

Sledování objektu v obraze s využitím částicového filtru

Jan Hamrský <xhamrs00@stud.fit.vutbr.cz>

Václav Maliňák <xmalin19@stud.fit.vutbr.cz>

Jan Wozniak <xwozni00@fit.vutbr.cz>

6. ledna 2012

1 Úvod

Cílem našeho projektu bylo vytvořit aplikaci schopnou sledovat pohybující se objekt v obraze pomocí částicového filtru, v našem případě varianty CONDENSATION algoritmu, a tu aplikovat na zvolenou úlohu. Jako tu jsme si vybrali sledování aut na cestě¹. Uživatel aplikace označí v rámci videosekvence oblast, kde se nachází objekt zájmu a aplikace tento objekt sleduje a vizuálně jej označí. Příznaky extrahované z obrazu byly použity HSV histogram, LBP histogram a BRIEF.

2 Teorie

Particle filter

Cílem částicového filtru je odhadnutí stavových proměnných popisující sledovaný objekt na základě naměřených hodnot z dat. Samotné částicové filtrování pak znamená, že je aproximována distribuční funkce neznámých proměnných pomocí částic, které reprezentují různě váhované vzorky distribuční funkce.

¹jako data byly použity videa z projektu http://www.eecs.qmul.ac.uk/~andrea/avss2007_d.html

CONDENSATION algoritmus ([IB98]) se skládá z několika kroků, které jsou cyklicky opakovány po dobu běhu trasování. V každém z nich je na konci kroku k dispozici aproximovaný stav objektu.

První krok, který je teoreticky volán pouze jednou, je inicializační krok, ve kterém je inicializován model objektu a částice s uniformní pravděpodobností. V našem programu je inicializace prováděna vždy při ručním znovuinicializování uživatelem pomocí GUI.

Stejně tak jako v původním CONDENSATION algoritmu je i naše částice ve tvaru (sn, π_n, c_n) , kde s_n je prvek stavového prostoru, π_n pravděpodobnost částice a c_n kumulativní pravděpodobnost.

Resampling

V kroku vzorkování jsou vybrány částice, které nezaniknou a budou distribuovány do dalších kroků algoritmu. Tyto částice vybíráme na základě jejich kumulativní pravděpodobnosti (rostoucí posloupnost čísel s π_n) a náhodně generovaného čísla r s uniformním rozložením.

Dynamický model

Dynamický model určuje způsob, jakým se mění stav částice mezi jednotlivými kroky. Často používanými modely jsou autoregresivní rovnice prvního a druhého řádu, která je určena rovnicí

$$X_n = A_0 * X_{n-1} + A_1 * X_{n-2} + B * w$$

kde X_n je stav objektu v čase n , A_0, A_1, B jsou koeficienty určující dynamiku modelu a w je vektor náhodných standardních normálních proměnných. Takto zapsaná rovnice určuje tzv. „deterministický drift“ a Brownovské pohyby (zašumění).

Observační model

Observační model určuje pravděpodobnost částice. Naším vybraným modelem je

$$p(z_t | x_t = s_t) = e^{-\kappa * D[h_0, h_t]}$$

kde D je metrikou na histogramech (euclidova vzdálenost, Bhattacharya, χ -kvadrát, ...). Koeficient κ lze určit z dat pomocí MLE, ale touto možností jsme se nezabývali.

Odhad

Na konci každého kroku je možné provést odhad stavu, který je váhovanou sumou stavů částic a dostat tak „průměrnou hodnotu“.

Příznaky

Z obrazu jsou extrahovány příznaky pro přesnější sledování objektu. HSV histogram funguje velmi spolehlivě pro barevný obraz avšak nedokáže si poradit s problémem, kdy naprosto odlišný objekt má podobné rozložení barevných složek a proto jsme se rozhodli najít jiný vhodný příznak, který by byl výpočetně nenáročný a reprezentoval by do jisté míry i strukturu hledaného objektu.

Textura oblasti v obraze může v daném problému sloužit jako velmi základní popis reflektující rovněž strukturu objektu, proto jsme se pokusili použít LBP histogram [CLS11]. Zkoušeli jsme dvě varianty LBP(8,1) rotačně invariantní a LBP(8,1) uniformní, a metriky χ^2 , L1, euklidovskou a Bhattacharya. Nejlepší výsledky podávaly uniformní LBP s metrikou L1.

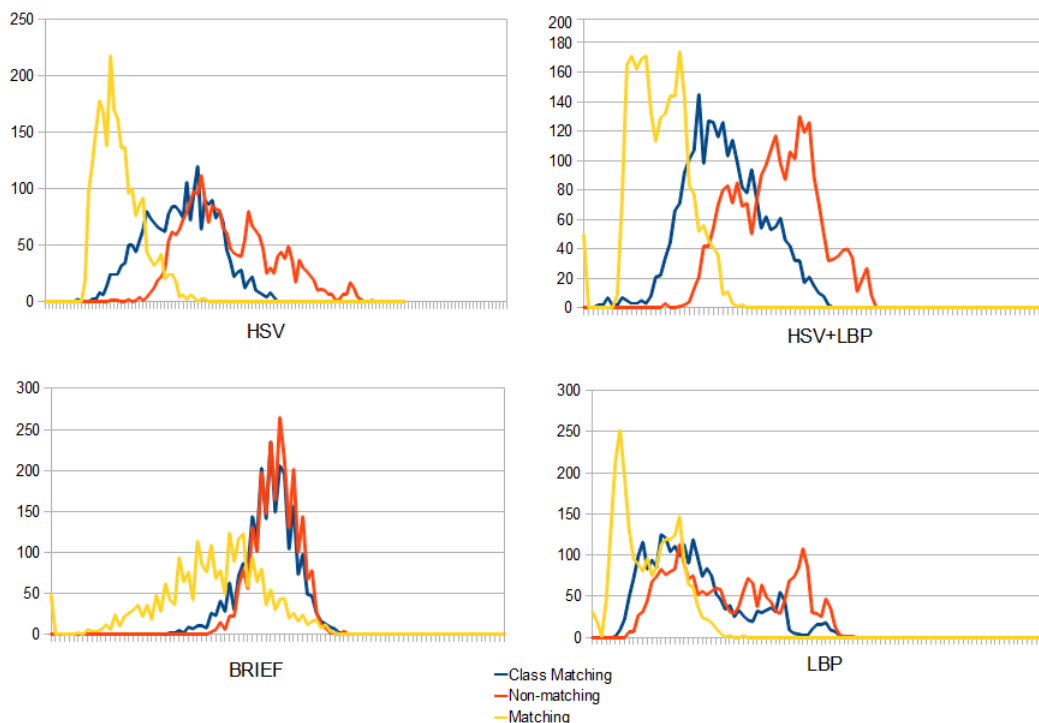
BRIEF [PIP03] je obvykle používán jako deskriptor významných bodů v obraze pracující na principu náhodně vybraných porovnání dvojic bodů v okně pevné velikosti. Testovali jsme základní variantu s uniformním rozložením lokace bodů (x_i, y_i) , tedy $(-\frac{S}{2}, \frac{S}{2})$. Metriku pro BRIEF jsme použili Hammingovu.

3 Vyhodnocení

Pro výběr vhodných příznaků jsme si vytvořili databázi, na které jsme se změřili diskriminabilitu hledaných objektů pomocí turnajové selekce. Cílem našeho projektu bylo sledovat jedoucí automobil v záznamu z kamery na základě výběru uživatele.

Na každém grafu z obrázku 1 jsou tři distribuční křivky, žlutá znázorňuje rozložení vzdáleností mezi příznaky hledaného objektu (obrázky konkrétního vozu), modrá mezi objektem a třídou těchto objektů (obrázky různých vozů), červená mezi objektem a okolím. Jak je z grafů patrné, LBP a HSV byly úspěšnější než BRIEF, proto jsme se pokusili spojit LBP a HSV. Podle distribuce by toto spojení mělo přinést zlepšení.

Pokusili jsme se naimplementovat algoritmus CONDENSATION. Kroky jsme implementovali stejně tak, jako byly popsány v článku. Jako stav ob-



Obrázek 1: Graf distribuce vzdáleností.

jektu jsme zvolili $[x, y]$. V rámci projektu jsme se také pokusili stanovit koeficienty dynamického modelu na základě článku [BIR95]. S pomocí programu *ViPER*² jsme provedli určení *ground truth*. Kvůli námi nezjištěné chybě v programu nebylo možné získané koeficienty použít, protože chyba znemožňovala použití autoregresivní rovnice 2. řádu.

O stanovení koeficientu κ jsme se nepokoušeli.

4 Závěr

HSV příznaky i LBP příznaky samy o sobě by měly rozlišit objekt od okolí. Spojením těchto dvou příznaků bychom měli dosáhnout zlepšení diskriminačních vlastností, to se nám ale v částicovém filtru z časových důvodů nepodařilo ověřit.

Algoritmus CONDENSATION je poměrně efektivní způsob pro sledování

²<http://vipier-toolkit.sourceforge.net/>

objektu v obraze, který funguje v téměř reálném čase v závislosti na velikosti inicializační oblasti a použitých příznamech. V rámci projektu se nám podařilo vytvořit sledování objektu na základě šumového modelu, ne však s použitím autoregresivní funkce 2. řádu. Původně zamýšlenou metriku překrytí detekovaných ploch s ground truth nebylo možné použít kvůli již uvedeným problémům.

Reference

- [BIR95] Andrew Blake, Michael Isard, and David Reynard. Learning to track the visual motion of contours. *Artificial Intelligence*, 78:101–134, 1995.
- [CLSF11] Michael Calonder, Vincent Lepetit, Christoph Strecha, and Pascal Fua. Brief: Binary robust independent elementary features, 2011. [Online; accessed 2-January-2012].
- [IB98] Michael Isard and Andrew Blake. Condensation - conditional density propagation for visual tracking. *International Journal of Computer Vision*, 29:5–28, 1998.
- [PIP03] Matti Pietikäinen, Jukka Iivarinen, and Maria Petrou. The local binary pattern approach to texture analysis, 2003. [Online; accessed 2-January-2012].