

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

Računalni vid
Projekt

Histogram orijentiranih gradijenata - detekcija i praćenje ljudi

Petra Bevandić
Dragan Drandić
Melita Kokot
Igor Smolković
Dino Šantl

Zagreb, prosinac 2013.

SADRŽAJ

1. Prikupljena literatura	1
1.1. HOG članci	1
1.1.1. Histogram of Oriented Gradients for Human Detection	1
1.1.2. Detection Using a Cascade of Histograms of Oriented Gradients	1
1.1.3. Pedestrian Detection Using Infrared Images and Histograms of Oriented Gradients	2
1.1.4. Visual Classification of Coarse Vehicle Orientation Using Histogram of Oriented Gradients Features	2
1.1.5. Enhancing Real-time Human Detection based on Histograms of Oriented Gradients	3
1.1.6. Integral Histogram: A fast way to Extract Histograms in Cartesian Spaces	3
1.2. Praćenje objekata	3
2. Dijagram planiranog sustava	5
3. Baza slika/videoa	7
3.1. Uvod	7
3.2. Dostupne baze podataka	7
3.2.1. INRIA	7
3.2.2. CAVIAR Test Case Scenarios	7
3.2.3. PETS 2009	7
3.2.4. PETS 2006	8
3.2.5. BEHAVE	8
3.2.6. Pedestrian	8
3.3. Inicijalni odabir baze	8

4. Programski alati	9
4.1. Programski jezik	9
4.2. Biblioteke	9
4.2.1. OpenCV	9
4.2.2. Scikit-image	9
4.3. RapidMiner	10
4.4. Repozitorij koda	10
5. Napredak projekta	11

1. Prikupljena literatura

1.1. HOG članci

1.1.1. Histogram of Oriented Gradients for Human Detection

- **Histogram of Oriented Gradients for Human Detection**
- Navneet Dalal, i Bill Triggs
- U International conference on Computer Vision & Pattern Recognition, Vol. 1, stranice 886-893, Lipanj 2005.
 - URL: <http://lear.inrialpes.fr/people/triggs/pubs/Dalal-cvpr05.pdf>
 - opisan histogram orijentiranih gradijenata
 - analiziran utjecaj raznih parametara histograma na performanse
 - HOG primjenjene na prepoznavanje pješaka (uz korištenje SVM-a)
 - MIT i Inria baza pješaka, cca 2500 slika, originalne reflektirane
 - izdvajanje "teških" primjera
 - postojanje okoline oko osobe

1.1.2. Detection Using a Cascade of Histograms of Oriented Gradients

- **Detection Using a Cascade of Histograms of Oriented Gradients**
- Qiang Zhu, Shai Avidan, Mei-Chen Yeh, i Kwang-Ting Cheng. Fast Human
- U International conference on Computer Vision & Pattern Recognition, Vol. 2, stranice 1491-1498, Lipanj 2006.
 - URL: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.68.6232&rep=rep1&type=pdf>

- ubrzanje traženja ljudi u slikama
- ne koriste blokove fiksne veličine - globalne karakteristike
- Ada-Boost
- korištenje "integralne slike" - koriste se HOG-ovi izračunati za manje blokove da bi se ubrzalo računanje histograma za veće blokove
- slučajan odabir blokova koji idu u klasifikator

1.1.3. Pedestrian Detection Using Infrared Images and Histograms of Oriented Gradients

- **Pedestrian Detection Using Infrared Images and Histograms of Oriented Gradients**
- Frédéric Suard, Alain Rakotomamonjy, Abdelaziz Bensrhair, i Alberto Broggi
- U Intelligent Vehicles Symposium, stranice 206-212, 2006.
 - URL: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.80.1379&rep=rep1&type=pdf>
 - detekcija ljudi pomoću infracrvenih slika
 - HOG značajke, SVM klasifikator
 - primjena u detekciji pješaka noću
 - odabir optimalnih parametara svih stavaka sustava za detekciju osoba
 - osrednji rezultati

1.1.4. Visual Classification of Coarse Vehicle Orientation Using Histogram of Oriented Gradients Features

- **Visual Classification of Coarse Vehicle Orientation Using Histogram of Oriented Gradients Features**
- Paul E. Rybski, Daniel Huber, Daniel D. Morris, i Regis Hoffman
 - URL: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.183.4072&rep=rep1&type=pdf>
 - detekcija orijentacije vozila iz slike (bez informacije o smjeru kretanja)
 - jednostavna primjena prethodno opisanog algoritma na prepoznavanje vozila
 - korištena javno dostupna implementacija HOG-a

- vrlo dobri rezultati

1.1.5. Enhancing Real-time Human Detection based on Histograms of Oriented Gradients

- **Enhancing Real-time Human Detection based on Histograms of Oriented Gradients**
- Marco Pedersoli, Jordi González, Bhaskar Chakraborty, i Juan J. Villanueva
 - URL: http://iselab.cvc.uab.es/files/Publications/2007/PDF/CORES07_MP.pdf
 - ubrzanje računanje HOG-a, izrada frameworka za detekciju osoba
 - (pristupi detekciji osoba - detekcija cijele osobe vs. detekcija dijelova koji su povezani)
 - AdaBoost
 - računanje značajki za veće blokove preko značajki za manje blokove
 - MIT baza osoba

1.1.6. Integral Histogram: A fast way to Extract Histograms in Cartesian Spaces

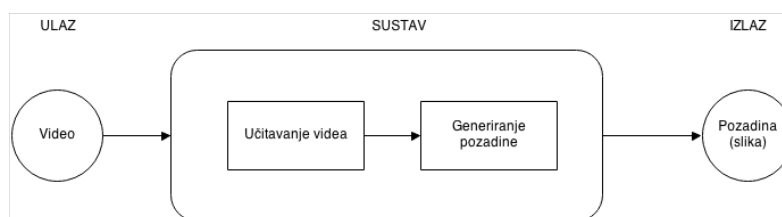
- **Integral Histogram: A fast way to Extract Histograms in Cartesian Spaces**
- Fatih Porikli
- U International conference on Computer Vision & Pattern Recognition, Vol. 1, stranice 829-836, Lipanj 2005.
 - URL: <http://www.merl.com/reports/docs/TR2005-057.pdf>
 - primjenjiv u sustavima koji rade u realnom vremenu
 - koristi se za bilo koju metodu koja traži maksimalno preklapanje histograma

1.2. Praćenje objekata

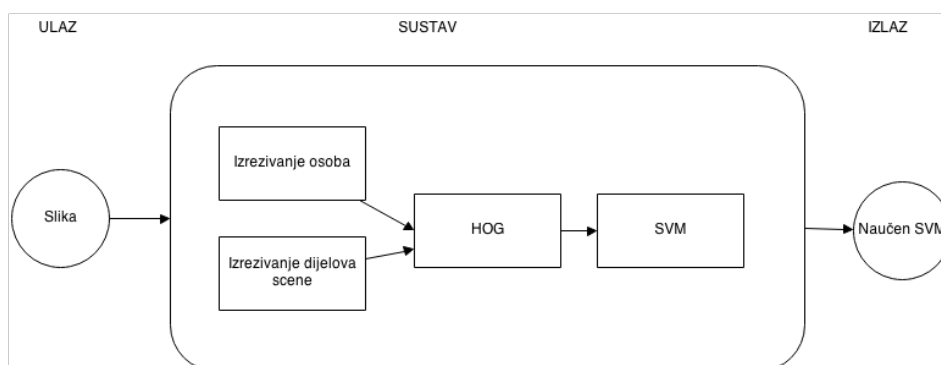
- **Multitarget Tracking of Pedestrians in Video Sequences Based on Particle Filters**

- Hui Li, Shengwu Xiong, Pengfei Duan, Xiangzhen Kong
- <http://www.hindawi.com/journals/am/2012/343724/>
- **Object Tracking in Crowded Video Scenes Based on the Undecimated Wavelet Features and Texture Analysis**
 - M. Khansari, H. R. Rabiee, M. Asadi, M. Ghanbari
 - <http://asp.eurasipjournals.com/content/pdf/1687-6180-2008-2.pdf>
- **Real-time object detection and tracking for industrial applications**
 - Selim Benhimane, Hesam Najafi, Matthias Grundmann, Ezio Malis, Yakup Genc, Nassir Navab
 - <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.119.2257>
- **Efficient Tracking of Many Objects in Structured Environments**
 - Nathan Jacobs, Michael Dixon, Scott Satkin, Robert Pless
 - http://cse.wustl.edu/Research/Lists/Technical%20Reports/Attachments/859/sheet_tracking_tr.pdf
- **Fast and accurate moving object extraction technique for MPEG-4 object-based video coding**
 - Ju Guo, Jongwon Kim, C.C: Jay Kuo
 - <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.6.6887>

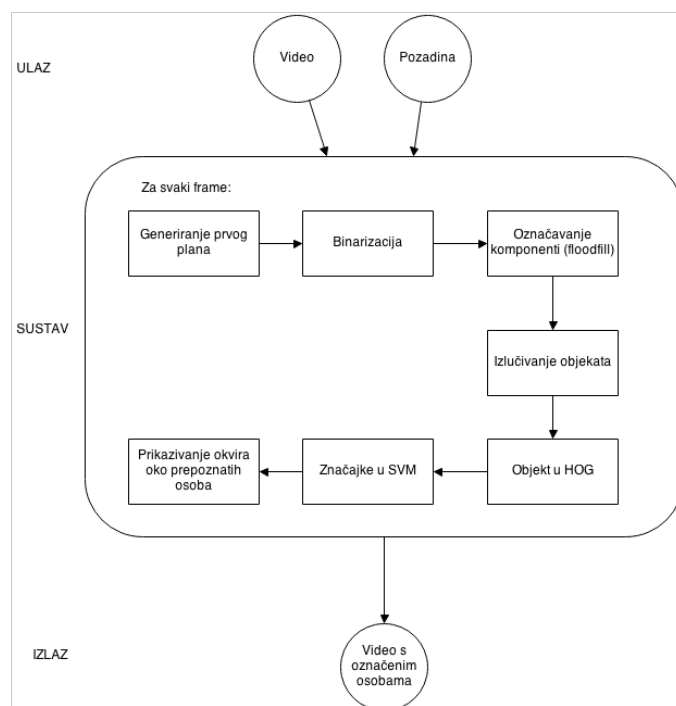
2. Dijagram planiranog sustava



Slika 2.1: Izlučivanje pozadine



Slika 2.2: Učenje modela strojnog učenja (npr. SVM)



Slika 2.3: Prikaz rada sustava u realnom vremenu u modu traženja osobe na slici

3. Baza slika/videoa

3.1. Uvod

Za promatrani problem koristi se dvije vrste baza podataka. Prva vrsta sastoji se od slika, a druga vrsta od video snimaka.

3.2. Dostupne baze podataka

3.2.1. INRIA

Baza se sastoji od slika. Navedena baza korištena je upravo za testiranje HOG algoritma. Podijeljena je na pozitivne (na slici se nalazi jedna ili više osoba) i negativne (osoba na slici nema). Uz to, baza sadrži i opisnike slike koji u sebi sadrže informaciji o pozicijama osoba na slikama. Slike su dostupne u više različitih dimenzija (pixel x pixel): 70x134, 96x160 i 64x128. Baza se sastoji od ukupno 1218 negativnih primjera slika i 614 pozitivnih slika. Za generiranje negativnih primjera negativne slike mogu se izrezivati na bilo koji način. Nad pozitivnim slikama definirani su opisnici tako da se ukupno može generirati 2500 pozitivnih primjera.

3.2.2. CAVIAR Test Case Scenarios

Baza se sastoji od video snimaka. Baza je bazirana na javnim nadzornim kamerama. Baza se sastoji od skupa ciljanih akcija osoba. Detekcija akcija nije u središtu ovog projekta pa se mogu koristiti sve snimke ravnopravno. Dimenzija slika u video snimcima je 384px x 288px. Format je MPEG2 i sadrži 25 slika u sekundi.

3.2.3. PETS 2009

Baza se sastoji od slika. Baza je korištena za praćenje skupa ljudi pomoću nadzornih kamera. Slike su u JPEG formatu. Dvije su različite dimenzije slika. Jedna vrsta slika

dimenzija je 768px x 576px, a druga vrsta 720px x 576px. Svojstvo baze je da su dimenzije osoba na slikama relativno malne s obzirom na cijelu sliku. Ovakva baza može ispitati koliko je sustav osjetljiv na dimenzije osoba.

3.2.4. PETS 2006

Baza se sastoji slika u JPEG formatu. Dimenzija slika je 768px x 576px i posljedica su uzrokovana 25 slika u sekundi. Baza je korištena za specifični problem praćenja ostavljene prtljage. Kako je u središtu proučavanja osoba, ova baza pogodna je i za naš promatrani problem.

3.2.5. BEHAVE

Baza se sastoji od slika u JPEG formatu. Dimenzije slike su 640px x 480px i slike predstavljaju slike koje su uzrokovane s frekvencijom 25 slika u sekundi. Baza je korištena za prepoznavanje određenih ponašanja kod osoba promatranih nadzornim kamerama.

3.2.6. Pedestrian

Baza se sastoji od slika u PPM formatu. Baza sadrži 924 slika. Slike su dimenzija 64px x 128px. Baza se sastoji od slika osoba. Procijenili smo da je dobra kod testiranja različitih modela strojnog učenja.

3.3. Inicijalni odabir baze

Odlučili smo odabrati jednu bazu podataka nad kojom ćemo promatrati problem i izraditi prototip. Odlučili smo se za INRIA bazu zbog toga što ima opisnike osoba na slici. Zbog toga možemo brzo izraditi primjere i testirati naš sustav.

4. Programski alati

4.1. Programski jezik

Za implementaciju sustava odabran je programski jezik python. Glavni razlog odabira jezika je lakoća pisanja programskog koda te širok izbor biblioteka u kojima su implementirani algoritmi računalnog vida. Python također sadrži veliku paletu drugih biblioteka koje nisu dostupne u nekim drugim programskim jezicima (primjerice C/C++). U konačnoj fazi projekta biti će potrebno izraditi grafičko korisničko sučelje što ne predstavlja problem jer za python postoji puno različitih frameworka koji olakšavaju tu fazu projekta. Python također omogućava pokretanje implementiranog sustava na različitim operacijskim sustavima (Windows, Linux).

4.2. Biblioteke

Za implementaciju sustava koriste se sljedeće programske biblioteke:

- openCV
- scikit-image

4.2.1. OpenCV

OpenCV (Open Source Computer Vision) je biblioteka koja sadrži implementacije algoritama računalnog vida koji su namijenjeni za korištenjem u realnom vremenu. Pruža podršku i za algoritme stajnog učenja koji će biti korišteni u izradi projekta (primjerice k-NN, SVM, neuronske mreže).

4.2.2. Scikit-image

Scikit-image je također biblioteka otvorenog koda. OpenCV više orijentiran na algoritme računalnog vida, dok scikit-image podržava više metoda za obradu slike. OpenCV

podržava više programskih jezika(C++, python, Java), dok je biblioteka scikit-image namijenjena za python. Scikit-image za razliku od openCV-a za interno spremanje slika koristi numpy.ndarray. Novije verzije openCV-a također mogu koristiti numpy pa se time postiže međusobna kompatibilnost. Prednost scikit-image je lakše pisanje programskog koda i bolja razumljivost dokumentacije biblioteke. Pojedini algoritmi su implementirani unutar oba dvije biblioteke te se može uočiti bolja efikasnost openCV implementacije. Razlog tome je namjena openCV za primjenu u industriji.

4.3. RapidMiner

RapidMiner je okruženje za strojno učenje, rudarenje podataka, itd. Alat će biti korišten za učenje, testiranje i validaciju modela koji će služiti za detekciju ljudi na video sekvencama. Alat pruža mogućnost istovremene primjene različitih postupaka učenja modela. Te na taj način možemo odabrati postupak koji daje najbolje rezultate. Dodatno podržava automatsku optimizaciju hiperparametara čime se model može dodatno poboljšati kako bi se postigli najbolji mogući rezultati. Cilj korištenja RapidMinera je pronaći najbolji model koji će kasnije biti implementiran korištenjem openCV/scikit-image biblioteka.

4.4. Repozitorij koda

Odabrali smo GitHub za repozitorij koda. Repozitorij koda je nužan budući da na razvoju sustava radi veći broj ljudi te postoji puno nadopuna programskog koda. GitHub pruža podršku za održavanje konzistencije i lakoću dijeljenja programskog koda.

5. Napredak projekta

Instalirane su biblioteke i alati potrebni za izradu projekta. Napravljena su prva testiranja alata i *HOG* implementacija. Ispitane različite mogućnosti obrade video snimaka. Isprobana su različita odstranjivanja pozadina i detekcija objekata na binarnoj slici. Iz baze podataka izvučene su značajke pomoću *HOG* algoritma i isprobani različiti modeli strojnog učenja.