $1$  inak
  - výstup - binárna maska
  - vhodné pre obrázky, ktorých histogram obsahuje 2 Gaussove rozdelenia - optimálny threshold je tam, kde sa pretínajú tieto rozdelenia:



- Otsu-ho metóda
  - vyskúšajú sa všetky možnosti thresholdu (0-255)
  - vyhodnotí sa variancia vnútri tried a medzi triedami (trieda je určená global thresholding-om ako 0 alebo 1)
  - zvolí sa taký threshold, kde je vnútorná variancia najmenšia a medzi-triedna najväčšia
- Mixture of Gaussians (MOG)
  - predpokladáme že histogram obsahuje mix gaussových rozdelení
  - snažíme sa odhadnúť stredné hodnoty a variancie týchto rozdelení
  - po odhade určíme threshold tam, kde sa krivky pretínajú (optim. thr.)
- Využitie predošlých znalostí:
  - pr.: máme vytlačený štítok a vieme, koľko % obrazu má zaberať plocha písmen (znakov)
  - nastavíme threshold tak, aby vo výstupnej maske bolo dané množstvo pixelov
- Typy prahovania (thresholding):
  - globálne - jeden threshold na celý obrázok
  - lokálne - viac thresholdov (snaha znížiť zmeny osvetlenia)
  - adaptívne - variabilných threshold (môže byť automaticky počítaný pre každý pixel zvlášť)
  - multispektrálne - kombinácia obrázkov alebo masiek, výstupom je 1 maska

### Growing regions

- pri tejto metóde je nutné predspracovanie, hlavne odstránenie šumu

- proces:
  - volba začiatočného bodu (seed point)
  - zväčšovanie plochy okolo bodu podľa daných pravidiel

### **Watershed segmentation**

- ďalšia na regiónoch založená segmentácia
- pixely s nižšou intenzitou tvoria "nádoby" do ktorých stečie voda a vytvoria sa tak oddelené segmenty

### **Graph-cut**

- algoritmus ktorý nájde globálne optimálnu segmentáciu
- rezy grafom:
  - minimálny rez
  - normalizovaný rez
- interaktívna segmentácia
  - plne automatické metódy nemajú vždy dostatočné výsledky
  - napr. spolupráca počítač + doktor
  - človek označí rôzne časti pozadia a popredia ako seed points a algoritmus podľa toho obrázok rozsegmentuje

Segmentácia ako priestorová klasterizácia:

- pixely s podobnými hodnotami sú spojené a vytvoria jeden komponent
- segmentácia chápana ako klasifikačný problém, kde máme K regiónov a každý pixel musí do jedného z nich patriť
- je možné použiť:
  - K-means
  - mean shift
  - fuzzy c-means
  - gaussian mixture

### **K-Means klasterizácia**

- každý región je reprezentovaný strednou hodnotou bodov patriacich do regiónu
- algoritmus:
  - 1. náhodná inicializácia stredov regiónov
  - 2. klasifikácia každého bodu do najbližšieho regiónu

$$y_i = \arg \min_{k \in \{1, \dots, K\}} \|x_i - \mu_k\|^2$$

- 3. prepočet stredov regiónov

$$\mu'_k = \frac{1}{N_{y_i=k}} \sum_{i:y_i=k} x_i$$

- opakuj 2. a 3. kým sa stredy regiónov neustália

K=6 in 3-Dim feature space: RGB color space:



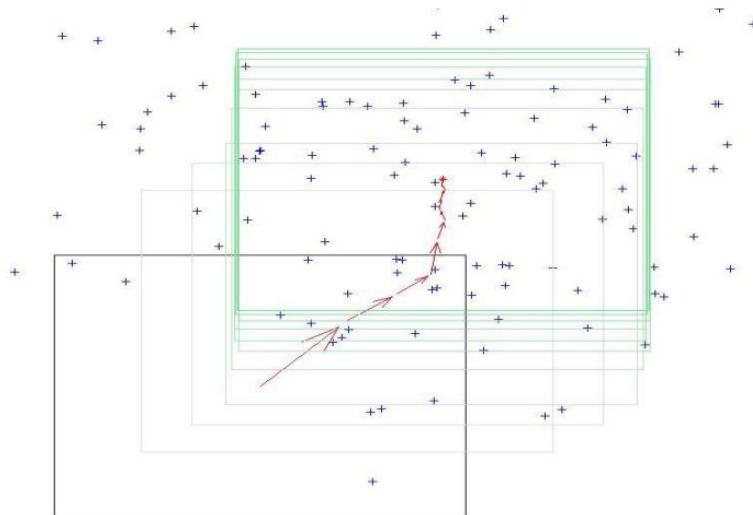
- vstupné dimenzie (informácie, ktoré sa použijú na klasterizáciu):
  - môže ich byť viac
    - farebné zložky (napr. *hue* a *saturation*)
    - pozícia pixla na obrázku
    - vlastnosť textúry
  - pokial použijeme len 1 dimenziu, tak je to len modifikácia thresholding-u

### SOM (Samo-Organizujúca mapa)

- typ neurónovej siete
- neuróny uložené v 2D mriežke

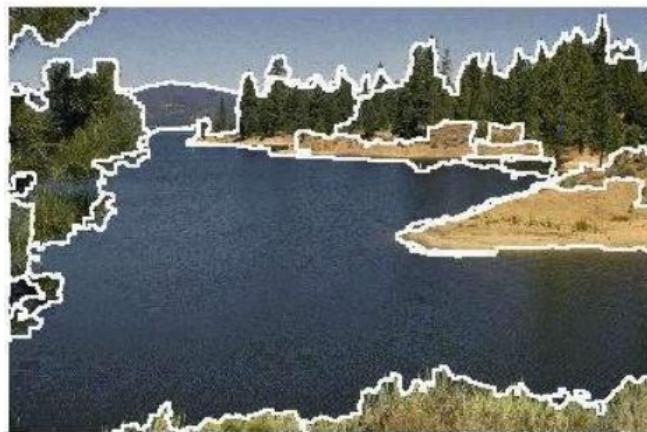
### Mean Shift Algorithm

1. zvol veľkosť prehľadávacieho okna
2. zvol počiatocnú pozíciu okna
3. nájdi lokáciu priemeru (centroid dát) v rámci okna
4. posuň okno tak aby jeho stred bol v danej lokácii
5. opakuj 3 a 4 až do konvergencie



### Mean Shift Segmentation Algorithm

1. skonvertuj obrázok na črtu (farby, gradienty...)
2. rovnomerne rozmiestni prehľadávacie okná
3. spusti mean shift algoritmus pre každé okno
4. spoj okná, ktoré skončili v rovnakých cieľoch
5. body v spojených oknách sú klastrované spolu do daného regiónu

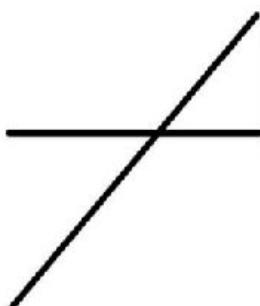


### Split & Merge Algorithm

- rekurzívne delenie obrázku na štvrtiny
- ak sa prilahlé časti podobajú (podľa stanovenej metriky), sú späť spojené
- ak nie, obe časti sa ďalej delia
- cyklus sa opakuje až kým sa nedosiahne minimálna veľkosť častí (napr. 10 px)

### Hough Transform

- metóda pre hľadanie čiar, kruhov alebo iných jednoduchých tvarov
- cieľom je nájsť objekty, ktoré majú takýto tvar pomocou "hlasovania"
- hlasovanie prebieha v tzv. akumulátore
- Hough Line Tr.
  - uvažuje sa nekonečné množstvo čiar
  - nájdú sa také z nich, ktoré najviac vyhovujú stanoveným podmienkam (prechádzajú cez najviac bodov)
  - čiara je reprezentovaná vzdialenosťou  $r$  od počiatku súradnicovej sústavy a uhlom  $\theta$  (*tetha*) kolmice (normály) čiary - takto definovaná čiara je v *Hough space*
  - postup:
    - pre každý bod sa určí niekoľko čiar, ktoré bod pretínajú v rôznych uhloch
    - vypočítajú sa parametre  $\theta$  (*uhol*) a  $r$  (*vzdialenosť*) každej čiary
    - čísla v matici na miestach určených parametrami  $\theta$  a  $r$  sa inkrementujú
    - po spracovaní všetkých bodov vznikne matica, ktorú môžeme vykresliť ako gray-scale obrázok (napravo):



vstupný obr.



Hough Line Transform

- vstupným obrázkom boli dve čiary (naľavo)
- najjasnejšie body (najväčšie čísla v matici) určia vŕťazné čiary
- na transformácii vidíme dva jasné body
- tá čiara, ktorá má najmenší rozptyl vzdialenosť pre všetky body
  
- podobná transformácia sa dá použiť aj na iné tvary (hocijaké tvary reprezentovateľné množinou parametrov)
  - krivky, kruhy atď.
- možné použiť aj na detekciu 3D objektov (napr. detekcia plochy v 3D point cloud-e)
  - z čiary sa v 3D stáva plocha, ktorá je definovaná:
    - $z = ax^*x + ay^*y + d$
  - pre hľadanie valcov sú potrebné dva kroky:
    - nájdenie orientácie valca (uhol)
    - nájdenie pozície a polomeru
- delenie Hough Transform:
  - klasické - pre hľadanie pravidelných tvarov (čiary, elipsy...)
  - generalizované - pre zložitejšie definované tvary

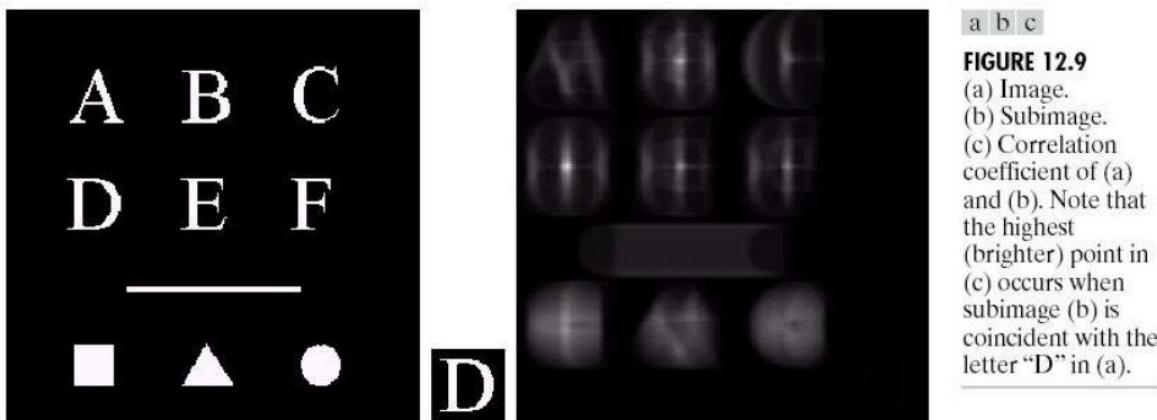
### **Superpixel segmentation**

- obrázok je rozdelený podľa mriežky veľkosti  $n$  pixelov
  - pre  $n$  sa preferuje malé číslo -> vzniká veľa segmentov (oversegmentation)
- okraje superpixelov by mali korešpondovať s okrajmi objektov na obrázku
- **Simple Linear Iterative Clustering (SLIC)**
  - obrázky v CIE L\*a\*b\* farebnom priestore
  - inicializácia: stredy klastrov sú rovnomerne rozmiestnené po obrázku
  - potom: stredy klastrov sú posunuté na pixely s najnižším gradientom z okolia 3x3 px
  - iterácia: stredy klastrov sú posúvané do priemerného pixlu z pixlov patriacich do daného klastra
  - opakuje sa, pokiaľ sa Euklidova vzdialenosť pohybu týchto stredov neustáli
  - na konci sa pospájajú oddelené pixely k blízkym superpixelom
  - tento algoritmus zabezpečí, aby stredy klastrov neboli na hranách ani na zašumených pixeloch
  - porovnanie s k-means:
    - k-means prehľadáva celý obrázok,
    - zatiaľ čo SLIC prehľadáva len blízke okolie.

## **Match-based segmentation**

### **Template segmentation:**

- **segmentácia podľa vzorov**
- posúdenie zhody pre všetky pozície a rotácie vzoru
- lokálne maximum zhodnosti reprezentuje pozíciu vzoru v obrázku
- kritérium zhodnosti je viac, najčastejšie používané je **korelácia**
  - podobné konvolúcii, avšak na obrázok prikladáme vzor



a b c

**FIGURE 12.9**  
 (a) Image.  
 (b) Subimage.  
 (c) Correlation coefficient of (a) and (b). Note that the highest (brighter) point in (c) occurs when subimage (b) is coincident with the letter "D" in (a).

#### Matching edge templates:

- hľadanie vzoru v hranách obrázku
- vzor posúvame po obrázku a počítame priemernú vzdialenosť pixelov
- tam kde je vzd. najmenšia bol vzor nájdený
- postup:
  - detekcia hrán
  - distance transform
  - hľadanie vzoru

#### Chamfer matching:

- priemerná vzdialenosť bodov vzoru a bodov (pixelov) obrázku
- dá sa tu využiť distance transform
- výpočtovo lacné

#### Feature-based segmentation:

- frekvenčné črty
- PCA

#### Segmentácia pomocou hlbokých neurónových sietí

- dosahujú najkvalitnejšie výsledky
- neurónové siete s veľkým počtom skrytých vrstiev
- využitie:
  - multi-segmentácia (U-Net)
  - generovanie salientných máp (GrabCut)

#### Segmentácia vo videu

- detekcia pozadia/popredia
- možnosť využiť informácie z predošlých snímkov
  - priemer
  - vážená kumulácia hodnôt  $b_{n+1} = a \cdot f_n + (1-a) \cdot b_n$ 
    - $b_n$  - hodnota akumulátora
    - $b_{n+1}$  nová hodnota akumulátora

## [6. prednáška]

### Detekcia, rozpoznávanie objektov

- používajú sa metódy:
  - **Rozpoznávanie vzorov** (pattern recognition) - Rule based/statistical/fuzzy pattern recognition
  - **Umelá inteligencia** - Detekcia črt a klasifikácia (Feature detection + classification), trénovanie s dátami
  - **Neurónové siete** - metódy hlbokého učenia - konvolučné neurónové siete CNN

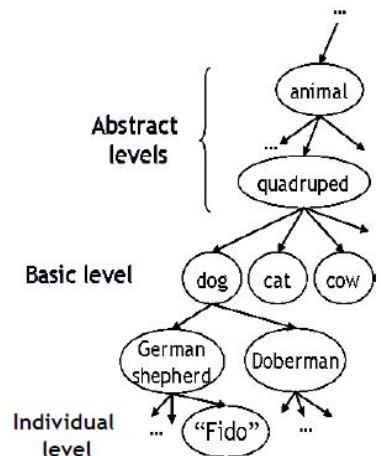
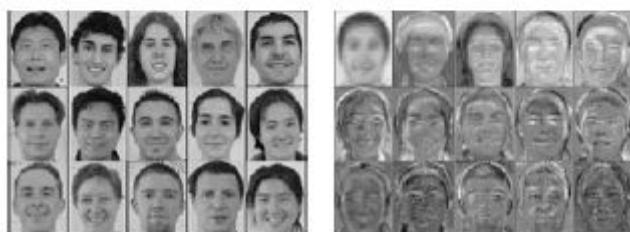
### Kategória objektu vs. inštancia

- detekcia, rozpoznanie kategórie objektu - mnoho variácií v kategórii => podstatná je generalizácia
- detekcia, rozpoznanie inštancie objektu - potreba rozlišovať medzi podobnými objektami

### Vizuálne kategórie objektu

- kategórie základ. levelu
- ľudia zvyčajne začínajú s kategorizáciou základ. levelu, predtým ako začnú s identifikáciou
- kategória základného levelu je jednoduchšia a rýchlejšia pre ľudí ako identifikácia objektov.

### Detekcia vs. rozpoznanie



### Výzvy

- zmeny v osvetlení, uhlo kamery, póza objektu, medzi-triedne variácie..

### Základné prístupy

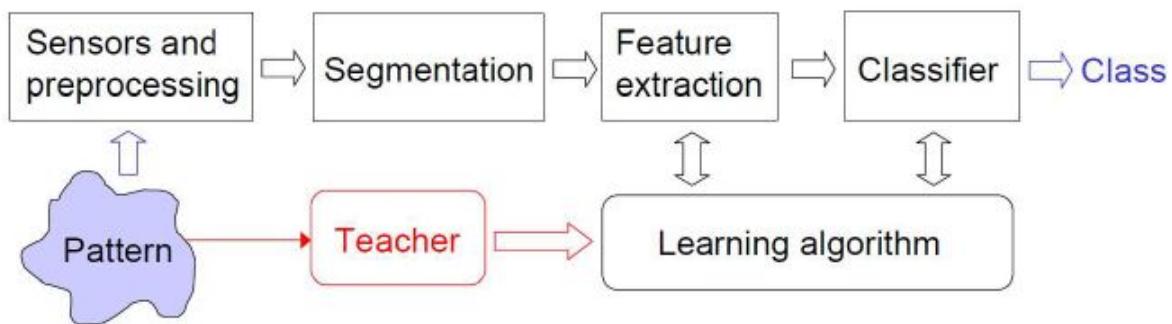
- *bottom-up* - part-based reprezentácie, **detekcia lokálnych črt + rozpoznanie**

- *top-down* - **segmentácia + rozpoznanie objektu**, Global appearance recognition - sliding window
- *hlboké učenie + konvolučné neurónové siete CNN*

### Segmentácia + rozpoznanie objektov

- Robustná segmentácia - geometrický objekt (značky na ceste), farebné dominantný objekt  
Príklad: detekcia znaciek na ceste, OCR...

### Rekognícia objektov s využitím segmentácie a klasifikácie



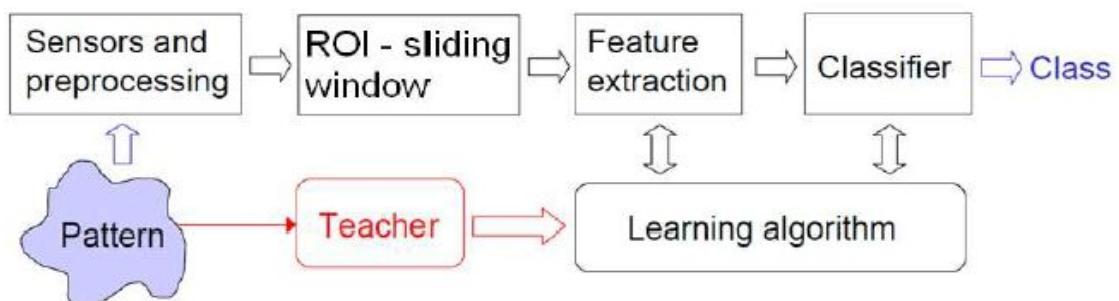
### Global Appearance & Sliding window

Sliding window



Examples: Face detection,

People detection, ....



## Úloha binárnej klasifikácie

- klasifikátor zodpovedá na otázku - nachádza sa v danom okne objekt? (Y/N)

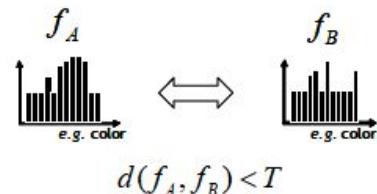
## Lokálne deskriptory

- robustnejšie
- occlusion of objects
- zmeny v pozícii pohľadu kamery
- nemennosť rotácie, veľkosti
- medzi kategóriké variácie



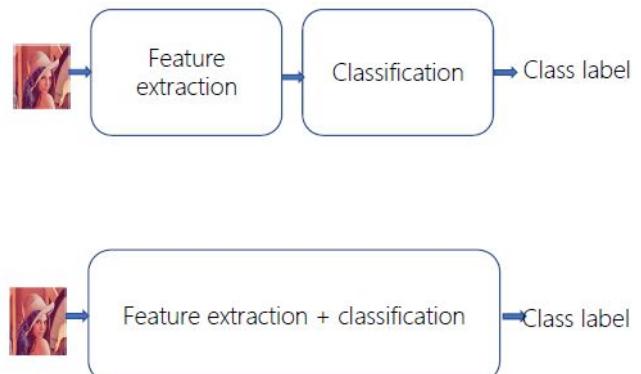
## Detekcia objektov s využ. lokál. deskriptorov

1. Nájdi set rozličných klúčových bodov.
2. Definuj oblasť okolo každého klúčového bodu
3. Extrahuj a normalizuj obsah oblasti
4. Vypočítajte lokálny deskriptor z normalizovanej oblasti
5. Porovnaj lokálne deskriptory



## Hlboké učenie a Klasifikácia

- Tradičný prístup
  - suboptimal
  - vyžaduje odborné znalosti
  - ručne vytvorené funkcie
- Hlboké učenie
  - klasifikátory aj črty sú učené automaticky



## Neurónové siete

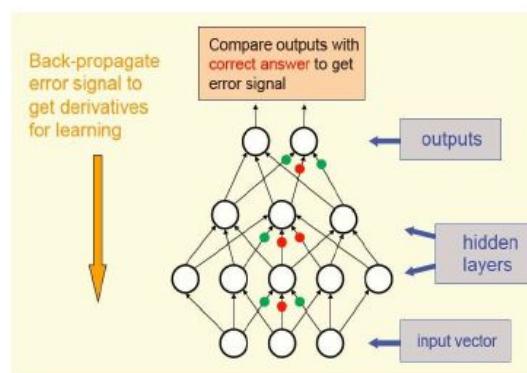
- viacvrstvový perceptron
- back-propagation (učenie s učiteľom)
- differentiable activation function
- stochastic gradient descent

Empirical risk  

$$Q(w) = \sum_{i=1}^n Q_i(w),$$

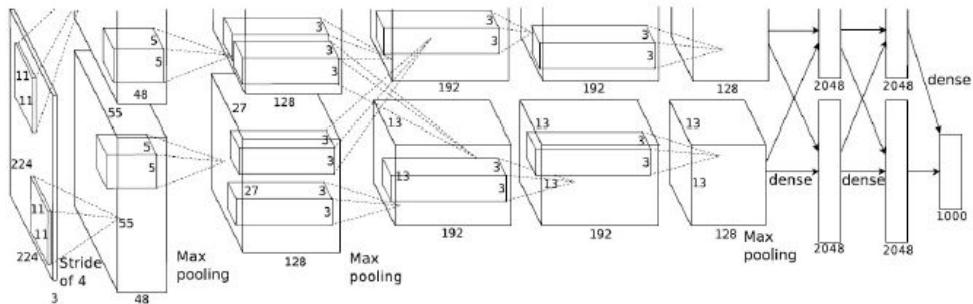
Update weights:

$$w := w - \alpha \nabla Q_i(w).$$



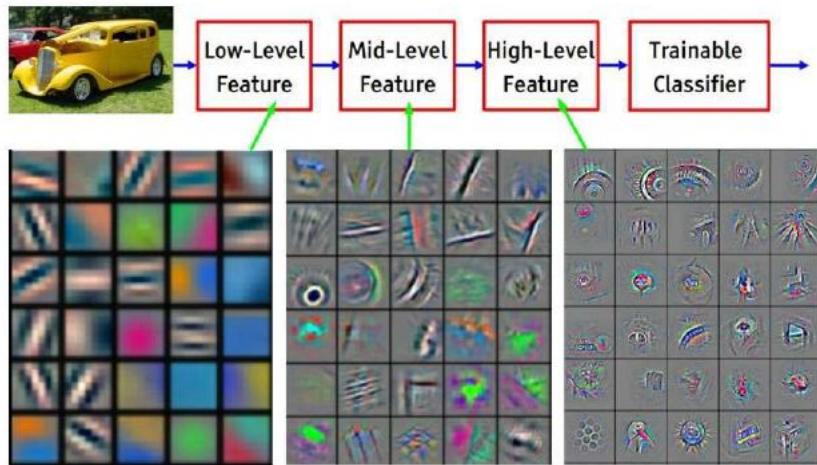
**Hlboké učenie** - je súbor algoritmov strojového učenia založený na viacvrstvových sieťach

## Deep convolutional neural network example



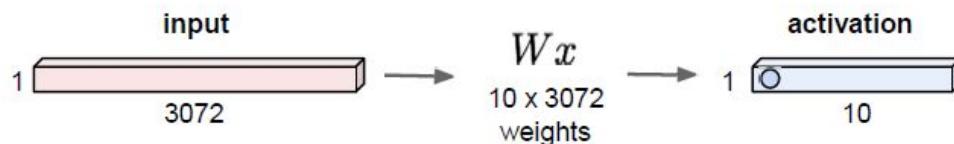
- 5 convolutional layers, 3 fully connected layers
- 60M parameters, trained on 1.2M images (~1000 examples for each category)

### Vizualizácia CNN čít



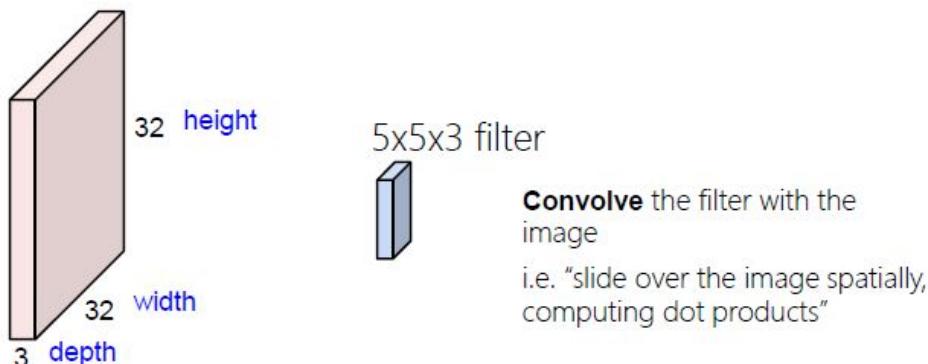
## Fully Connected Layer

32x32x3 image -> stretch to 3072 x 1



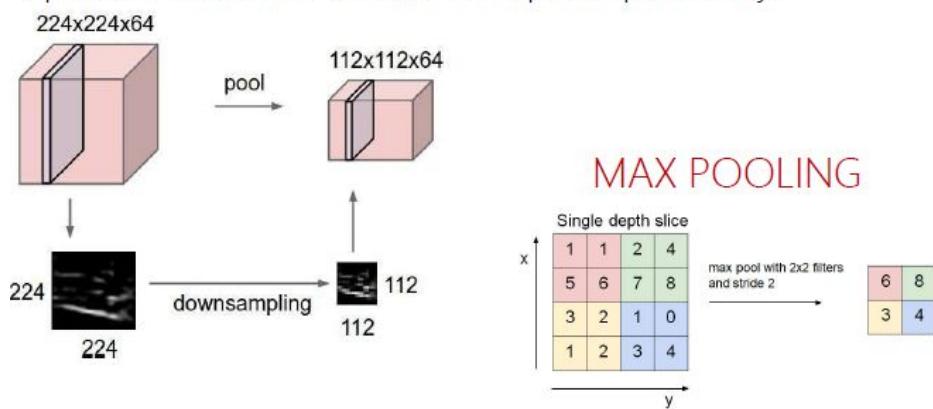
## Convolution Layer

32x32x3 image -> preserve spatial structure



## Pooling layer

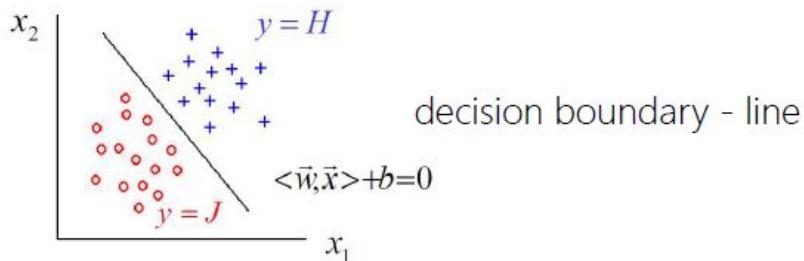
makes the representations smaller and more manageable  
operates over each activation map independently:



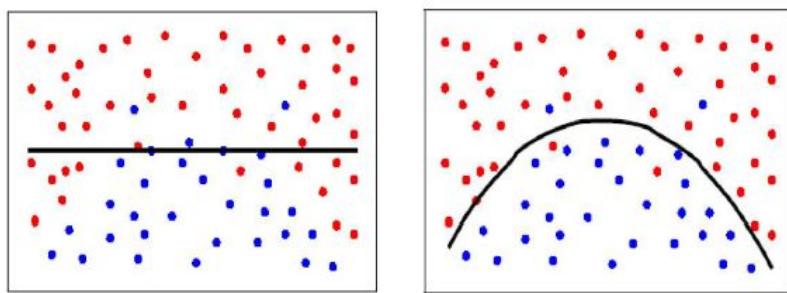
**Črty pre detekciu/rozpoznanie objektov** - farba, tvar, textúra, hrany...

**Vektor črt**

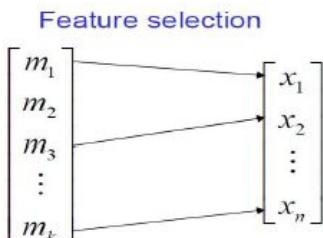
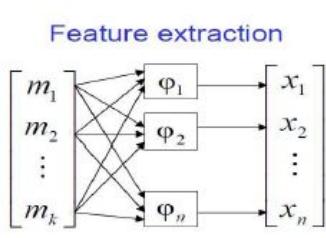
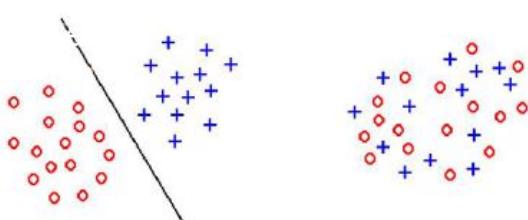
$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$



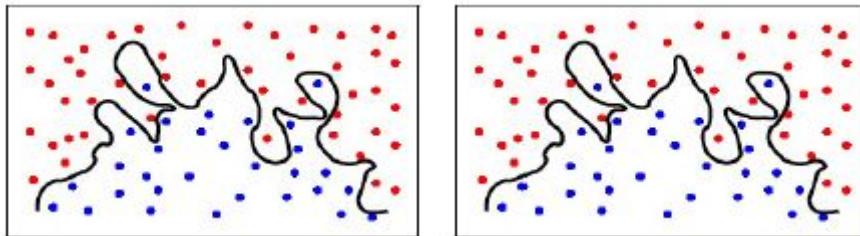
**Črty** - lineárne/ nelineárne oddeliteľné triedy



Feature extraction  
/ feature selection

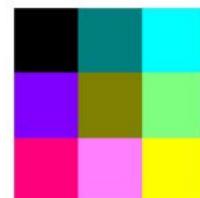


Problém - overfitting - dôležitá je generalizácia, Cross-validation

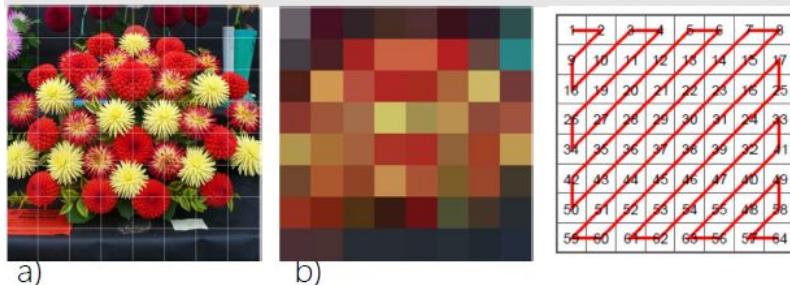


### Farebné črty

- dominantné farby
  - dominantná farba pokrýva veľkú časť obrázka
  - 1 alebo viac
  - deskriptor dominantnej farby => farba + %
  - problém: priestorová informácia nie je zahrnutá
  - vizu. -> 1D histogram (graf/matica)
- The scalable color descriptor
  - je odvodený z farebného histogramu definovaného vo farebnom priestore HSV s kvantizáciou s pevným farebným priestorom.
  - Využíva Haar transformáciu (HT), kódovanie koeficientov, ktoré umožňujú škálovateľnú reprezentáciu opisu, ako aj komplexnosť škálovateľnosti postupov extrakcie a párovania funkcií. HT predstavuje histogramy s rôznym počtom tried



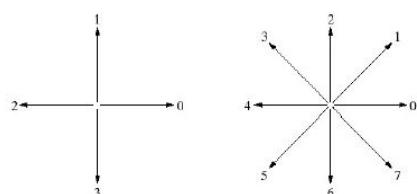
### The Scalable Colour Descriptor (MPEG7)



- a) the image is divided into  $8 \times 8$  blocks
- b) the average colour of blocks.
- c) Zig-zag ordering of coefficients in the descriptor distribution of colours

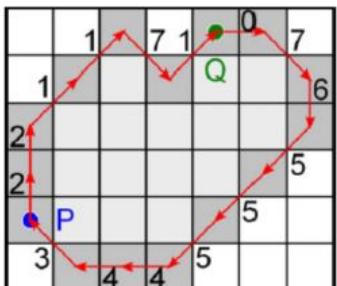
### Tvarové črty

- typicky používané na binárny obr., ktorý sme dostali po segmentácii obrázka
- reprezentácia tvaru
  - chain codes - reprezentujú hranicu spojenou sekvenciou rovných čiar špecifickej dĺžky a smeru.
    - 4 directional chain codes
    - 8 directional chain codes



- Freeman chain code

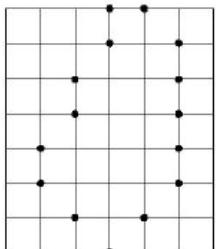
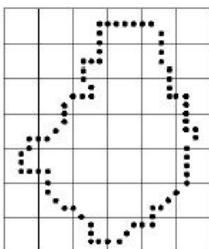
## Freeman Chain code



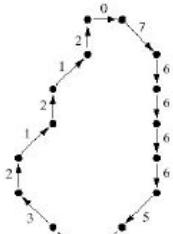
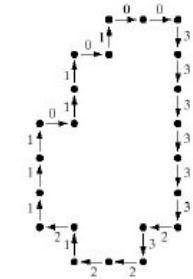
For 8-neighbor code a limit of 8-neighborhood of the point 0, ..., 7 as shown in Figure

The chain (Freeman) code boundary object is then a sequence of numbers that contain information that direction limit from the point continues.

- tvarová reprezentácia
  - normalizácia pre rotáciu
  - normalizácia pre začiatočný bod
  - normalizácia pre veľkosť



**FIGURE 11.2**  
 (a) Digital boundary with resampling grid superimposed.  
 (b) Result of resampling.  
 (c) 4-directional chain code.  
 (d) 8-directional chain code.

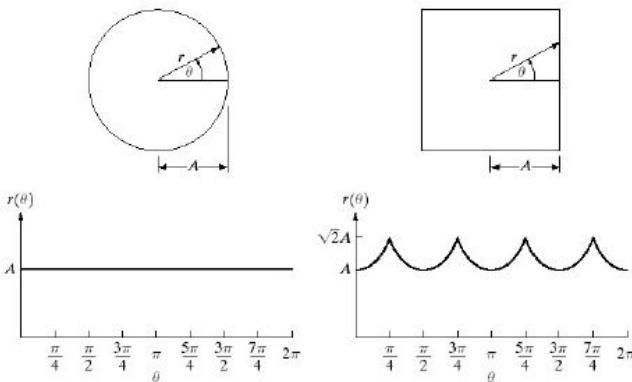


- signatures

- 1-D funkčná reprezentácia hranice
- Základná myšlienka: Znížte hraničnú reprezentáciu na 1-D funkciu, ktorá by sa dala ľahšie opísať ako 2-D hranica
- jednoduchý prístup: použitie vzdialenosť od ťažiska k hranici ako funkciu uhla. (nemení transláciu ale rotáciu a škálovanie áno)
- rotácia - začiatočný bod = najvzdialenejší bod od ťažiska
- škálovanie - normalizuj funkciu varianciou

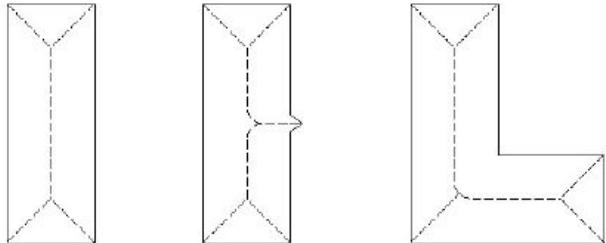
a b

**FIGURE 11.5**  
Distance-versus-angle signatures.  
In (a)  $r(\theta)$  is constant. In (b),  
the signature consists of  
repetitions of the  
pattern  
 $r(\theta) = A \sec \theta$  for  
 $0 \leq \theta \leq \pi/4$  and  
 $r(\theta) = A \csc \theta$  for  
 $\pi/4 < \theta \leq \pi/2$ .



- skeleton of region

- použí na reprezentovanie regiónu
- skeletonizing (thinning) regiónu
- výpočtovo náročné



a b c

**FIGURE 11.7**  
Medial axes  
(dashed) of three  
simple regions.

### Textúra - črty

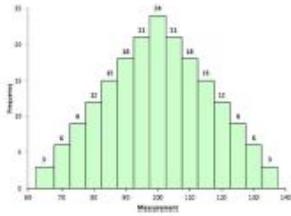
- Textúra je opis priestorového usporiadania farieb alebo intenzít v obraze alebo vybranej oblasti obrazu.
  - štrukturálne vs. štatistické prístupy
    - statistical texture measures
      - segmentovanie textonov (skupiny vizuál. vlastností)
      - Numerické veličiny alebo štatistiky, ktoré opisujú textúru, sa dajú vypočítať zo samotných sivých tónov (alebo farieb)
      - málo intuitívny ale efektívny prístup
      - použiteľný pre klasifikáciu aj segmentáciu
      - Statistical moments

## 2. Standard deviation

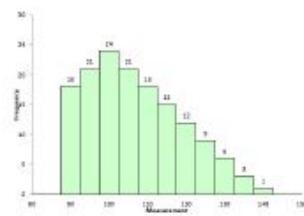
$$s_x = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

## 3. Skewness – degree of symmetry in the distribution

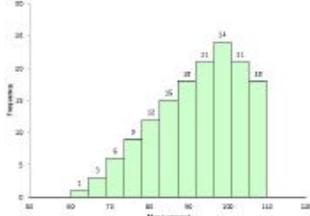
$$\text{Skewness} = \frac{n}{(n-1)(n-2)} \sum \frac{(X_i - \bar{X})^3}{s^3}$$



Symmetrical Dataset with Skewness = 0



Dataset with Positive Skewness



Dataset with Negative Skewness

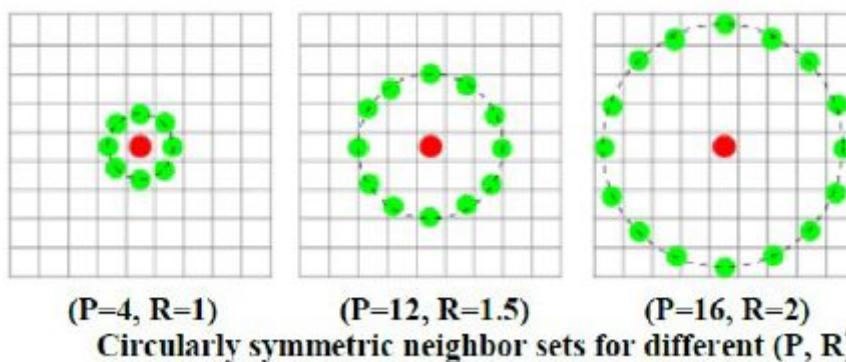
## 4. Kurtosis – peakedness of the distribution

$$\text{Kurtosis} = \left\{ \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum \frac{(X_i - \bar{X})^4}{s^4} \right\}$$

- edge-based measures
- local binary patterns - LBP
  - pre každý pixel p vytvor 8bit nr. b1-b8, kde bi = 0 ak sused i má hodnotu <= ako p, inak bi = 1
  - Reprezentuje textúru v obrazu (alebo oblasti) histogramom týchto čísel.

	1	2	3	
8	100	101	103	4
	40	50	80	→
	50	60	90	
7	6		5	

[ 1 1 1 1 1 1 0 0 ]



- co-occurrence matrices
  - 2d pole C, v ktorom:
    - riadky aj stĺpce reprezentujú set možných hodnôt obrázka

- $C_d(i,j)$  - koľkokrát sa hodnota  $i$  opakovane objavuje s hodnotou  $j$  v konkrétnom priestorovom vzťahu
- priestor. vzťah - špecifikovaný vektorom  $d = (dr, dc)$

GLCM

1	1	5	6	8
2	3	5	7	1
4	5	7	1	2
8	5	1	2	5

1	1	2	0	0	1	0	0	0
2	0	0	1	0	1	0	0	0
3	0	0	0	0	1	0	0	0
4	0	0	0	0	1	0	0	0
5	1	0	0	0	0	1	2	0
6	0	0	0	0	0	0	0	1
7	2	0	0	0	0	0	0	0
8	0	0	0	0	1	0	0	0

co-occurrence matrix

From  $C_d$  we can compute  
**N- the normalized co-occurrence matrix**,  
where each value is divided by the sum of all the values.

64

## Co-occurrence Features

Energy measures uniformity of the normalized matrix.

$$\begin{aligned}
Energy &= \sum_i \sum_j N_d^2(i,j) \\
Entropy &= - \sum_i \sum_j N_d(i,j) \log_2 N_d(i,j) \\
Contrast &= \sum_i \sum_j (i-j)^2 N_d(i,j) \\
Homogeneity &= \sum_i \sum_j \frac{N_d(i,j)}{1 + |i-j|} \\
Correlation &= \frac{\sum_i \sum_j (i - \mu_i)(j - \mu_j) N_d(i,j)}{\sigma_i \sigma_j}
\end{aligned}$$

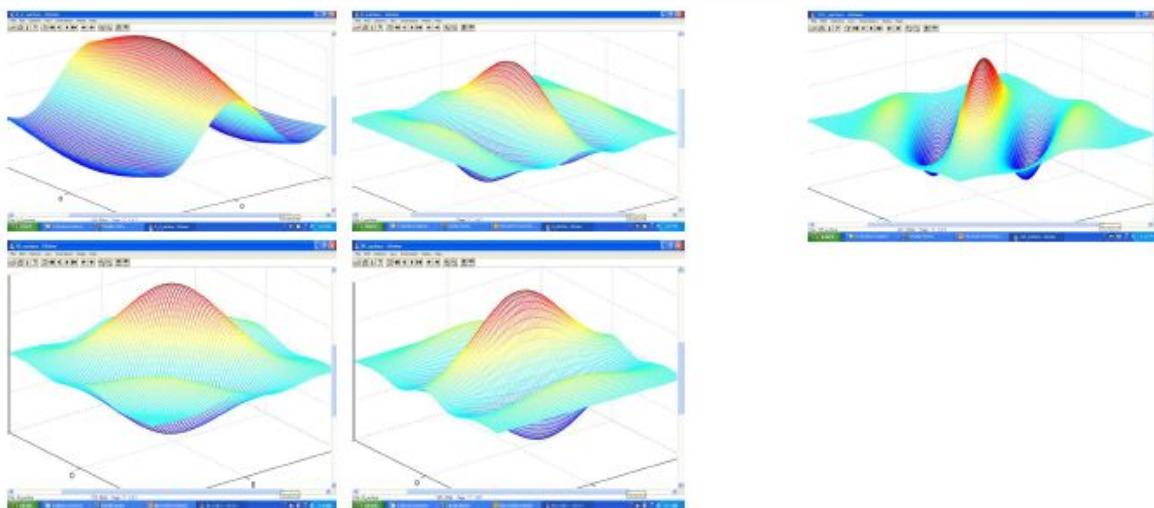
where  $\mu_i, \mu_j$  are the means and  $\sigma_i, \sigma_j$  are the standard deviations of the row and column

- gabor filters
  - wavelets na rôznych frekvenciach a orientáciách

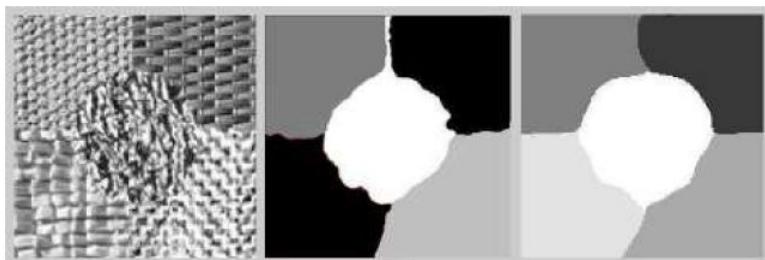
Generalised Gabor functions :

$$\gamma(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp \left[ -\frac{1}{2} \left( \frac{(x-x_0)^2}{\sigma_x^2} + \frac{(y-y_0)^2}{\sigma_y^2} \right) + 2\pi\sqrt{-1}(u_0x + v_0y) \right]$$

## Gabor Filters Convolution kernels – examples in 3D viz.



## Gabor Filters – segmentation example

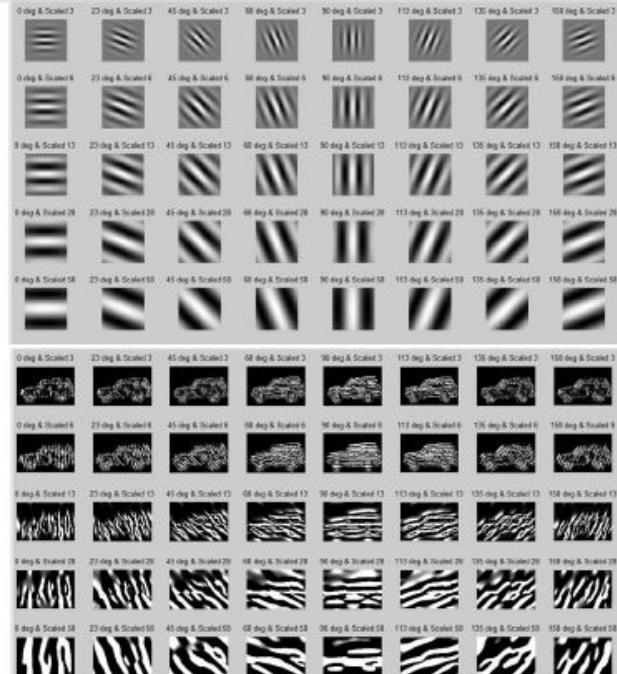


# Gabor Filters – segmentation example



Input image

Gabor kernels



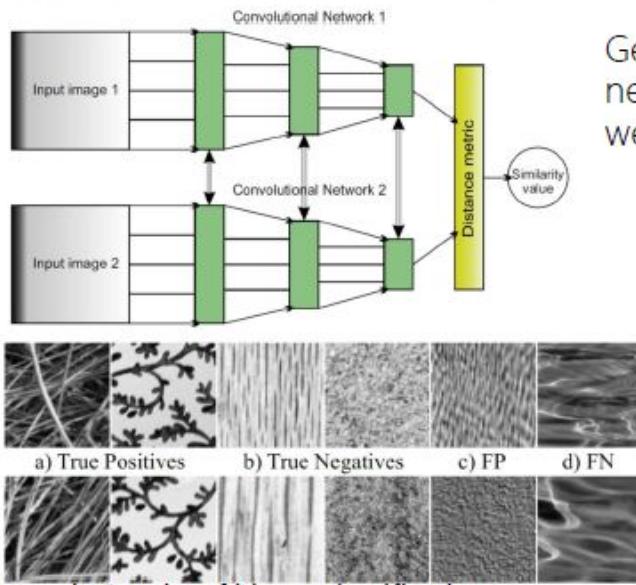
Convolution output

Zheng, Danian, Yannan Zhao, and Jiaxin Wang. "Features extraction using a gabor filter family." *Proceedings of the sixth>Lasted International conference, Signal and Image processing, Hawaii.* 2004.

Computer vision vgg.fiiit.stuba.sk

## Textúra a konvolučné neurónové siete

Texture Similarity Evaluation via Siamese Convolutional Neural Network



Example Results of binary classification

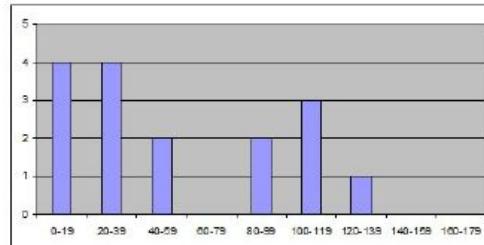
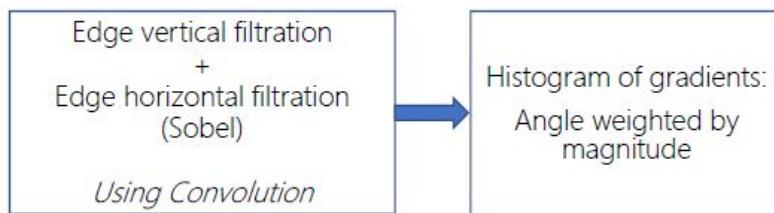
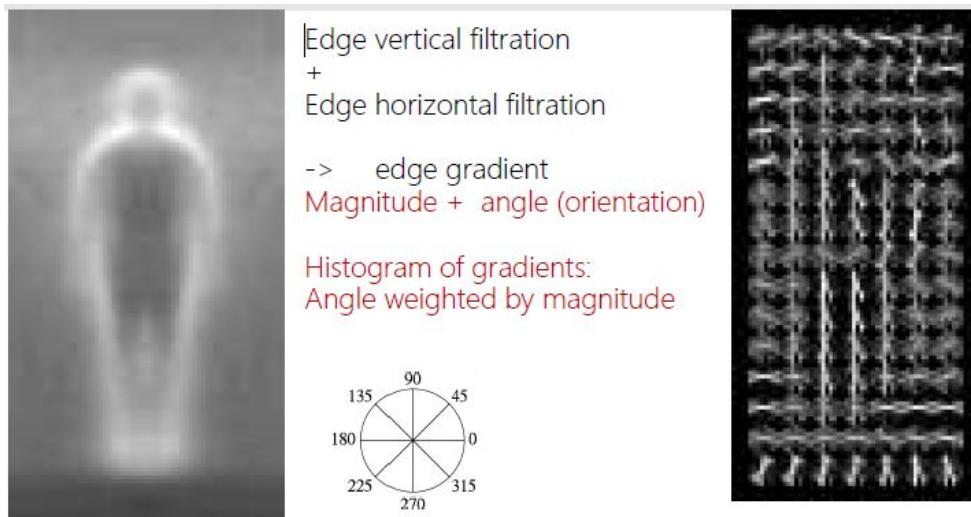
General architecture of Siamese network, two networks sharing weights and metric loss

OUR NETWORK ARCHITECTURE.

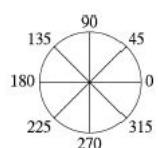
Layer	Kernel Size	Stride	Output chn.
conv1	5x5	3x3	96
pool1	3x3	3x3	
conv2	5x5	1x1	265
pool2	3x3	2x2	
conv3, conv4	3x3	1x1	384
conv5	3x3	1x1	256
conv-fc1	1x1	1x1	1024
conv-fc2	1x1	1x1	128

## Črty hrán

- Histogram of gradient orientations HOG

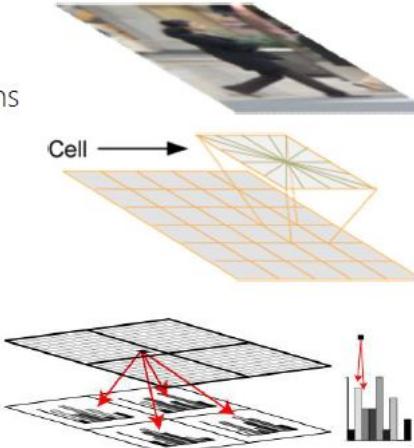


Cell histograms  
typically  
8 (or 9) bins for gradient orientations  
(0-180 degrees)  
Filled with magnitudes



HOG feature: chain of data  
4 cells

$$f = (h_1^1, \dots, h_9^1, h_1^2, \dots, h_9^2, h_1^3, \dots, h_9^3, h_1^4, \dots, h_9^4)$$



## Haar črty

- suma pixelov, ktorá je v bielom obdĺžniku je odčítaná od sumy pixelov v sivých obdĺžnikoch -> vyrátanie rozdielov medzi sumami pixelov v obdĺžnikoch.
- podobné Haar wavelets
- efektívne na výpočet s použitím integral image
- napr. face recognition

