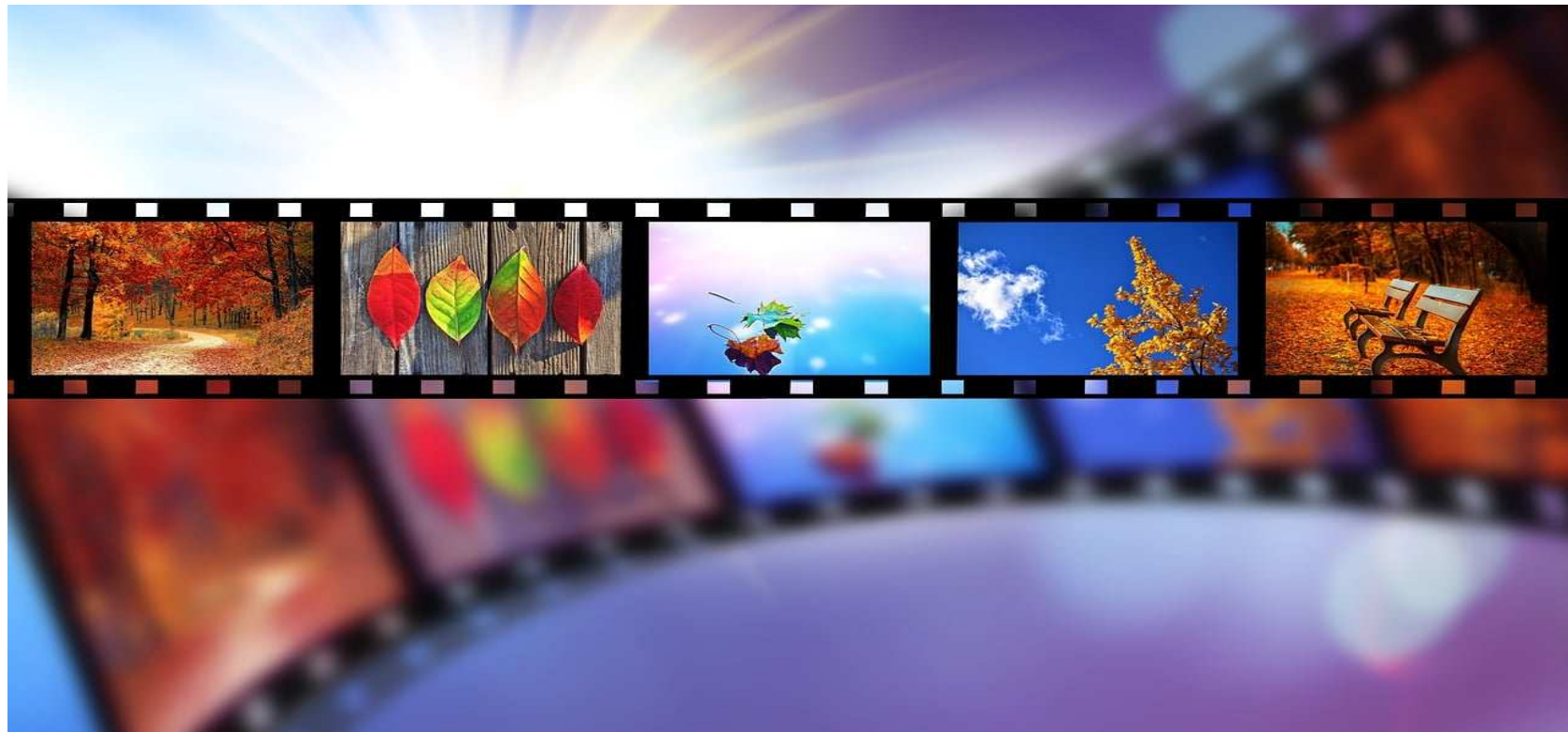


# Analýza videa, vizuální segmentace Camshift

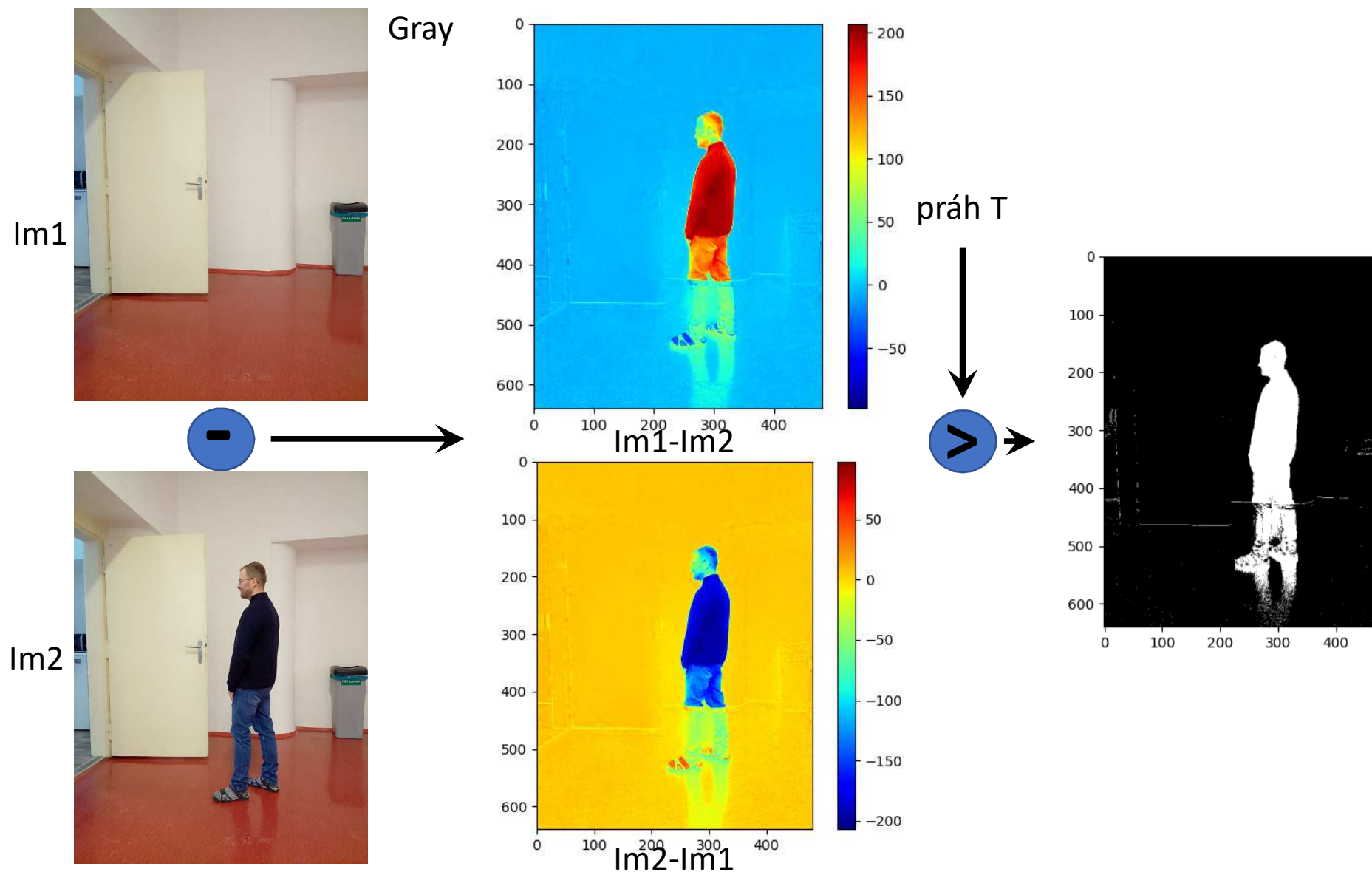
doc. Ing. Josef Chaloupka, Ph.D.



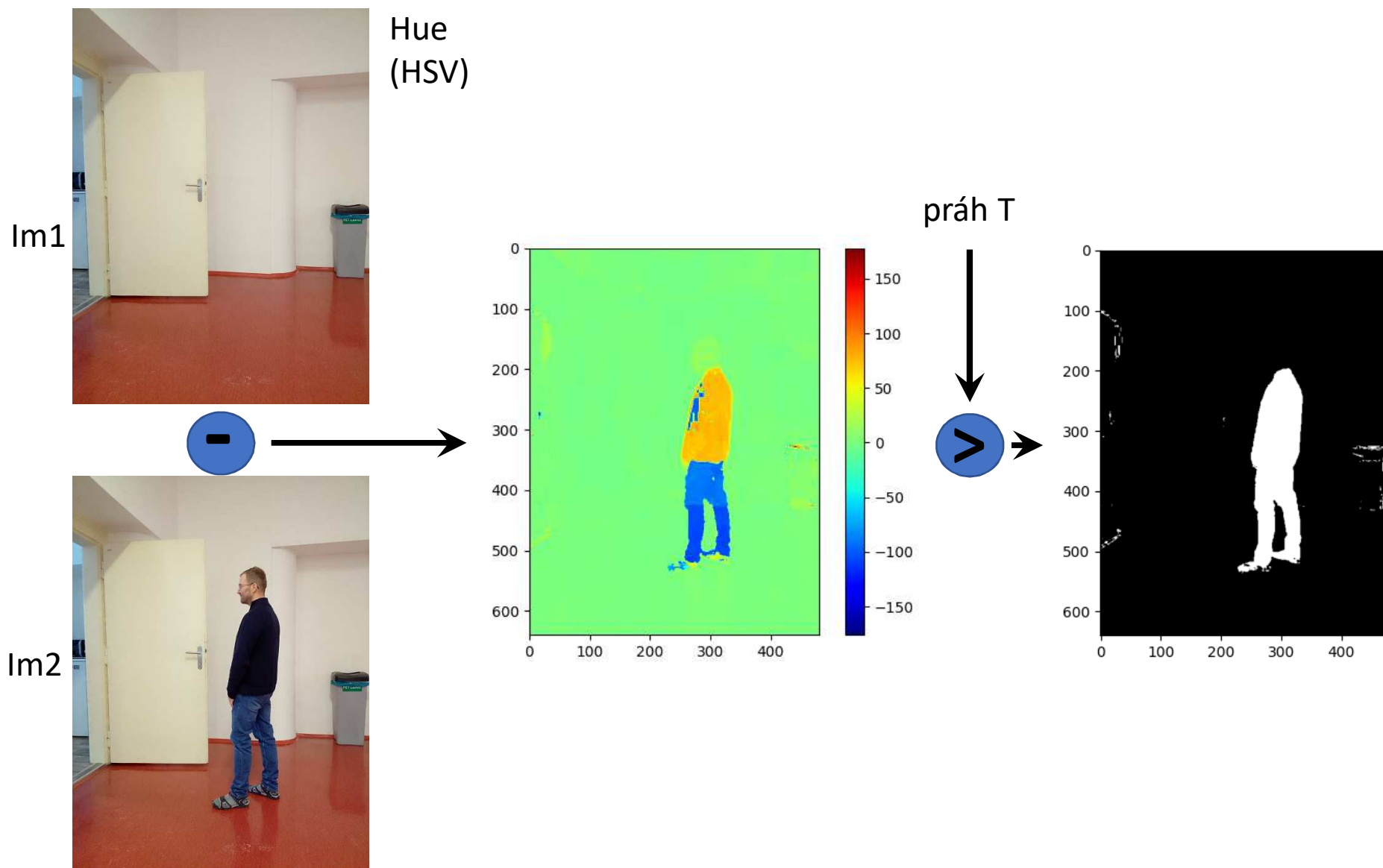
## Video: Metoda odečítání pozadí (Background Subtraction)



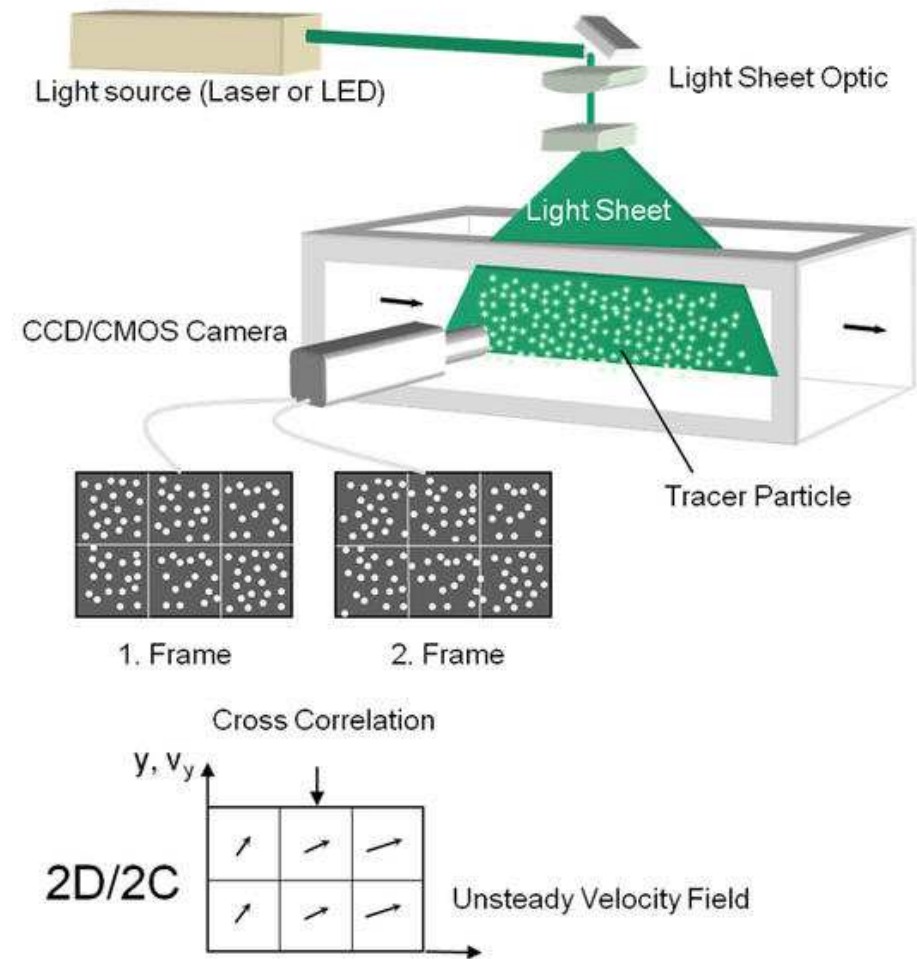
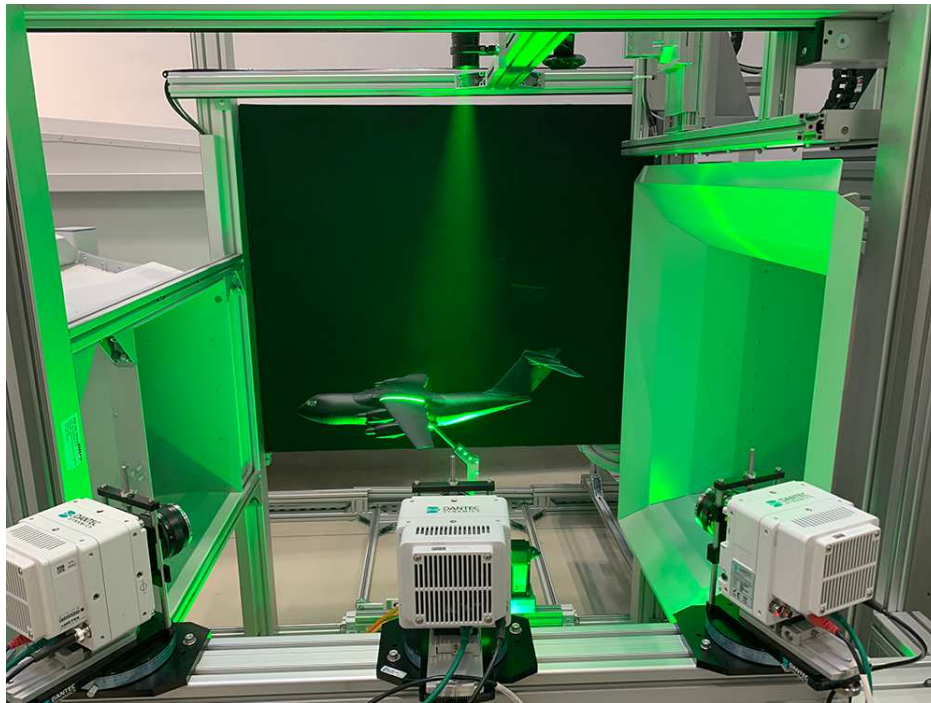
# Video: Metoda odečítání pozadí (Background Subtraction)



# Video: Metoda odečítání pozadí (Background Subtraction)

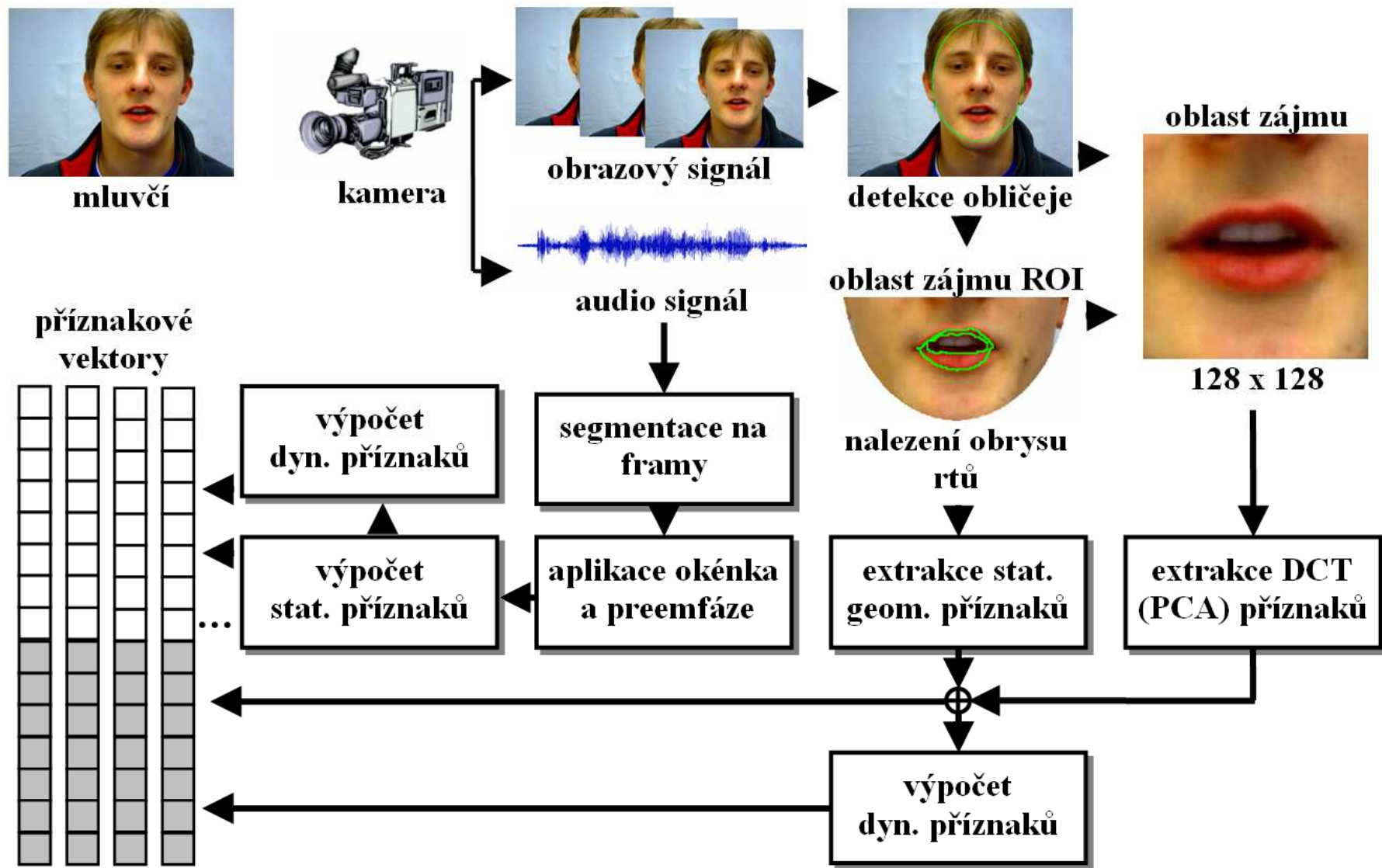


# Video: PVI - Particle Image Velocimetry

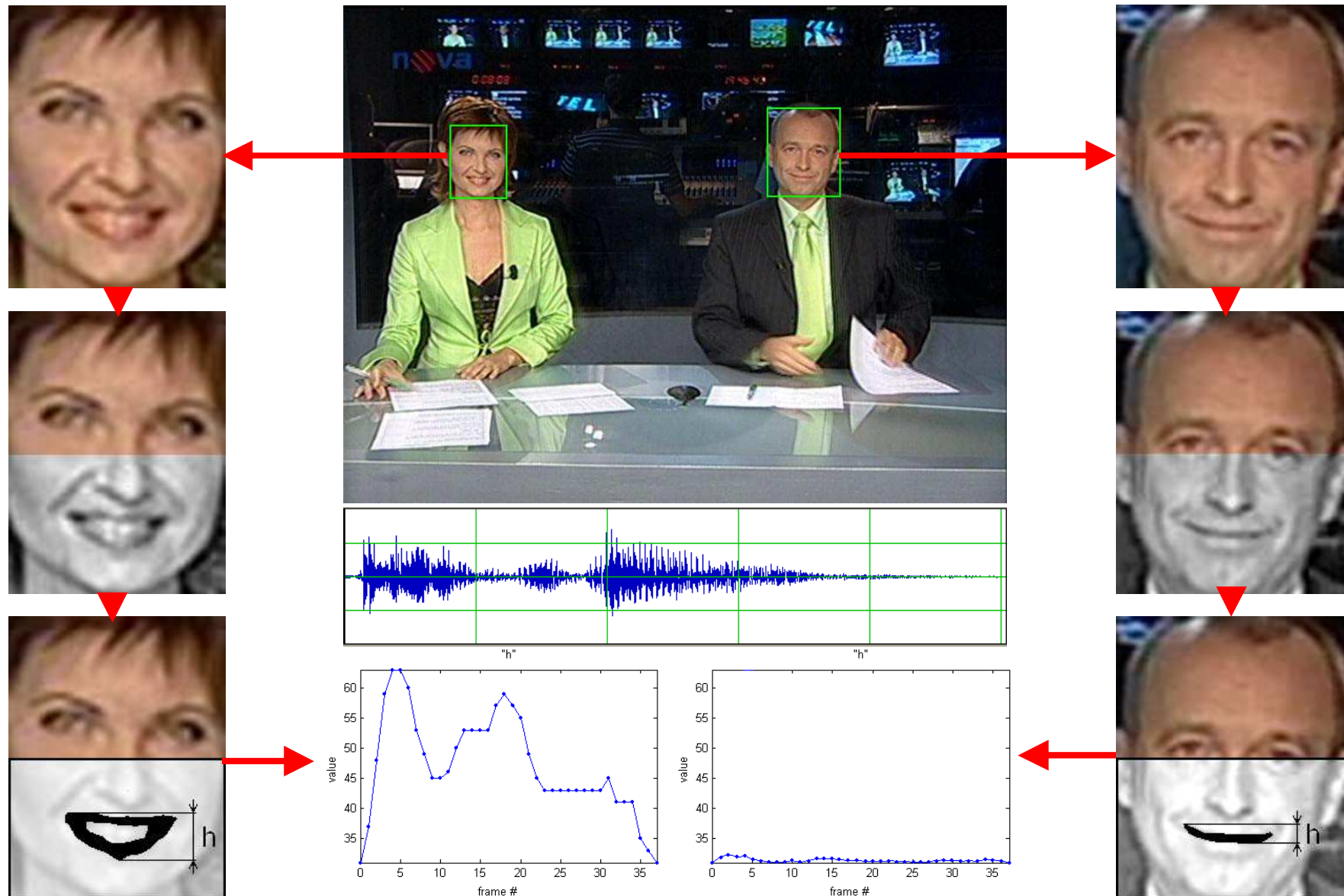




# Video: Audio-vizuální rozpoznávání řeči

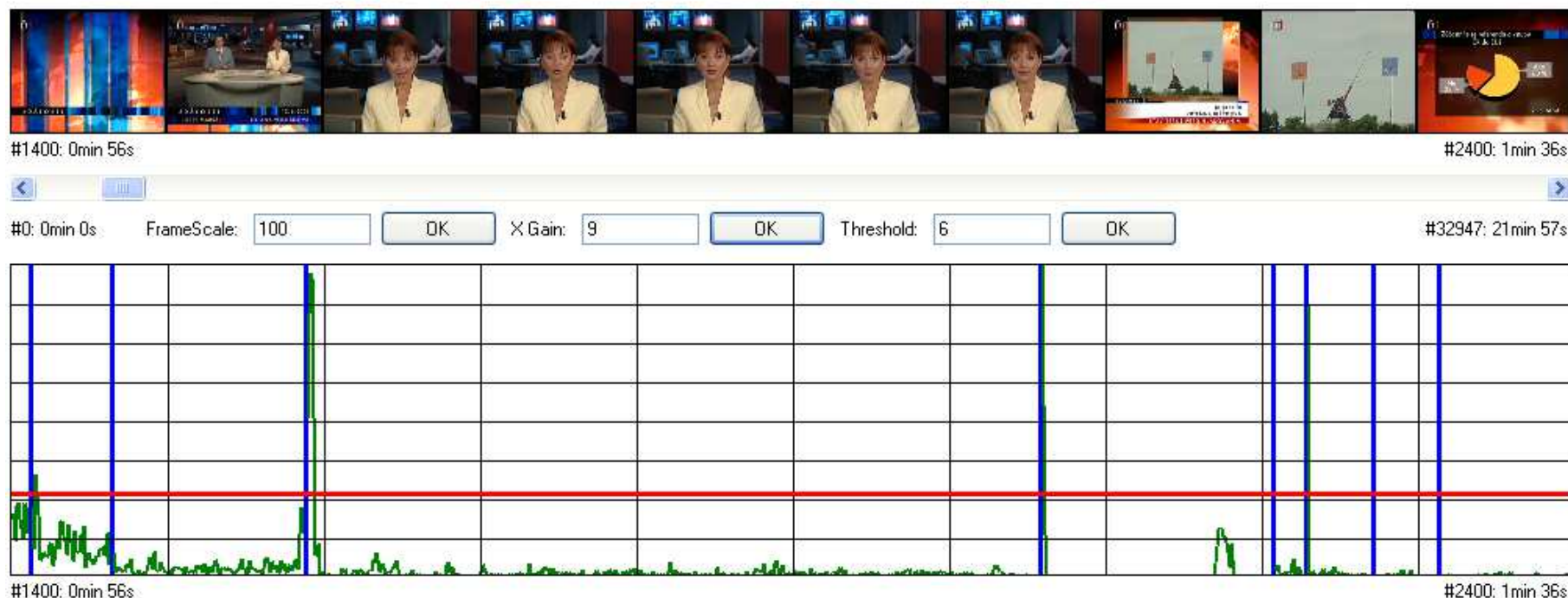


# Video: vizuální detektor řeči



# Video: vizuální segmentace obrazu

- Rozdělení vizuální signálu na homogenní části, kde se vizuální informace mění v čase málo nebo vůbec
- Metody pro vizuální segmentaci: 1) Metody založené na porovnání informace v obrazových bodech, 2) Histogramově zaměřené metody, 3) Blokově zaměřené metody, 4) Metody využívající příznaky, 5) Metody vycházející ze změny pohybu, ...





# Vizuální segmentace obrazu



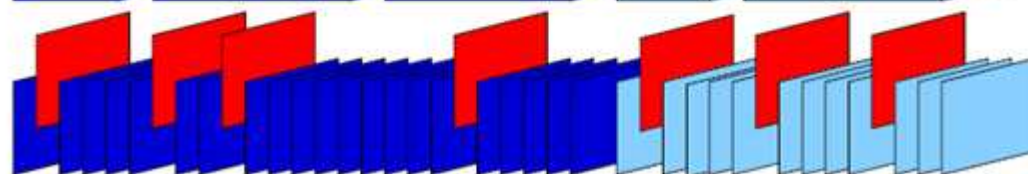
Video sekvence



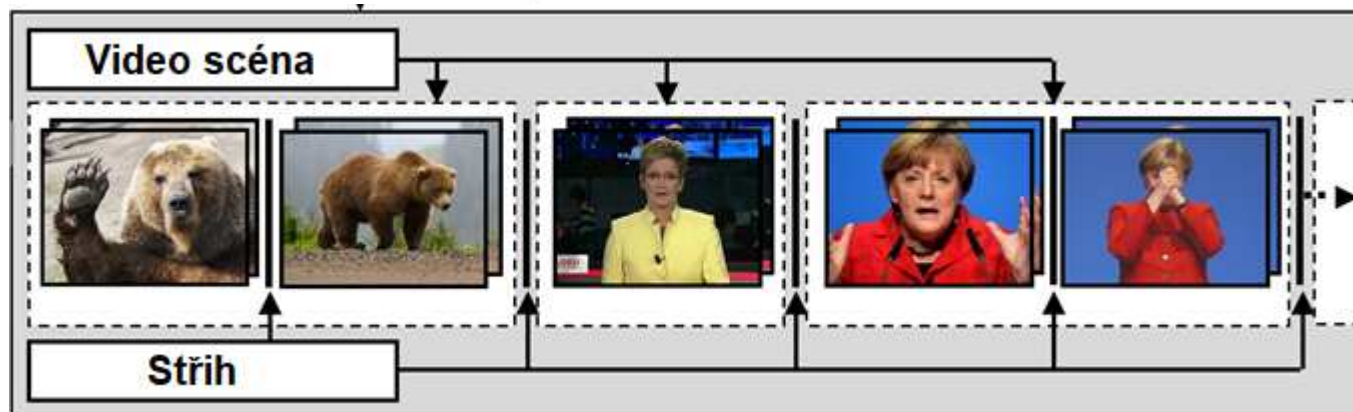
Video scény



Střihy



Klíčové snímky



# Rozpoznávání objektů v klíčových snímcích

Anténa L: Samsung L: Mc Donald OCR: Vogue L: Coca Cola Obličej Lampa

Dům

Auto

OCR: Dog Night Time

Autobus

Auto

Lidé

Dům

TDK

L: TDK

Hyundai

L: Hyundai

OCR: Hyundai

# Rozpoznávání objektů v klíčových snímcích



# Vizuální segmentace obrazu - přechody



Ostrý střih



rozsvětlení



ztmavení



# Vizuální segmentace obrazu - přechody



průnik



clona



změna scény pohybem

# Vizuální segmentace – metody

- **Metody založené na porovnání informace v obrazových bodech**

- Suma barevných hodnot pixelů ve dvou následujících snímcích

$$(1) \left| \sum_{x=1}^X \sum_{y=1}^Y f_i(x, y) - \sum_{x=1}^X \sum_{y=1}^Y f_{i+N}(x, y) \right| > T \quad (2) \sum_{x=1}^X \sum_{y=1}^Y |f_i(x, y) - f_{i+N}(x, y)| > T$$

- T – segmentační práh, N – posun snímků, obvykle 1



8606 (00:05:44) 8710 (00:05:48) 8814 (00:05:53) 8815 (00:05:53)



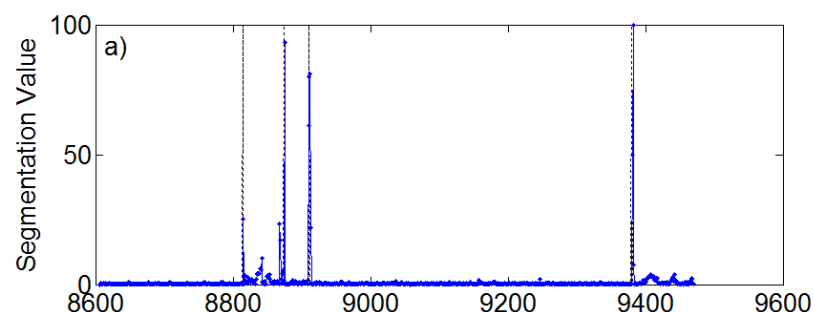
8844 (00:05:54) 8874 (00:05:55) 8875 (00:05:55) 8893 (00:05:56)



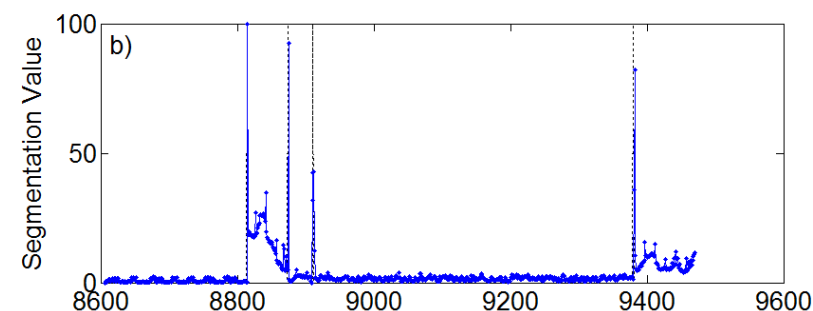
8909 (00:05:56) 8912 (00:05:56) 9146 (00:06:06) 9380 (00:06:15)

# frame (hodina:minuta:sekunda)

(1)



(2)

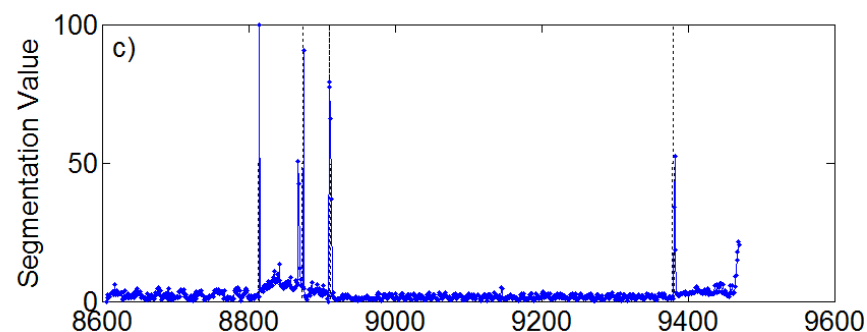
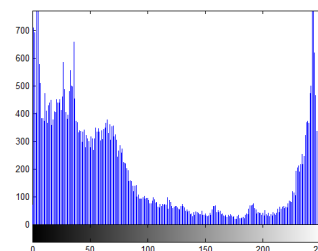
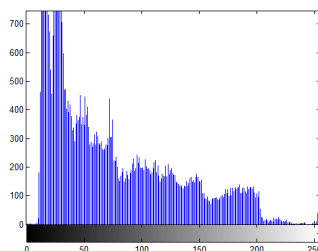
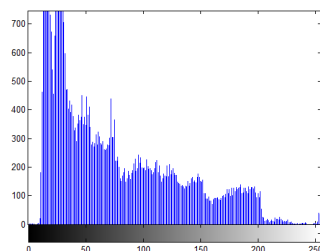
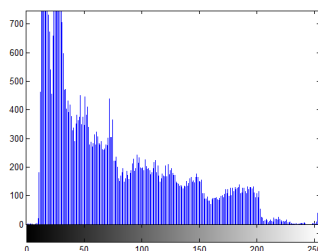


# Vizuální segmentace – metody

- **Histogramově zaměřené metody**

- 1 obraz = 1 histogram, 1 histogram != 1 obraz

$$\sum_{v=1}^V |H_i(v) - H_{i+N}(v)| > T$$



# Vizuální segmentace – metody

- **Metody využívající obrazové příznaky**
- Video snímek charakterizován vektorem příznaků
- Vizuální příznaky: momenty počítané z obrazu, hrany, příznaky získané z barevných přechodů, příznaky získané na základě použití různých statistických metod a přístupů (střední hodnoty, rozptyly, kovarianční matice, Bayesův model)
- Jako příznaky lze také využít koeficienty z různých 2D transformací obrazu, např. parametry získané pomocí Houghovy transformace nebo diskretní Fourierovy transformace DFT, vlnkové transformace DWT, diskretní kosinové transformace DCT



# Vizuální segmentace – metody

- **Metody využívající obrazové příznaky – výpočet z DCT**

- 1) Převod snímku do barevné složky zvoleného barevného prostoru, 2) výpočet DCT:

$$F(u,v) = \frac{2c(u)c(v)}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \cos\left(\frac{2x+1}{2N}u\pi\right) \cos\left(\frac{2y+1}{2N}v\pi\right) \quad c(k) = \begin{cases} \frac{1}{\sqrt{2}} & \text{pro } k = 0 \\ 1 & \text{jinde} \end{cases}$$

- 3) Z jednotlivých DCT koeficientů je následně spočítána mocnina:

$$E(u,v) = F(u,v)^2$$

- 4) P-nejvyšších umocněných DCT koeficientů je poté použito jako vizuální příznaky, které většinou dobře popisují charakteristickou informaci v daném video snímku.
- 5) Příznakový vektor je vhodné normalizovat, např. je počítán logaritmus z příznakového vektoru.
- 6) Pro dva po sobě následující snímky se poté počítá vzdálenost mezi jednotlivými vizuálními příznaky:

$$\sum_{p=1}^P |VP_i(p) - VP_{i+N}(p)| > T$$

# Vizuální segmentace – metody

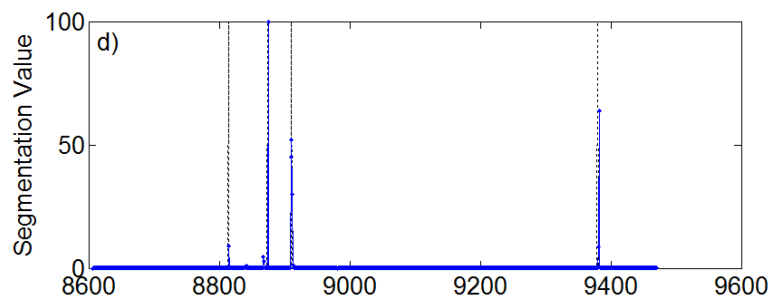
- Metody využívající obrazové příznaky – výpočet z DCT



vzdálenost = 0.274



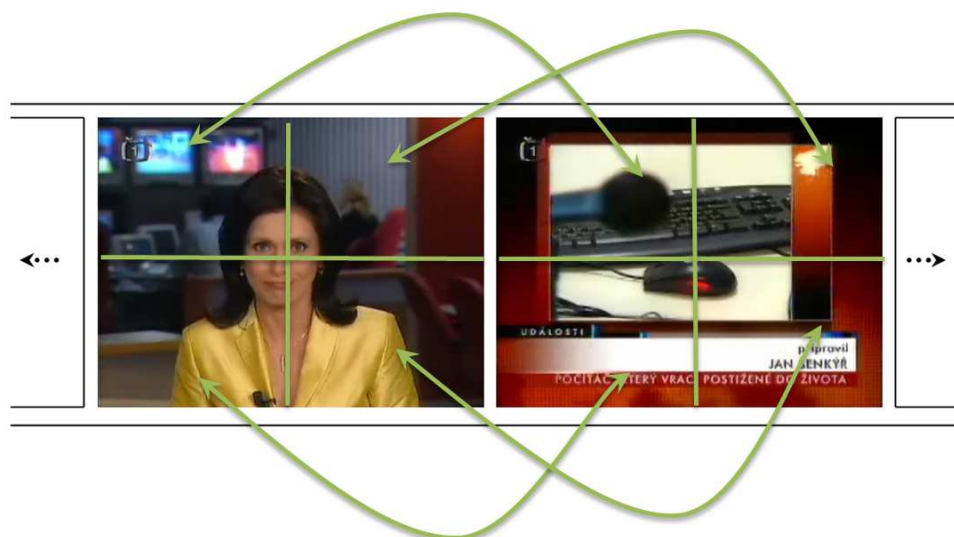
vzdálenost = 21.477



# Vizuální segmentace – metody

- **Blokově zaměřené metody**

- Obraz rozdělen na několik bloků (2x2, 4x4...), rozdíly jsou počítány ze stejných bloků dvou následujících snímků
- Bloky mohou být různě velké, výpočet jenom z několika bloků, jednotlivým blokům lze nastavit různé váhy
- Pro zjištění rozdílu bloků: pixelově- , histogramově- , příznakově-zaměřené vizuálně segmentační metody



# Vizuální segmentace – metody

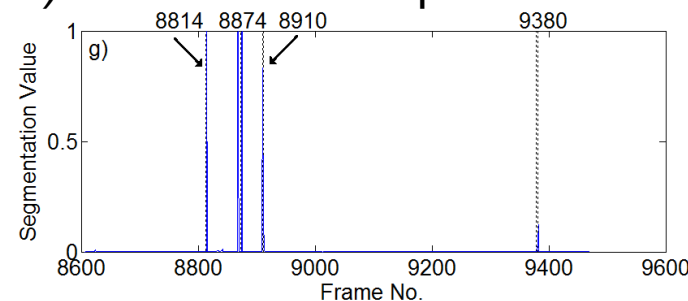
- **Blokově zaměřené metody – využití statistických koeficientů**

- Výpočet statistických veličin v jednotlivých blocích, např. rozptyl a střední hodnota, pro dva korespondující bloky je pak počítána funkce  $L(i, b)$ :

$$L(i, b) = \frac{\left( (\sigma_{i,b}^2 + \sigma_{i-N,b}^2) / 2 + ((\mu_{i,b} - \mu_{i-N,b}) / 2)^2 \right)^2}{\sigma_{i,b}^2 \sigma_{i-N,b}^2}$$

- kde  $\sigma_{i,b}^2$  je rozptyl a  $\mu_{i,b}$  je střední hodnota barevných hodnot v jednotlivých blocích  $b$  v  $i$ -tém snímku
- Hodnota  $L(i, b)$  je následně porovnávána s určitým prahem  $T_b$ . Pokud je větší než tento práh, tak  $L(i, b) = 1$ , jinak  $L(i, b) = 0$ . Kritérium pro detekování hranice vizuálního segmentu je poté:

$$\sum_{b=1}^B w_b L(i, b) > T$$



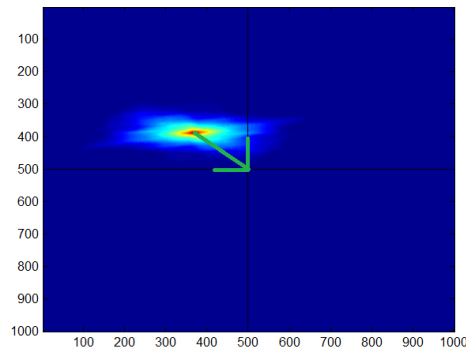
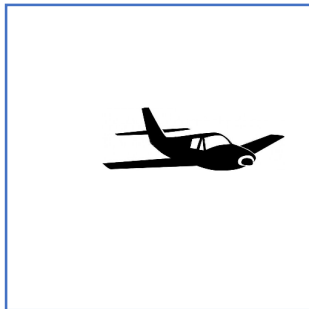
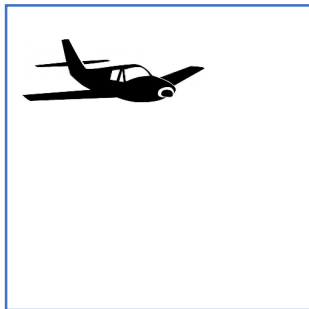
- kde  $B$  je počet bloků,  $T$  je práh a  $w_b$  je hodnota váhy pro jednotlivé bloky
- Nevýhodou u této metody je nutnost stanovení dvou prahů  $T$  a  $T_b$ , kritérium



# Vizuální segmentace – metody

## ● Další metody

- Kombinace předchozích metod, urychlení >>> segmentace hrubá (N následujících snímků) a jemná v oblastech změn
- Metody vycházející ze změny pohybu >>> 1) rozdělení obrazu na bloky, 2) výpočet pohybových vektorů v každém bloku (v dvou následujících snímcích), 3) analýza pohybových vektorů

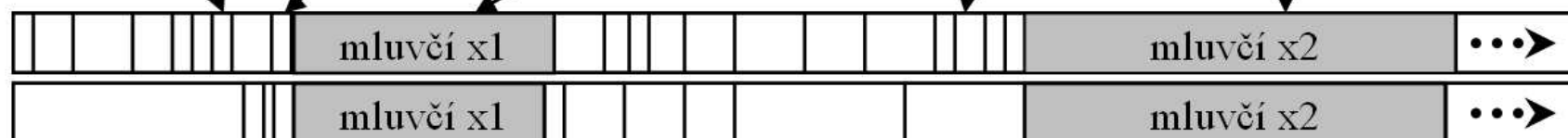


# Vizuální segmentace TV nahrávek

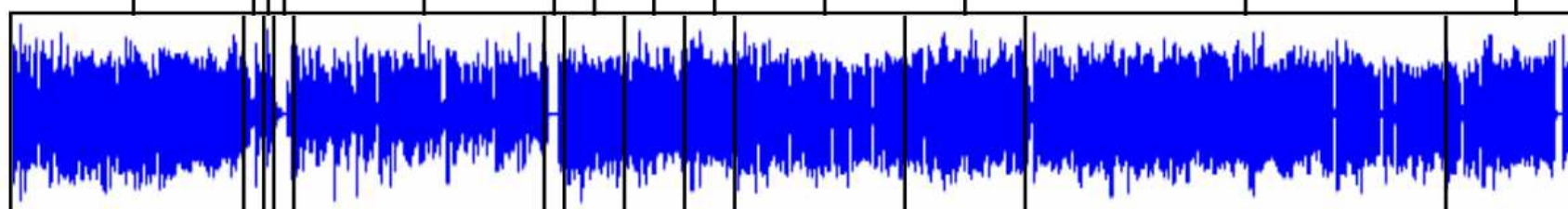
## Vizuální signál



## Vizuální segmenty

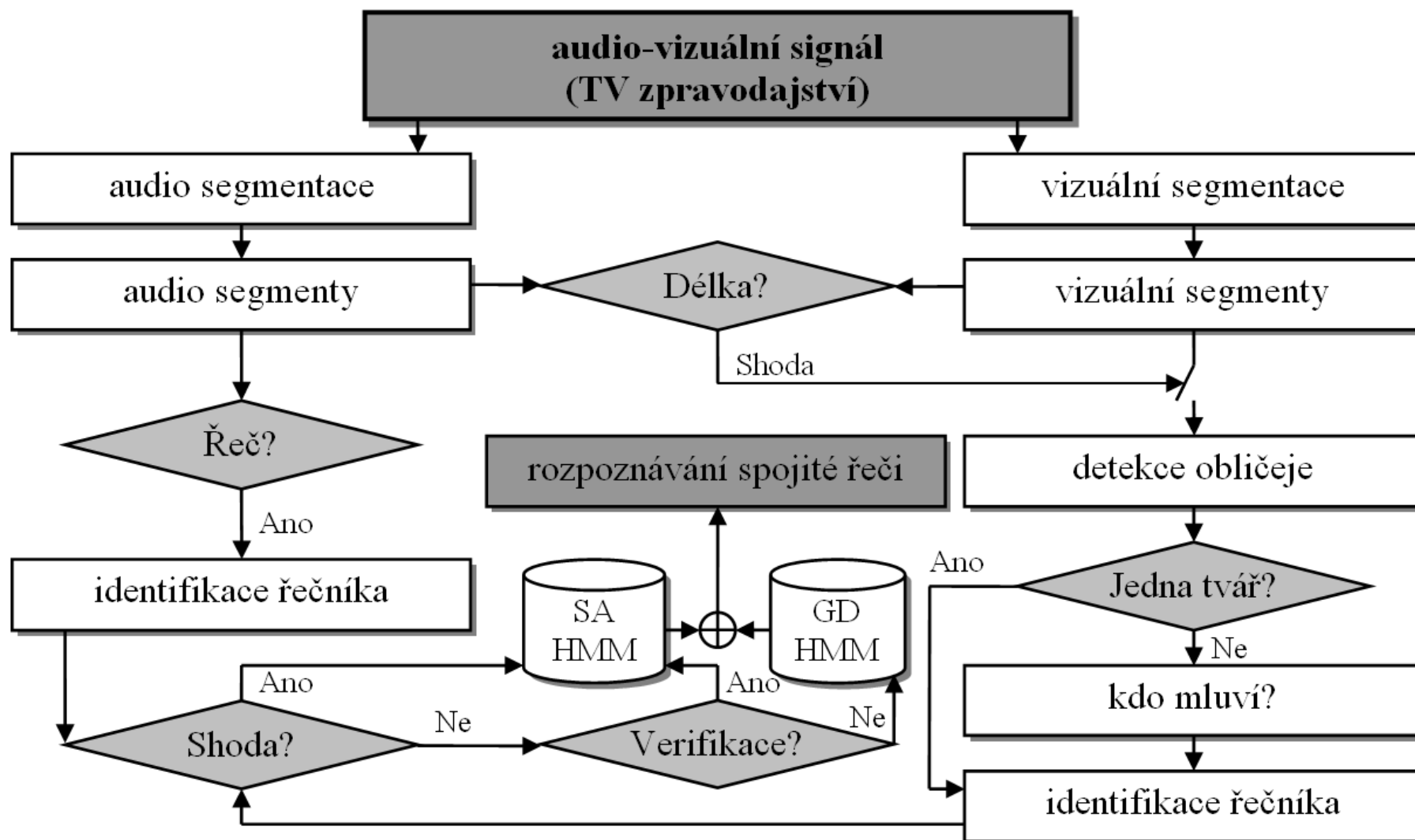


## Audio segmenty



## Audio signál

# Vizuální segmentace TV nahrávek



# Camshift

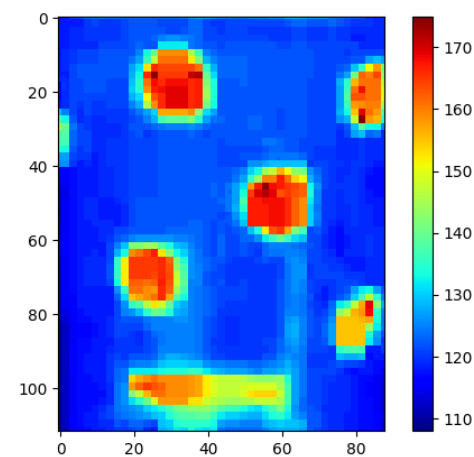
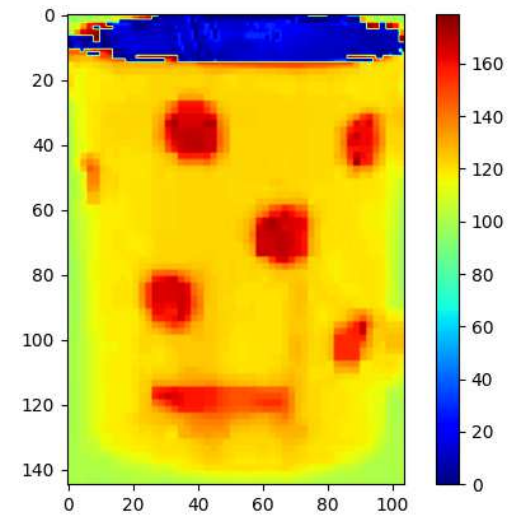
- Jednoduché sledování objektů v obraze





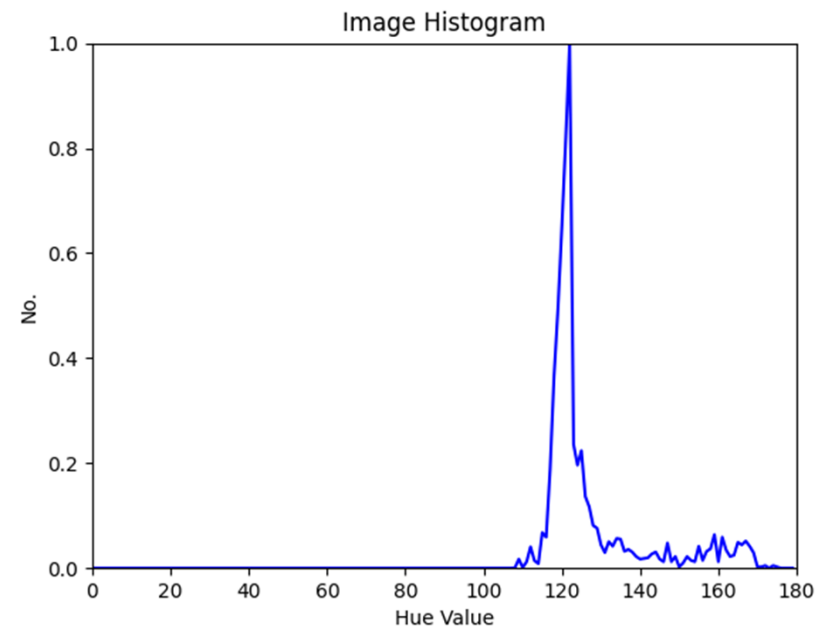
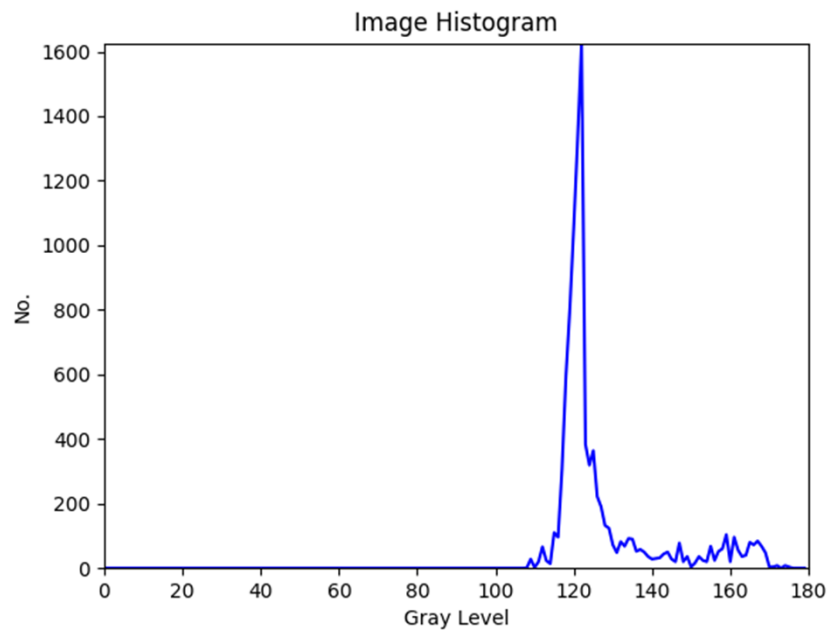
# Camshift

- 1) Výběr sledované barevné oblasti >>> obraz  $I(x, y)$ , převod do HUE



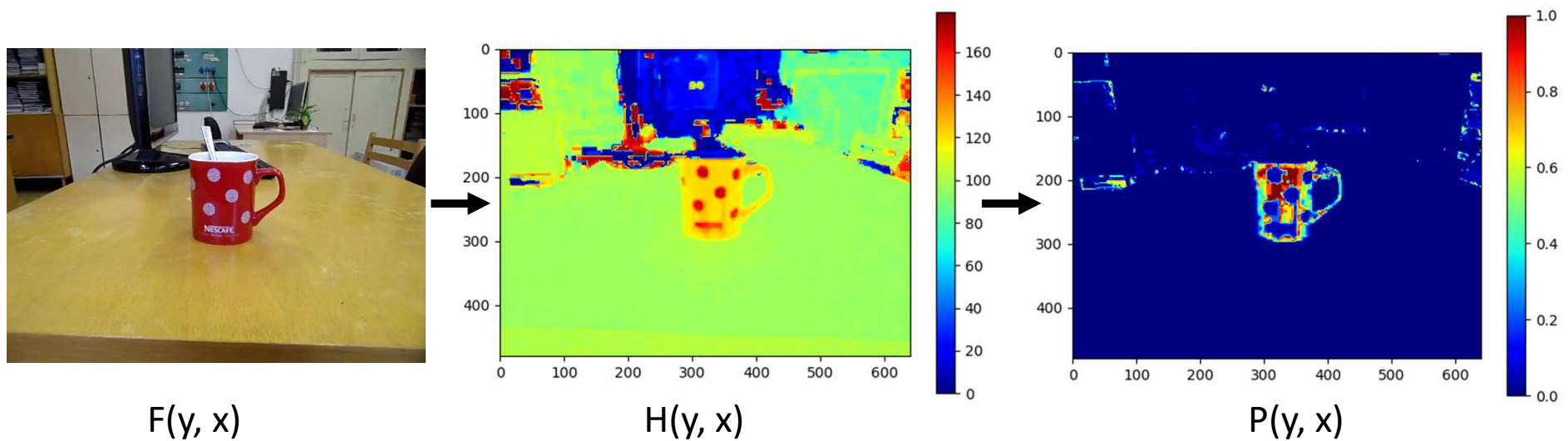
# Camshift

- 2) Výpočet histogramu ze sledované oblasti >>> *hist*



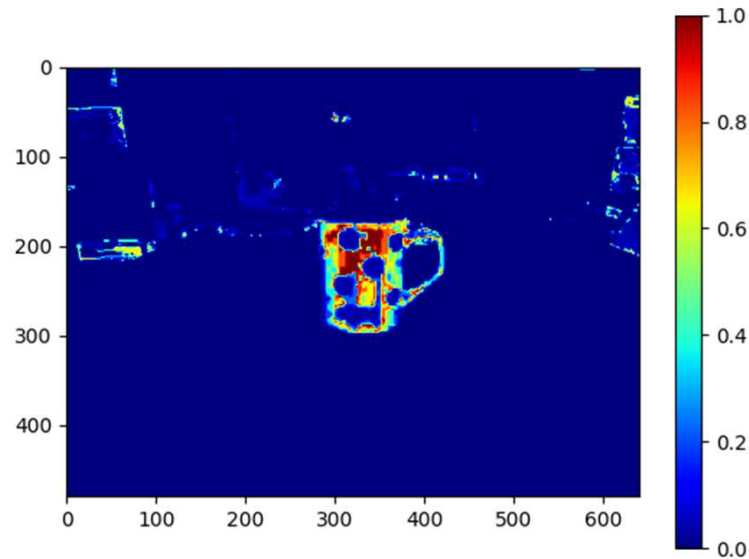
# Camshift

- 3) První video snímek  $F(y, x) \gg \gg$  převod do Hue:  $H(y, x) \gg \gg$  přepočet:  
 $P(y, x) = \text{hist}[h(y, x)]$



# Camshift

- 4) První video snímek – výpočet těžiště + oblast zájmu (ROI)



$$x_t = \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x P(x, y)}{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} P(x, y)}$$

$$y_t = \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} y P(x, y)}{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} P(x, y)}$$

# Camshift

- 5) Další video snímky – výpočet jen z oblast zájmu (ROI) v předchozím snímku



snímek č. 01



snímek č. 02



snímek č. 150