

VŠB – Technická univerzita Ostrava
Fakulta elektrotechniky a informatiky
Katedra informatiky

Analýza řidiče za pomoci sférických kamer

Driver Analysis Using Spherical Cameras

Prehlasujem, že som túto diplomovú prácu vypracoval samostatne. Uviedol som všetky literárne
pramene a publikácie, z ktorých som čerpal.

V Ostrave 1. apríla 2020

.....

Súhlasím so zverejnením tejto diplomovej práce podľa požiadaviek čl . 26, odst. 9 Študijného a skúšobného poriadku pre štúdium v bakalárskych programoch VŠB-TU Ostrava.

V Ostrave 1. apríla 2020

.....

Rád by som poďakoval môjmu vedúcemu práce Ing. Radovanovi Fusekovi za pomoc a ochotu pri vypracovaní diplomovej práce

Abstrakt

Hlavnou témou diplomovej práce je rozpoznávanie a analýza vodiča v aute pomocou sférických kamier. Táto práca je rozdelená do viacerých samostatných častí. Prvá časť spočíva v samotnej detekcii ľudí a ich aktivít sférickou kamerou, hľadanie nedostatkov a nájdenie optimálnych parametrov pre čo najefektívnejšiu detekciu. Druhá časť je zameraná na porovnanie jednotlivých knižníc a metód, ktoré sa používajú na analýzu ľudského tela a tváre v obraze. Posledná časť je venovaná porovnaniu týchto metód s použitím reálnych dát zachytených sférickou kamerou a zhrnutie výsledkov.

Kľúčové slová: Sférická kamera, detekcia obrazu, analýza ľudskej tváre, detekcia ľudí, vodič

Abstract

Main focus of this Diploma thesis is detection and analysis of driver in car with help of spherical cameras. This thesis is divided into few parts. The first part is about detection itself, detection of people by spherical cameras, research of disadvantages and finding optimal parameters for most efficient detection. The second part is focused on comparison of libraries used for human body and face detections. The last part is about comparison of libraries with real data captured by spherical camera and summary of results.

Keywords: Spherical camera, image detection, analysis of human face, pedestrian detection, driver

Obsah

Zoznam použitých skratiek a symbolov	7
Zoznam obrázkov	8
1 Úvod	9
2 Detekcia a analýza ľudského tela v obrazoch	10
2.1 Haar	10
2.2 HOG	12
3 Detekcia pozície vodiča	14
3.1 OpenPOSE	15
3.2 TensorFlow	16
3.3 Ostatné metódy	17
4 Využitie sférických kamier v automobiloch	18
4.1 Technické parametre	18
4.2 Použitie v analýze videa	19
5 Program	20
5.1 Požiadavky a návrh programu	21
5.2 Pozícia vodiča	22
5.3 Orientácia hlavy	25
5.4 Výstup programu	26
5.5 Porovnanie výsledkov	27
5.6 Využitie zozbieraných dát	28
5.7 Používateľská príručka	29
6 Možnosti vylepšenia detekcie	30
7 Záver	31
Literatúra	32

Zoznam použitých skratiek a symbolov

CPU	– Central processing unit
FPS	– Frames per second
GPU	– Graphical processing unit
HOG	– Histogram oriented gradients
OpenCV	– Open Source Computer Vision
PX	– Pixel

Zoznam obrázkov

1	Haar - dvoj-obdĺžnikové príznaky (A, B), troj-obdĺžnikové príznaky (C) a štvor-obdĺžnikové príznaky(D) [1]	11
2	HOG - séria krokov [2]	12
3	HOG - vstupný obraz (a), normalizácia gradientu (b), orientácia gradientu (c), rozdelenie do buniek (d) vypočítaný histogram (e). [3]	13

1 Úvod

V dnešnom modernom svete sú autá takmer každodennou súčasťou života ľudí. Mnohokrát sa ani nezamýšľame nad ich bezpečnosťou, ktorá je v prípade zrážky kľúčová. V súčasnosti nám pri jazde autom asistuje veľké množstvo systémov, ktoré zvyšujú bezpečnosť posádky, ale aj ostantých účastníkov cestnej premávky. Aj keď tieto systémy ešte stále nedokážu vodiča úplne nahradiť, dokážu mu výrazným spôsobom pomôcť napríklad v krízových situáciách. Výhodou takýchto systémov je ich rýchlejší reakčný čas oproti človeku. Takéto systémy spočívajú v použití rôznych snímačov alebo kamier, ktoré aktívne sledujú okolie ale aj interiér vozidla. Vďaka takýmto moderným technickým riešeniam je možné predísť rôznym častokrát aj smrteľným dopravným nehodám. Výrobcovia áut sa čoraz častejšie snažia svoje systémy vylepšovať na čo najvyššiu možnú úroveň a poskytnúť tak vysoký level ochrany.

Táto diplomová práca sa zameriava hlavne na problematiku analýzy vodiča pomocou detekcie obrazu zo sférickej (360-stupňovej) kamery. V diplomovej práci som sa venoval analýze videa z kamery umiestnenej v interiéri vozidla. Vhodným umiestnením kamery je možné získať obraz zpred auta, ale aj obraz vodiča sediaceho za volantom. V tejto práci som sa zameriaval na analýzu a spracovanie videa z interiéru vozidla na zachytenie ľudských aktivít vodiča. Aby som získal čo najväčšiu časť tela vodiča, je potrebné mať dostatočne veľký uhol záberu. Bežné kamery majú uhol záberu veľmi nízky, aby dokázal z malej vzdialenosti zachytiť celý snímaný objekt. Takýto problém sa naskytuje najpríklad aj v interiéri vozidla, kde je vzdialenosť kamery od snímaného objektu menej ako 1 meter, čo nemusí byť dostatočné na zosnímanie tela celého vodiča. Práve v takejto situácii je vhodné použiť širokouhlú prípadne sféricкую kameru. Počas práce som mal k dispozícii viaceré kamery, s ktorými som zhotovil niekoľko desiatok videí v rôznych situáciách. Z takýchto videí som dokázal analyzovať a zistiť mnoho užitočných informácií, ktore sú spracované v tejto diplomovej práci. Tieto informácie som zbieral nahrávaním videa sférickými kamerami za rôznych svetelných podmienok a pozícií vodiča. V tejto práci sú taktiež spomenuté problémy takejto analýzy, riešenia vzniknutých problémov, ale aj zhrnutie celkovej problematiky sledovania vodiča vo vozidle. V práci sú tiež zhrnuté ďalšie možnosti vylepšenia detekcie a porovnanie oproti klasickým kamerám.

V nasledujúcich kapitolách je postupne rozobratá problematika snímania ľudských postáv v obrazoch, a skúmanie ich aktivít. Pre snímanie postavy som sa rozhodol použiť viacero metód, ktoré som následne porovnal a zanalyzoval. Aby som vedel vyhodnotiť správnu pozíciu vodiča, rozhodlo sa použiť neurónovú sieť, ktorú som trénoval na vlastnom datasete.

V súčasnosti som taktiež nenašiel veľa riešení na spracovanie videa zo sférickej kamery a preto by som sa snažil zamerať túto prácu hlavne na túto oblasť. Pri analýze vodiča som taktiež nenašiel vhodné datasety z interiéru vozidla snímané sféricickou kamerou.

2 Detekcia a analýza ľudského tela v obrazoch

História detekcie postáv v obrazoch siaha až do polovice 20. storočia. Mnoho inžinierov videlo obrovský potenciál detekcie obrazu napríklad v oblastiach medicíny, priemyslu, dopravy a mnohých ďalších oblastiach. S nárastom technických možností postupne rástla aj motivácia využiť detekciu obrazu aj v praxi. Jeden z prvých vedeckých článkov v oblasti spracovania obrazu [4] rozoberal napríklad jednoduchú analýzu obrazu a spracovanie obrazov s dostupnými prostriedkami. Postupom času sa však počítačová technika vylepšovala a bolo možné pracovať na vývoji metód pre analýzu a detekciu objektov v obrazoch. Na detekciu chodcov alebo iných ľubovoľných objektov existuje mnoho metód. Veľkým fenoménom v posledných rokoch sa stali neurónové siete. Okrem neurónových sietí však stále existujú aj tradičné metódy, ktoré fungujú aj bez tréningových dát. Vo svojej práci som pracoval hlavne s metódami Haar a HOG, ktoré sa radia medzi najpoužívanejšie tradičné metódy a sú im venované samostatné podkapitoly 2.1 a 2.2.

Každý obraz sa skladá z pixelov. Analýza obrazu však nespočíva v prehľadávaní jednotlivých pixelov, ale v hľadaní jednotlivých objektov v obraze. Tieto objekty je možné určovať do samostatných tried. Triedy nám určujú, aký druh objektu sa v obraze nachádza (Napríklad chodec, vozidlo, dopravná značka a podobne). Aby bolo možné tieto objekty (v našom prípade ľudí) nájsť, bolo potrebné nájsť spoľahlivý spôsob detekcie.

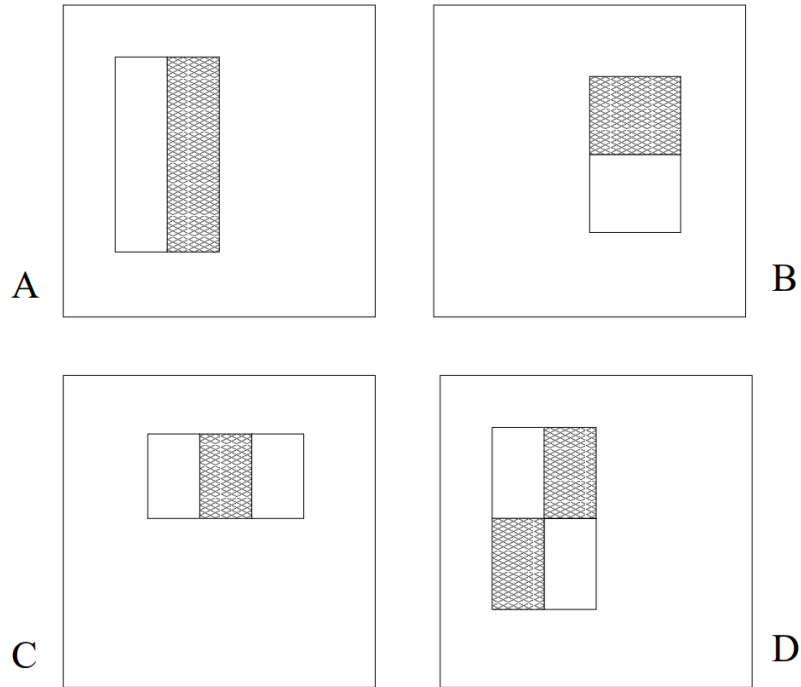
2.1 Haar

Táto metóda bola popísaná autormi Viola a Jones. [5]. Medzi jej hlavné výhody patrí vysoká rýchlosť a spoľahlivá detekcia. Metóda funguje na porovnávaní celých blokov pixelov. Tieto bloky môžu mať rôzne tvary, veľkosť a natočenie. Tieto bloky môžu nadobúdať rôzne tvary ale vo všeobecnosti sa používajú 3 hlavné typy príznakov:

- **Dvoj-obdĺžnikové** (*angl. two-rectangle*) - porovnávajú sumu pixelov v obdĺžnikových oblastiach, ktoré sa nachádzajú vedľa seba, vodorovne, alebo zvislo
- **Troj-obdĺžnikové** (*angl. three-rectangle*) - porovnávajú sumu obdĺžnikových oblastí, ktoré sa nachádzajú po oboch stranách aktuálnej oblasti a sumu aktuálnej oblasti.
- **Štvor-obdĺžnikové** (*angl. four-rectangle*) - počítajú rozdiel medzi dvoma aktuálnymi obdĺžnikovými областami, ktoré sa dotýkajú svojimi rohmi, a obdĺžnikovými областami medzi nimi

Znázornenie jednotlivých typov môžeme vidieť na obrázku 1. Jednotlivé príznaky môže byť použité dostatočne efektívne. Efektivita klesá pri aplikácii príznaku na celý obraz. Vhodným riešením je preto skombinovať menšie množstvo príznakov. Pre výber príznakov existuje mnoho algoritmov. Je použitie na celý obrázok sa efektivita stráca. Na vytvorenie efektívneho klasifikátora je potrebné skombinovať veľmi malé množstvo príznakov. Na vybratie správnych efektívnych

príznakov sa používajú špeciálne algoritmy. Jedným s najpoužívanějších algoritmov pre zvýšenie efektivity je AdaBoost [6], ktorý vytvoril profesor Yoav Freund. Kvôli rýchlemu prístupu k dátam na obraze a rýchlym výpočtom obdĺžnikových príznakov sa zaviedol pojem Integrálny obraz. Integrálny obraz je reprezentáciou pôvodného obrazu. V umiestnení x, y obsahuje sumu pixlov nahor a naľavo od x, y , vrátane (viď. Rovnica 3)



Obr. 1: Haar - dvoj-obdĺžnikové príznaky (A, B), troj-obdĺžnikové príznaky (C) a štvor-obdĺžnikové príznaky (D) [1]

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y'), \quad (1)$$

2.2 HOG

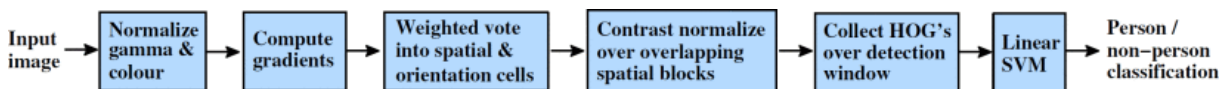
S nápadom vylepšiť detekciu objektov použitím príznakov prišli v roku 2005 Navneed Dalal a Bill Triggs [2], kde postupne vyskúšali niekoľko typov deskriptorov. V práci taktiež podobne rozobrali možnosti a spôsoby ako správne určiť parametre ich detekčnej metódy pre správne fungovanie detekcie jednotlivých tried. Podstatou ich metódy je, že objekt môže byť charakterizovaný viacerými spôsobmi. Táto metóda je rozdelená do niekoľkých samostatných krokov (Obr. 2):

- **Úprava obrazu** - v tomto kroku je potrebné v obraze upraviť kontrast a jas , ktoré by mohli spôsobovať problémy v nasledujúcich krokoch. Okrem tejto úpravy je možné obraz upraviť napríklad gamma filtrom.
- **Výpočet gradientov** - veľkosť gradientov sa počíta na základe vstupného obrazu a masky. Masky, ktoré sa používajú v tomto kroku sú $[-1, 0, 1]$ alebo $[-1, 0, 1]^T$. Gradienty je nutné vypočítať v oboch osiach, čím získame I_x a I_y . Po získaní gradientov je potrebné vypočítať veľkosť gradientov $m(x,y)$ a ich smer $\theta(x,y)$.

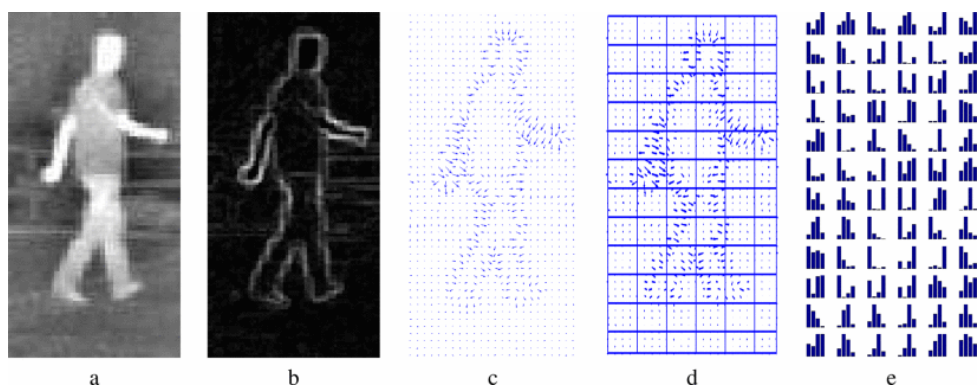
$$m(x,y) = \sqrt{I_x^2 + I_y^2} \quad (2)$$

$$\theta(x,y) = \left(\frac{I_y}{I_x} \right) \quad (3)$$

- **Normalizácia** - Pre správne fungovanie je potrebné obraz normalizovať, aby sa minimalizovali rozdiely medzi jednotlivými bunkami. Tento krok spočíva v skladaní viacerých buniek, čím následne vznikajú bloky.
- **Deskriptor** - Je vytvorený zo vstupného obrazu do jednotlivých blokov. Jednotlivé bloky sa posúvajú a prekrývajú o daný počet pixelov. Výsledok deskriptoru je odovzdaný klasifikátoru, ktorý následne určuje do akej triedy objekt patrí. Jeden z často používaných klasifikátorov je Support vector machine (SVM), ktorý napríklad používali autori práce na efektívnu detekciu chodcov. [7]



Obr. 2: HOG - séria krokov [2]



Obr. 3: HOG - vstupný obraz (a), normalizácia gradientu (b), orientácia gradientu (c), rozdelenie do buniek (d) vypočítaný histogram (e). [3]

3 Detekcia pozície vodiča

Nieco o detekciach pozície

3.1 OpenPOSE

openPose

3.2 TensorFlow

tensorflow

3.3 Ostatné metódy

WrncAI

4 Využitie sférických kamier v automobiloch

Počas vypracovania diplomovej práce som mal zapožičané 2 sférické kamery.

4.1 Technické parametre

Go PRO:

THETA:

Senzor FishEye CMOS 2x12MPix Maximálne rozlíšenie (Video) 4K 30fps Maximálne rozlíšenie (Fotografia) 14.5MP (5376x2688px) Svetelnosť f2.0 Vnútorná pamäť 19GB

$$\left(\sum_{n=1}^{\infty} a_n b_n \right)^2 \leq \sum_{n=1}^{\infty} a_n^2 \cdot \sum_{n=1}^{\infty} b_n^2 \quad (4)$$

4.2 Použitie v analýze videa

Problem s formatom videa, rozlisenim , deformaciou hroe a dole

5 Program

Zhrnutie vysledkov

5.1 Požiadavky a návrh programu

Architektura , schemy

5.2 Pozícia vodiča

5.2.1 Detekcia

NN klasifikator

5.2.2 Neurónová sieť

NN klasifikátor

5.3 Orientácia hlavy

5.4 Výstup programu

obrazky

5.5 Porovnanie výsledkov

porovnanie

5.6 Využitie zozbieraných dát

porovnanie

5.7 Používatelská příručka

`python program.py --use-openPose=true`

6 Možnosti vylepšenia detekcie

Zhrnutie výsledkov

7 Záver

Zhrnutie výsledkov

Literatúra

- [1] Paul Viola, Michael Jones, et al. Robust real-time object detection. *International journal of computer vision*, 4(34-47):4, 2001.
- [2] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, volume 1, pages 886–893. IEEE, 2005.
- [3] Massimo Bertozzi, Alberto Broggi, Mike Del Rose, Mirko Felisa, Alain Rakotomamonjy, and Frédéric Suard. A pedestrian detector using histograms of oriented gradients and a support vector machine classifier. In *2007 IEEE Intelligent Transportation Systems Conference*, pages 143–148. IEEE, 2007.
- [4] Azriel Rosenfeld. Picture processing by computer. *ACM Computing Surveys*, 1(3):147–176, Jan 1969.
- [5] Paul Viola and Michael Jones. Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001*, volume 1, pages I–I. IEEE, 2001.
- [6] Yoav Freund and Robert E Schapire. A desicion-theoretic generalization of on-line learning and an application to boosting. In *European conference on computational learning theory*, pages 23–37. Springer, 1995.
- [7] Yanwei Pang, Yuan Yuan, Xuelong Li, and Jing Pan. Efficient hog human detection. *Signal Processing*, 91(4):773–781, 2011.