

Background Subtraction

Ing. Lukáš Bureš
Katedra kybernetiky

Plzeň
26. června 2014

1 Metoda odečítání pozadí

Metoda odečítání pozadí je technika v oblasti zpracování obrazu a počítačového vidění, při níž dochází k extrakci popředí pro další zpracování (například rozpoznávání objektů atd.). Obecně tvoří oblasti zájmu objekty (lidé, auta atd.), které jsou v popředí. Po fázi předzpracování, která může například obsahovat odstranění šumu, je nasazena metoda odečítání pozadí a její výstupy mohou sloužit jako vstupy pro další metody lokalizace objektu. Odečítání pozadí je široce používaný přístup pro detekci pohybujících se objektů ve videích ze statických kamer. Princip metody spočívá v detekci pohybujících se objektů z rozdílu mezi aktuálním snímkem a referenčním snímkem, kterým může být například obrázek pozadí. Tato metoda se většinou používá v případě, že je snímek součástí video sekvence. V této kapitole byly čerpány informace z [1], [2] a [3].

1.1 Obecný přístup

V této podkapitole budou stručně nastíněny některé základní přístupy k metodě odečítání pozadí, dále zde bude označováno popředí jako F (Foreground) a pozadí obdobně jako B (Background). Všechny metody spojuje základní myšlenka, spočívající ve statickém pozadí a pohybujících se objektech v popředí. Za předpokladu, že pohybující se objekt má v čase t barvu (nebo rozložení barev) lišící se od pozadí B může být tento princip shrnut do následujícího vzorce,

$$F_t(s) = \begin{cases} 1 & \text{když } d(I_{s,t}, B_s) > \tau, \\ 0 & \text{jinak.} \end{cases} \quad (1)$$

Kde $F_t(s)$ je popředí F v čase t na pozici pixelu s , $d(I_{s,t}, B_s)$ označuje vzdálenost mezi aktuálním obrázkem I v čase t na pozici pixelu s a obrázkem pozadí B na pozici pixelu s , τ je hodnota prahu. Největší rozdíl mezi většinou metod odečítání pozadí je, jakým způsobem je modelováno pozadí B a jaká vzdálenostní metrika je použita pro výpočet d . V následujících podkapitolách bude představeno několik základních metod.

1.2 Základní přístup

Nejjednodušší cesta jak lze získat model pozadí B je použít obrázek ve stupních šedi nebo barevný obrázek, který neobsahuje žádné pohybující se objekty v popředí F . Za účelem udržení stále aktuálního modelu pozadí může být využito následujícího iterativního postupu aktualizace modelu pozadí B ,

$$\mathbf{B}_{s,t+1} = (1 - \alpha)\mathbf{B}_{s,t} + \alpha\mathbf{I}_{s,t}, \quad (2)$$

kde α je aktualizací konstanta, jejíž hodnota může nabývat hodnot $\alpha \in \langle 0, 1 \rangle$. V extrémních případech pro $\alpha = 1$ nebude brán model pozadí $B_{s,t}$ v potaz, naopak v případě $\alpha = 0$ se model pozadí nebude aktualizovat. Pixely náležící do popředí mohou být detekovány prahem s různými metrikami, například

$$d_0 = |I_{s,t} - B_{s,t}|, \quad (3)$$

$$d_1 = |I_{s,t}^R - B_{s,t}^R| + |I_{s,t}^G - B_{s,t}^G| + |I_{s,t}^B - B_{s,t}^B|, \quad (4)$$

$$d_2 = (I_{s,t}^R - B_{s,t}^R)^2 + (I_{s,t}^G - B_{s,t}^G)^2 + (I_{s,t}^B - B_{s,t}^B)^2, \quad (5)$$

$$d_\infty = \max\{|I_{s,t}^R - B_{s,t}^R|, |I_{s,t}^G - B_{s,t}^G|, |I_{s,t}^B - B_{s,t}^B|\}, \quad (6)$$

kde exponenty R , G a B reprezentují jednotlivé barevné kanály: červený, zelený a modrý (Red, Green, Blue). Metrika d_0 je speciálně pro obrázky v odstínech šedi.

1.3 Filtrace mediánem

Filtrace mediánem patří mezi nelineární filtraci, která vybírá z blízkého, většinou uživatelem definovaného okolí hodnotu mediánu, kterou dosadí na aktuální pozici. Často se využívá ve fázi předzpracování pro zlepšení výsledků následného zpracování. Tento filtr je velmi účinný při odstraňování zrnitosti/šumu v obraze. Jeho nevýhodou je, že může měnit tvary hran objektů. Po fázi předzpracování se využije postup naznačený v (1.2).

1.4 Gaussovské rozložení

Tato metoda je založená na modelování každého pixelu pozadí pomocí funkce pravděpodobnosti, která je učena pomocí sekvence trénovacích snímků pozadí. V tomto případě základní problém určení prahu přechází v problém nalezení prahu pro funkce pravděpodobnosti. S přihlédnutím k obrazovému šumu, lze trénovat každý pixel pomocí Gaussovského rozložení $\eta(\boldsymbol{\mu}_{s,t}, \Sigma_{s,t})$, kde $\boldsymbol{\mu}_{s,t}$ jsou střední hodnoty pixelu s do času t a $\Sigma_{s,t}$ je kovarianční matice pixelu s do času t . Gaussovské rozložení $\eta(\boldsymbol{\mu}_{s,t}, \Sigma_{s,t})$ v tomto případě má tvar

$$\eta(\mathbf{I}_{s,t}, \boldsymbol{\mu}_{s,t}, \Sigma_{s,t}) = \frac{1}{(2\pi)^{\frac{K}{2}} |\Sigma|^{-\frac{1}{2}}} \cdot e^{-\frac{1}{2}(\mathbf{I}_{s,t} - \boldsymbol{\mu}_{s,t})^\top \Sigma_{s,t}^{-1} (\mathbf{I}_{s,t} - \boldsymbol{\mu}_{s,t})}, \quad (7)$$

kde K je dimenze. Metriku lze určit pomocí Mahalanobisovi vzdálenosti

$$d(\mathbf{I}_{s,t}, \boldsymbol{\mu}_{s,t}) = \sqrt{(\mathbf{I}_{s,t} - \boldsymbol{\mu}_{s,t})^\top \Sigma_{s,t}^{-1} (\mathbf{I}_{s,t} - \boldsymbol{\mu}_{s,t})}, \quad (8)$$

kde $\mathbf{I}_{s,t}$ a $\boldsymbol{\mu}_{s,t}$ jsou vektory a $\Sigma_{s,t}$ je kovarianční matice. Aby byly zohledněny změny osvětlení, tak se střední hodnoty a kovarianční matice mohou iterativně přepočítávat podle následujících vztahů

$$\boldsymbol{\mu}_{s,t+1} = (1 - \alpha) \boldsymbol{\mu}_{s,t} + \alpha \mathbf{I}_{s,t}, \quad (9)$$

$$\Sigma_{s,t+1} = (1 - \alpha) \Sigma_{s,t} + \alpha (\mathbf{I}_{s,t} - \boldsymbol{\mu}_{s,t})^\top (\mathbf{I}_{s,t} - \boldsymbol{\mu}_{s,t}). \quad (10)$$

Kovarianční matice může být plná s velikostí $K \times K$ nebo může být pouze diagonální z důvodu ušetření výpočetního času. V případě, že se jedná o barevný třidimenzionální RGB prostor bude $K = 3$.

1.5 Gaussovská směs (GMM)

Pro zohlednění pozadí, které obsahuje animované textury, jako například vlny na vodě, nebo listy zmítané ve větru, byly použity Gaussovské směsi (Gaussian Mixture Model - GMM). Princip spočívá v modelování každého pixelu pozadí jako K Gaussovských směsí. Tedy pravděpodobnost výskytu barvy v daném pixelu je zastoupena následovně,

$$P(I_{s,t}) = \sum_{i=1}^K \omega_{i,s,t} \cdot \eta(I_{s,t}, \boldsymbol{\mu}_{i,s,t}, \Sigma_{i,s,t}), \quad (11)$$

kde $\eta(I_{s,t}, \boldsymbol{\mu}_{i,s,t}, \Sigma_{i,s,t})$ je i -tý Gaussovský model a $\omega_{i,s,t}$ je jeho váha. Jednotlivé parametry se mohou iterativně měnit podle vztahů,

$$\omega_{i,s,t} = (1 - \alpha)\omega_{i,s,t-1} + \alpha, \quad (12)$$

$$\boldsymbol{\mu}_{i,s,t} = (1 - \rho)\boldsymbol{\mu}_{i,s,t-1} + \rho I_{i,s,t}, \quad (13)$$

$$\Sigma_{i,s,t} = (1 - \rho)\Sigma_{i,s,t} + \rho (I_{i,s,t} - \boldsymbol{\mu}_{i,s,t})^\top (I_{i,s,t} - \boldsymbol{\mu}_{i,s,t}), \quad (14)$$

kde α je uživatelem volený parametr učení a ρ je uživatelem druhý volený parametr, který je definován jako

$$\rho = \alpha \cdot \eta(I_{s,t}, \boldsymbol{\mu}_{i,s,t}, \Sigma_{i,s,t}). \quad (15)$$

Parametry μ a σ nevyužitých rozdělení zůstávají stejné, zatímco jejich váhy jsou redukovány podle vztahu $\omega_{i,s,t} = (1 - \alpha)\omega_{i,s,t-1}$, pro dosažení nulového ovlivnění výsledné pravděpodobnosti. Pokaždé, když neodpovídá žádná komponenta $I_{s,t}$, tak je rozložení s nejmenší váhou nahrazenou Gaussovským rozložením s velkou počáteční variancí σ_0 a malou počáteční váhou ω_0 . Při každé aktualizaci Gaussovské směsi je K vah $\omega_{i,s,t}$ normalizováno tak, aby v součtu dávaly 1. Následně je K rozdělení seřazeno podle hodnoty $\omega_{i,s,t}$ nebo $\sigma_{i,s,t}$ a pouze H nejvíce se hodících, je použito jako část pozadí. Následně pixely, které jsou vzdálené více než 2.5 násobek směrodatné odchylky od některého z těchto H rozdělení, jsou teprve označeny za pohybující se objekty v popředí.

1.6 Kernel Density Estimace (KDE)

Neparametrický přístup může být také využit pro modelování Gaussovských směsí. V tomto ohledu byl navržen odhad pomocí Parzen-okénka. Pokud se jedná o barevné snímky z videa, může být jednodimenzionální jádro vytvořeno následovně,

$$P(\mathbf{I}_{s,t}) = \frac{1}{N} \sum_{i=t-N}^{t-1} \prod_{j=\{R,G,B\}} K\left(\frac{(I_{s,t}^j - I_{s,i}^j)}{\sigma_j}\right), \quad (16)$$

kde K je jádro (typicky Gaussovské rozložení) a N je počet předchozích snímků použitých pro odhad. Pixel je označen jako popředí, pokud je nepravděpodobné, že pixel pochází z tohoto rozdělení. Tedy pokud je $P(\mathbf{I}_{s,t})$ menší než předdefinovaná prahová hodnota. Kde σ_j může být neměnná nebo může být odhadnuta předem.

1.7 Modelování histogramem

Mezi další možnost jak oddělit pozadí od pohybujících se objektů v popředí, pomocí metody odečítání pozadí, je modelování pomocí histogramu. Hlavní myšlenka spočívá v rozdělení obrázku na menší obrazové části, ze kterých se vytvoří 1D histogramy v případě obrázku v odstínech šedi, nebo 3D histogramy v případě barevného obrázku. Tento postup se opakuje pro každý snímek videosekvence pro všechny obrazové části. Na základě porovnání dvou snímků lze určit, zda nastala změna v histogramech a vyhodnotit pohybující se objekt. Tedy pokud H_1 a H_2 označují histogramy ve stejné obrazové části ve dvou snímcích, je možno vzdálenost $d(H_1, H_2)$ určit pomocí Pearsonovi korelace r_{H_1, H_2}

$$d(H_1, H_2) = 1 - r_{H_1, H_2}, \quad (17)$$

kde

$$r_{H_1, H_2} = \frac{\sum_{i=1}^N (H_1^i - \bar{H}_1) (H_2^i - \bar{H}_2)}{\sqrt{\sum_{i=1}^N (H_1^i - \bar{H}_1)^2 \cdot \sum_{i=1}^N (H_2^i - \bar{H}_2)^2}}. \quad (18)$$

\bar{H}_1 a \bar{H}_2 jsou střední hodnoty

$$\bar{H}_j = \frac{1}{N} \sum_{i=1}^N H_j^i, \quad j \in \{1, 2\} \quad (19)$$

a N je počet intervalů (bins) umocněný na počet dimenzí histogramu (v případě, že je počet intervalů ve všech dimenzích stejný). Mezi další možné metriky patří Chi-Square vzdálenost

$$d(H_1, H_2) = \sum_{i=1}^N \frac{(H_1^i - H_2^i)^2}{H_1^i} \quad (20)$$

a Bhattacharyyaova vzdálenost

$$d(H_1, H_2) = \sqrt{1 - \frac{1}{\bar{H}_1 \bar{H}_2 N^2} \sum_{i=1}^N \sqrt{H_1^i H_2^i}}. \quad (21)$$

Na závěr se určí prahová hodnota τ , při které pixely patří do popředí a tvoří pohybující se objekty.

Reference

- [1] M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis, and Machine Vision*. Nelson Education Limited, 2008.
- [2] Y. Benezeth, P.-M. Jodoin, B. Emile, H. Laurent, and C. Rosenberger, “Review and evaluation of commonly-implemented background subtraction algorithms,” in *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, 2008, pp. 1–4.
- [3] M. Piccardi, “Background subtraction techniques: a review,” in *Systems, Man and Cybernetics, 2004 IEEE International Conference on*, vol. 4, 2004, pp. 3099–3104 vol.4.