

HUMAN DETECTION IN COMMERCIAL APPLICATIONS

Jan Černín

Master Degree Programme (5), FEEC BUT

E-mail: xcerni04@stud.feec.vutbr.cz

Supervised by: Karel Horák

E-mail: horak@feec.vutbr.cz

Abstract: This paper is related to the development and implementation of human tracking algorithm which should be used for commercial applications. The detector is based on Histogram of Oriented Gradients which creates a descriptor. The linear Support Vector Machine is used for classification. Tracking algorithm follow the path of the person with the possibility of prediction.

Keywords: Computer Vision, Human Detection, HOG, SVM, Kalman Tracker

1 ÚVOD

Problematika počítačového vidění v dnešní době proniká stále hlouběji do běžného života. Člověk, ať už chce nebo nechce, se každým dnem dostává více a více pod drobnohled visuálních kamer. Tato práce se zabývá vývojem a implementací algoritmů určených pro robustní sledování osob, vhodné především pro marketingové účely jako jsou například systémy optimalizace prodeje, z anglického point-of-sales (POS).

2 DETEKCE OSOB

Systémy počítačového vidění implementující algoritmy pro detekci osob nachází své uplatnění v širokém spektru aplikací. Algoritmy užívané k detekci se ve své podstatě liší ve způsobu zpracování obrazových dat a v jejich následném popisu. Výzkum posledních let přišel s řadou metod, jejichž nejčastějším základem je extrakce příznaků, které následně tvoří deskriptor. K takovým metodám patří například metoda Normalizovaného obrazového gradientu, která je také základem detekčního algoritmu použitého při zpracování této práce.

3 NORMALIZOVANÝ OBRAZOVÝ GRADIENT

Jedná se o robustní obrazový deskriptor, který využívá přímého pohledu na postavu či skupinu postav nacházející se ve sledované scéně ve vzpřímené poloze. Deskriptor může být statický nebo dynamický. Statický deskriptor poslouží při detekci ve statických snímcích, zatímco dynamický využívá informaci o pohybu objektu a je používán při detekci v sekvenci snímků. Klasifikátorem příznakového prostoru je jednoduchý kaskádový Support Vector Machine (SVM).

3.1 PRINCIP ALGORITMU

Objekt je popsán pomocí lokálního měření gradientu nebo směru hran. Jádrem deskriptoru je pohyblivé okno sestávající se z menších spojených podoken, přes která je histogram směrů gradientů nebo směr hran sestavován. Každé buňce odpovídá rozložení gradientů nebo hran v 1-D oblasti. Kombinace těchto histogramů reprezentuje výsledný deskriptor, který je sestavován přes celý obraz. Rozdělením detekčního okna do sítě překrývajících se normalizovaných bloků a použitím kombinovaného příznakového vektoru k natrénování lineárního SVM vede na řetězec detekce podle obrázku 1.



Obrázek 1: Řetězec exktrakce příznaků a následná detekce (převzato z [1]).

3.2 EXTRAKCE PŘÍZNAKŮ

K sestavení výsledného deskriptoru se hojně užívá metoda SIFT (z anglického "Scale Invariant Feature Transform"). Výsledkem SIFT algoritmu jsou na měřítku nezávislé významné body. Díky svým přednostem je SIFT hojně používán v mnoha detekčních algoritmech [2].

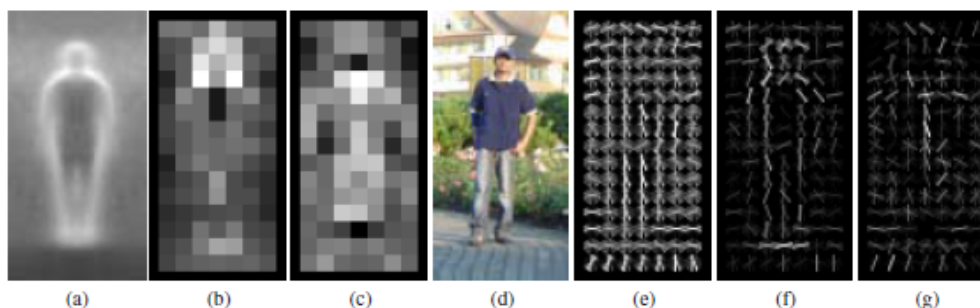
Kombinací HOG/SIFT dostáváme robustní deskriptor, který zohledňuje ty hrany nebo gradienty, nesoucí jedinečnou informaci vzhledem ke svému okolí. Algoritmus je zároveň odolný vůči malým změnám vlivem translace a rotace lidského těla. Tento přístup zároveň těží z toho, že se lidská postava v obraze často nachází ve vzpřímené pozici.

Deskriptor popisující okolí významného bodu se skládá z akumulovaných směrových histogramů přes dané okolí. Velikost složek histogramu je dána příspěvky gradientů tohoto okolí. Pro větší odolnost vůči změně osvětlení a kontrastu je velikost gradientu lokálně normalizovaná. Toho je dosaženo seskupením buněk do větších, prostorově spojených bloků. Deskriptor poté tvoří vektor všech vytvořených histogramů daného bloku [1].

Uvažujeme-li deskriptor s překrytím jednotlivých bloků, dávají koeficienty natrénovaného lineárního SVM míru, jakou jednotlivé buňky přispívají k výslednému rozhodnutí. Na obrázku 2b) a 2f) je vidět, že nejvýznamnější buňky jsou ty které se nacházejí v oblasti hlavy, ramen a dolních končetin (pozitivní váhy). Podobně, obrázek 2c) a 2g), zvýrazňuje gradienty uvnitř postav (negativní váhy).

3.3 SVM KLASIFIKÁTOR

Podstatnou částí detektoru je konečná klasifikace vektorů příznaků do skupin. V tomto případě se jedná o dichotomickou klasifikační úlohu člověk/nečlověk. K tomuto účelu je využito jednoduchého SVM klasifikátoru. Podstatou klasifikátoru je transformace lineárně neseparabilního prostoru příznaků do prostorů s vyšší dimenzí. V takovémto prostoru je poté možné provést separaci pomocí nadrovin. Parametry této nadroviny určují pomocné vektory, které jsou v průběhu učení přizpůsobovány. Po klasifikaci nastává samotné trasování pohybu objektu, viz následující kapitola.



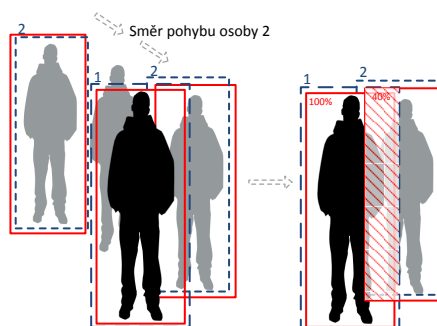
Obrázek 2: HOG detektor založený především na konturách lidského těla (hlava, ramena a dolní končetiny). (a) Průměrný gradient trénovacího vzoru. (b) Pozitivní váhy SVM zvýrazňují gradienty nacházející se v oblast hlavy, ramen a dolních končetin. (c) Negativní váhy SVM zvýrazňují především vertikální gradienty. (d) Testovací obrázek. (e) R-HOG deskriptor testovacího obrázku. (f,g) R-HOG deskriptor váhovaný podle pozitivních resp. negativních SVM vah. (Převzato z [1])

4 TRASOVÁNÍ POHYBU OBJEKTU

Určení pozice objektu je často zatíženo chybou vzniklou šumem a výpadky měření. V takovémto případě je vhodné užít metodu trasování, která do systému vnese potřebnou dynamiku a zároveň je schopná predikce následujícího stavu systému na základě znalosti modelu systému. Pro tento účel byl vybrán jednoduchý Kalmanův filtr. Predikce spočívá v určení pravděpodobného výskytu objektu, jež je popsán souřadnicemi nálezu a jeho velikostí. Při ztrátě měření se pokračuje v predikci pro další krok, dokud nedojde k získání měření, po kterém je provedena korekce filtru [3].

4.1 PROBLÉM KORESPONDENCE

Každý detekovaný objekt vlastní svoji instanci Kalmanova filtru. V tomto případě je třeba zaručit, aby byl detekovaný objekt přiřazen právě té instanci filtru, která je s ním spojena. Tato situace byla vyřešena jednoduchým způsobem. Pro každý filtr je spočítáno procento překrytí se všemi nalezenými objekty. Region, který pokrývá největší procento plochy odhadované pozice objektu (výstup filtru), náleží příslušajícímu filtru, jak ilustruje obrázek 3.



Obrázek 3: Řešení problému korespondence. První osoba (1) stojí na místě, druhá osoba (2) míjí stojící osobu v popředí. Vpravo je spočítána plocha, kterou pokrývají jednotlivá měření predikovanou oblast filtru 1. Filtru koresponduje to měření, které pokrývá největší plochu v procentech. Výstup detektoru - plná čára, výstup filtru - čárkovaná čára.

5 ZÁVĚR

Pro účel sledování osob v komerčních aplikacích byla vybrána metoda orientovaného gradientu HOG, která tvoří výsledný deskriptor. Vektor deskriptoru, který je tvořen přes celý obraz, je klasifikován lineárním SVM, jež je na měřítku nezávislý a vykazuje stejné rychlosti při detekci jedné i více osob. Výslednou trajektorii pohybu tvoří vektor výstupu odhadu Kalmanova filtru, který slouží jako trasovací algoritmus.

REFERENCE

- [1] Dalal, N.: Finding people in images and videos. Grenoble, 2006, 149 s., Disertační práce. Institut National Polytechnique de Grenoble. Vedoucí disertační práce M. William J. Triggs.
- [2] Lowe, D. G.: Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision, 2004, vol. 60, no. 2, pp. 91-110.
- [3] Gutman, P., Velger, M.: Tracking Targets Using Adaptive Kalman Filtering. IEEE Transactions on Aerospace and Electronic Systems, 1990, vol. 26, no. 5, pp. 691-699