

Odečítání pozadí a sledování lidí z nehybné kamery

Ondřej Šerý

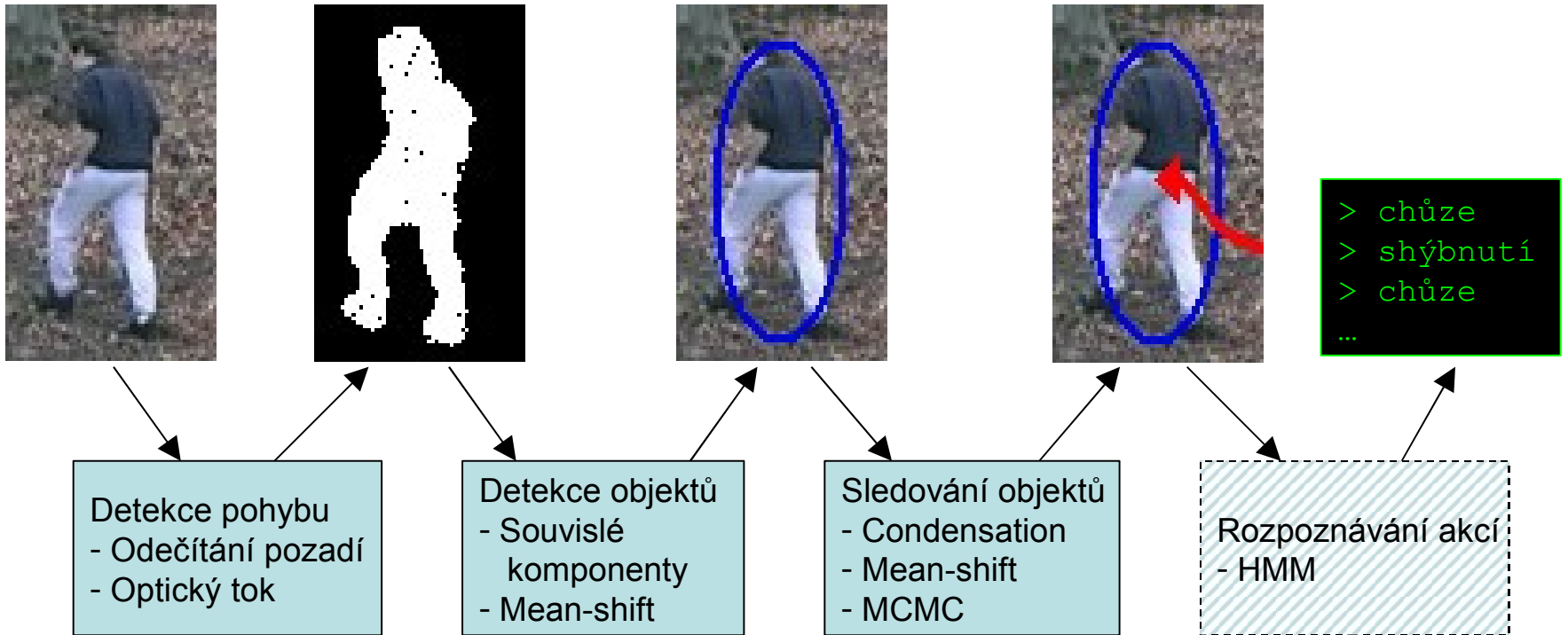
Plán

- Motivace a popis úlohy
- Rozdělení úlohy na tři části
 - Detekce pohybu
 - Detekce objektů
 - Sledování objektů
- Rozbor každé z částí a nástin několika používaných algoritmů
- Praktická ukáзка

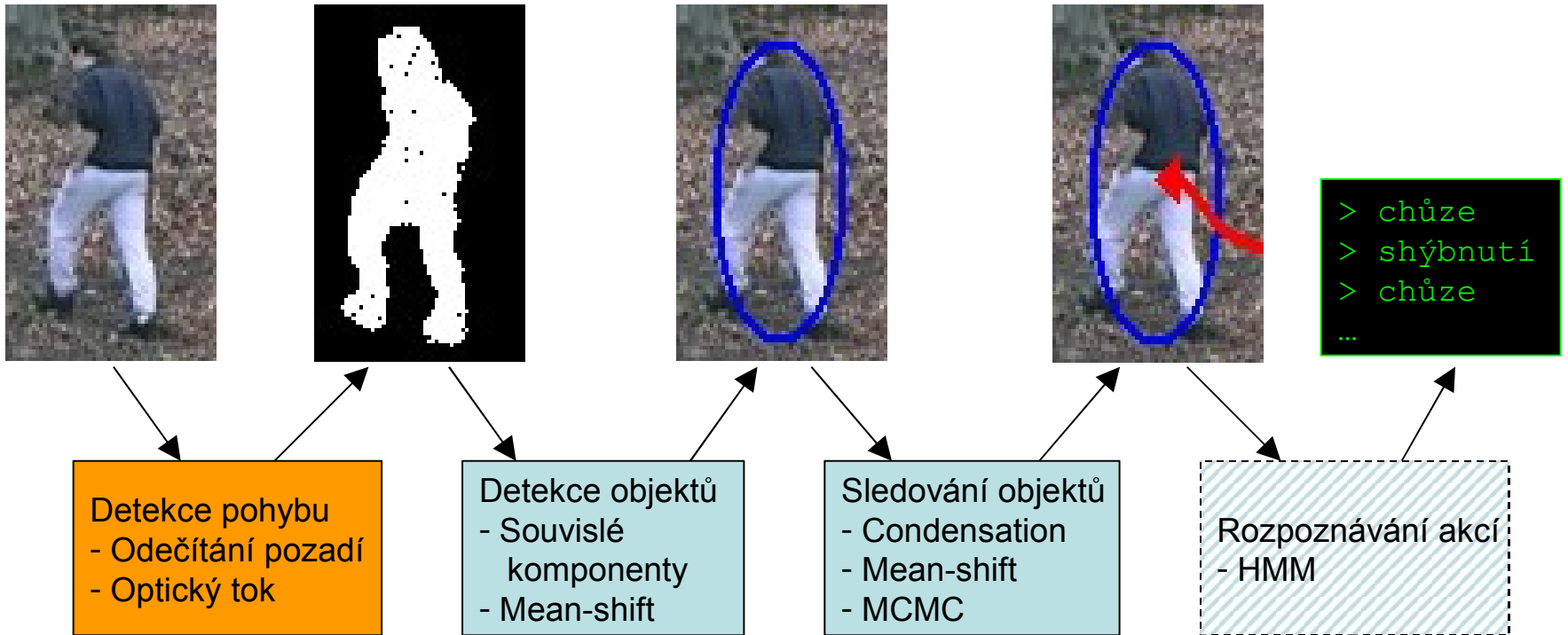
Motivace

- Úkol
 - Detekovat pohyblivé cíle
 - Sledovat cíle při pohybu scénou
 - Zjistit aktivitu cílů
- Motivace
 - Zabezpečovací systémy
 - Automatické zpracování
 - Menší přenos dat (přenos jen důležitých částí)
 - Sledování bezpečnosti práce
 - Továrny – narušení bezpečnostních zón
 - Letiště – sledování letištní plochy
 - Metro – detekce podezřelého chování
 - Statistiky a řízení
 - Řízení křižovatek

Rozdělení úlohy



Detekce pohybu



Detekce pohybu

- Vstup
 - Aktuální snímek
- Výstup
 - Binární obrázek: 0-pozadí, 1-popředí
 - Pravděpodobnostní obrázek: $[0,1]$ – pravd'. příslušnosti k popředí
- Metody
 - Modelování pozadí
 - Optický tok

Detekce pohybu – modelování pozadí

- Máme model pozadí
 - Typicky na úrovni pixelů
 - Nejpravděpodobnější barva v daném místě a v daném okamžiku
 - Aproximace rozložení pravděpodobnosti
- Porovnáme s aktuálním snímkem
 - rozdíl od pozadí
 - pravděpodobnost náležení pixelu k pozadí

Detekce pohybu – modelování pozadí

- Model
 - Statický
 - Předpočítaný v učící fázi
 - Nevýhody:
 - Je třeba získat sekvenci bez popředí
 - Nedokáže reagovat na dlouhotrvající změny scény
 - Špatně reaguje na změny osvětlení
 - Dynamický
 - Model se po dobu běhu výpočtu obměňuje
 - Nevýhody:
 - Problém s nastavením parametrů rychlosti učení
 - Duchové, splynutí s pozadím, nedostatečné přizpůsobování

Stauffer & Grimson [1998]

- Dynamický model
- Model pozadí v jednom pixelu reprezentován směsí n gausiánů
 - n obvykle 3 až 5
 - Každý gausián má přiřazenu váhu w_i , střední hodnotu μ_i a rozptyl σ_i^2
 - Pravděpodobnost náležení pixelu do pozadí

$$P([r, g, b]) = \sum_{i=1}^n w_i \eta([r, g, b], \mu_i, \sigma_i^2)$$

$$\eta([r, g, b], \mu, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{\frac{3}{2}}} e^{-\frac{(r-\mu_r)^2 + (g-\mu_g)^2 + (b-\mu_b)^2}{2\sigma^2}}$$

Stauffer & Grimson [1998]

- Obměna parametrů

- Pixel v aktuálním snímku je přiřazen některému gausiánu pokud:

$$\| [r, g, b] - [\mu_r, \mu_g, \mu_b] \| < 2.5\sigma$$

- Změna vah:

$$w_i := (1 - \alpha)w_i + \alpha * M_i$$

$$M_i = \begin{cases} 1 & i\text{-tý gausián reprezentuje tento pixel} \\ 0 & \text{jinak} \end{cases}$$

- Změna parametrů vybraného gausiánu:

$$[\mu_r, \mu_g, \mu_b] := (1 - \rho)[\mu_r, \mu_g, \mu_b] + \rho[r, g, b]$$

$$\sigma^2 := (1 - \rho)\sigma^2 + \rho((r - \mu_r)^2 + (g - \mu_g)^2 + (b - \mu_b)^2)$$

$$\rho = \alpha\eta([r, g, b], \mu, \sigma^2)$$

Stauffer & Grimson [1998]

- Přiřazení pixelu k pozadí nebo popředí

- Uspořádání gausiánů sestupně podle $\frac{w_i}{\sigma_i}$

- Necht':

$$B = \min \left\{ b \in \{1, \dots, n\}; \sum_{i=1}^b w_i > T \right\}$$

- Pak řekneme, že prvních B gausiánů reprezentuje pozadí

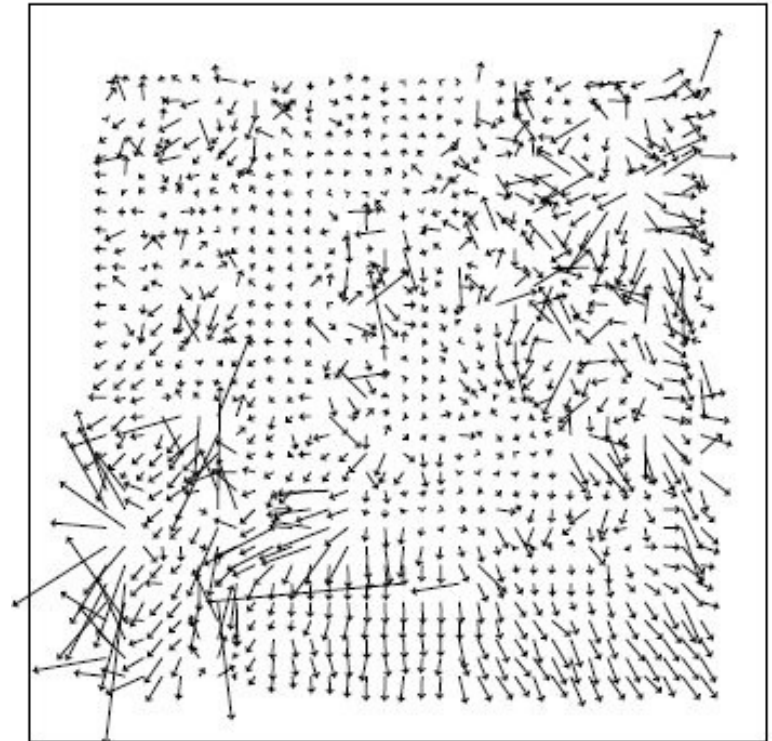
- Předpoklad: pozadí se málo mění a v sekvenci snímků se vyskytuje často
- T je parametr určující práh citlivosti

Další možnosti

- Histogram pozadí
 - Histogram za N posledních snímků pro každý pixel
 - Pravděpodobnost náležení do pozadí se určí jako hodnota v příslušné buňce histogramu
- Přímé odčítání několika snímků
 - Pro každý pixel se určí pravděpodobnost náležení do popředí jako „vzdálenost“ jeho hodnot v několika předcházejících snímcích

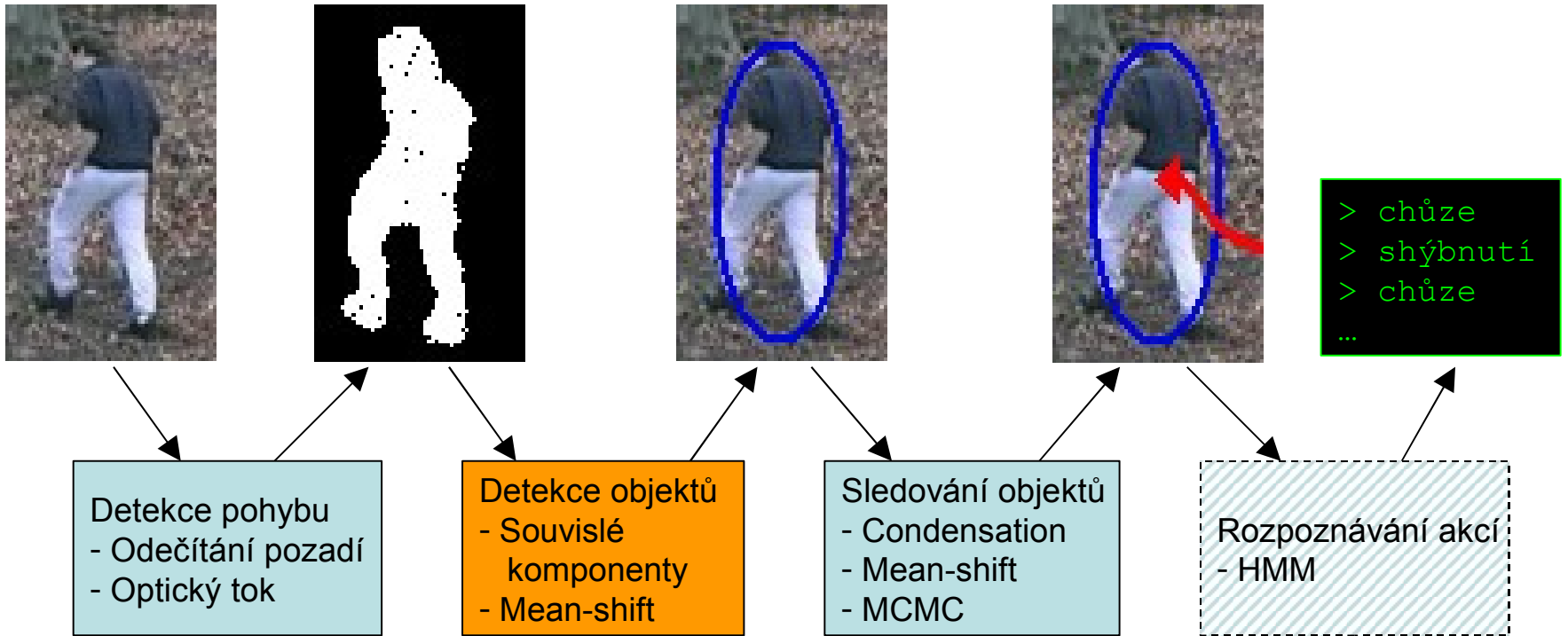
Optický tok

- Snaha nalézt nejpravděpodobnější směr pohybu pro každý pixel (nebo malé oblasti)
- Odtud rozdělení na pozadí a popředí
- Nevýhody:
 - Velice výpočetně náročné
 - Nepřesné
- Výhody:
 - Více informací využitelných i v dalších fázích zpracování



Optický tok (příklad z [1])

Detekce objektů



Detekce objektů

- Vstup
 - Binární obrázek: 0-pozadí, 1-popředí
 - Pravděpodobnostní obrázek: $[0, 1]$ – pravd'. příslušnosti k popředí
- Výstup
 - Identifikace „zajímavých“ objektů v popředí
 - Zajímavý \approx dostatečně veliký, vhodného tvaru (podle aplikace)
 - Při sledování lidí typicky eliptického tvaru
- Metody
 - Souvislé komponenty
 - Mean-shift

Mean-shift

- Úloha:
 - Množina X bodů v prostoru
 - Vágně
 - Úkolem je najít co největší „chumel“ těchto bodů
 - Lépe
 - Maximalizovat funkci:
$$E(y) = \sum_{x \in X} g(\|x - y\|^2)$$
 - Kde $g(a)$ je klesající nezáporná funkce, nenulová na omezené množině

Mean-shift

- Iterativní algoritmus
 - Zvolíme počáteční bod y
 - V každé iteraci uděláme krok ve směru gradientu

$$\|x - y\|^2 = \sum_i (x_i - y_i)^2$$

$$\frac{\partial \|x - y\|^2}{\partial y_i} = -2(x_i - y_i)$$

$$\frac{\partial E}{\partial y_i} = \sum_{x \in X} -2g'(\|x - y\|^2)(x_i - y_i)$$

$$\nabla E = \sum_{x \in X} -2g'(\|x - y\|^2)(x - y)$$

Mean-shift

- Nová pozice: $\bar{y} = y + \xi$

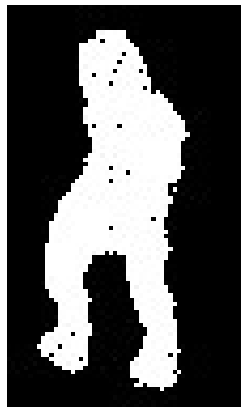
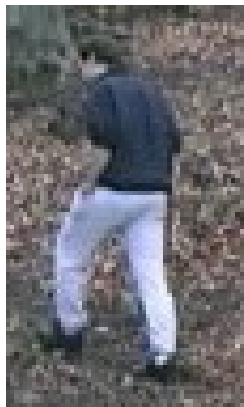
- Kde
$$\xi = \frac{\nabla E}{\sum_{x \in X} -2g'(\|x - y\|^2)} = \frac{\sum_{x \in X} -g'(\|x - y\|^2)(x - y)}{\sum_{x \in X} -g'(\|x - y\|^2)}$$

- Pak novou pozici lze spočítat

$$\begin{aligned} \bar{y} = y + \xi &= y + \frac{\sum_{x \in X} -g'(\|x - y\|^2)x}{\sum_{x \in X} -g'(\|x - y\|^2)} - \frac{\sum_{x \in X} -g'(\|x - y\|^2)y}{\sum_{x \in X} -g'(\|x - y\|^2)} = \\ &= \frac{\sum_{x \in X} -g'(\|x - y\|^2)x}{\sum_{x \in X} -g'(\|x - y\|^2)} \end{aligned}$$

- Což není nic jiného než střední hodnota

Sledování objektů



```
> chůze  
> shýbnutí  
> chůze  
...
```

Detekce pohybu
- Odečítání pozadí
- Optický tok

Detekce objektů
- Souvislé
komponenty
- Mean-shift

Sledování objektů
- Condensation
- Mean-shift
- MCMC

Rozpoznávání akcí
- HMM

Sledování objektů

- Vstup:
 - Pozice sledovaných objektů v minulém snímku
 - Aktuální snímek
 - Modely objektů
- Výstup:
 - Poloha sledovaných objektů v tomto snímku
 - Ztracené objekty
- Metody:
 - Condensation
 - Mean-shift

Používané modely

- Tvar
 - Eliptický
 - Obdélníkový
 - Skvrny (blobs)
- Určení podobnosti
 - Histogram
 - Z celé oblasti
 - Vážený histogram
 - Více histogramů (rozdělení oblasti)
 - Míra podobnosti
 - Bhattacharyya: $\rho(p, q) = \sum_{u=1}^m \sqrt{p_u q_u}$
 - další...

Condensation

- Sledování množiny N potenciálních pozic X_i pro daný cíl a jejich míry podobnosti s modelem Π_i
- Algoritmus:
 1. Výběr N cílů z množiny potenciálních cílů podle pravděpodobnosti Π_i
 2. Změna parametrů nově vybraných potenciálních cílů
 - Náhodně
 - Podle odhadu rychlosti
 3. Přepočítání pravděpodobností Π_i
 4. Pozice cíle v tomto snímku je vážený průměr množiny N potenciálních cílů s váhami Π_i

Mean-shift

- Vážený histogram s lineárním jádrem $g(a) = \begin{cases} 1 - a & a < 1 \\ 0 & \text{jinak} \end{cases}$

$$p_u(y) = C \sum_x g(\|x - y\|^2) \delta(b(x), u);$$

kde $b(x)$ index do hist., δ Kroneckerova delta

- Podobnost histogramů – Bhattacharyya: $\rho(p, q) = \sum_{u=1}^m \sqrt{p_u q_u}$

- Taylorův rozvoj kolem $p(y_0)$

$$\rho(p(y), q) \approx \sum_{u=1}^m \sqrt{p_u(y_0) q_u} + \sum_{u=1}^m \frac{q_u (p_u(y) - p_u(y_0))}{2 \sqrt{p_u(y_0) q_u}} = \frac{1}{2} \sum_{u=1}^m \sqrt{p_u(y_0) q_u} + \sum_{u=1}^m p_u(y) \sqrt{\frac{q_u}{p_u(y_0)}}$$

- Maximalizace podobnosti znamená maximalizaci druhého členu
– Použití alg. mean-shift

$$\sum_{u=1}^m p_u(y) \sqrt{\frac{q_u}{p_u(y_0)}} = C \sum_x w_x g(\|x - y\|^2); \quad \text{kde } w_x = \sum_{u=1}^m \delta(b(x), u) \sqrt{\frac{q_u}{p_u(y_0)}}$$

Mean-shift

1. Startovní pozice y_0 je koncová pozice v minulém snímku. Spočteme $\{p_u(y_0)\}_{u=1}^m$
2. Spočítáme váhy w_x a novou pozici y_1 jako těžiště s váhou w_x
3. Spočítáme $\{p_u(y_1)\}_{u=1}^m$
4. Dokud $\rho(p(y_1), q) < \rho(p(y_0), q)$
dosad' $y_1 := \frac{1}{2}(y_0 + y_1)$ a jdi na 3.
6. $y_0 := y_1$ dokud nezkonvergovalo jdi na 1.
7. y_0 je konečná pozice pro tento snímek

Literatura

- C. BELEZNAI; B. FRUSTUCK; H. BISCHOF; W. KROPATSCH. Detecting humans in groups using fast mean shift procedure. *28th Workshop of the Austrian Association for Pattern Recognition*, pages 71-78, June 1998
- D. COMANICIU; V. RAMESH. Mean-shift and optimal prediction for object tracking. *IEEE Int. Conf. Image Processing*, 2000
- M. ISARD; A. BLAKE. Condensation – conditional density propagation for visual tracking. *International Journal on Computer Vision*, pages 5-28, 1998
- C. STAUFFER; E. GRIMSON. Adaptive background mixture model for real-time tracking. *CVPR 1998*

Ukázky...