

UNIVERZA V LJUBLJANI  
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Tomaž Kunst

**Aplikacija za analizo promocij na  
osnovi računalniškega vida**

DIPLOMSKO DELO

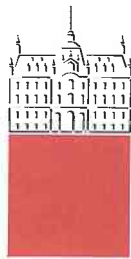
VISOKOŠOLSKI STROKOVNI ŠTUDIJSKI PROGRAM PRVE  
STOPNJE RAČUNALNIŠTVO IN INFORMATIKA

MENTOR: doc. dr. Peter Peer

Ljubljana 2012

Rezultati diplomskega dela so intelektualna lastnina avtorja in Fakultete za računalništvo in informatiko Univerze v Ljubljani. Za objavlanje ali izkoriščanje rezultatov diplomskega dela je potrebno pisno soglasje avtorja, Fakultete za računalništvo in informatiko ter mentorja.

*Besedilo je oblikovano z urejevalnikom besedil  $\text{\LaTeX}$ .*



Št. naloge: 00339/2012

Datum: 04.09.2012

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko izdaja naslednjo nalogo:

Kandidat: **TOMAŽ KUNST**

Naslov: **APLIKACIJA ZA ANALIZO-PROMOCIJ NA OSNOVI  
RAČUNALNIŠKEGA VIDA  
APPLICATION FOR ANALYSIS OF PROMOTIONS BASED ON  
COMPUTER VISION**

Vrsta naloge: Diplomsko delo visokošolskega strokovnega študija prve stopnje

Tematika naloge:

Trgovci za uspešno poslovanje izvajajo analize prodaje. Implementirajte aplikacijo za samodejno analizo oseb v okolici promocijske točke, ki deluje na osnovi računalniškega vida. Rešitev naj vsebuje detekcijo obrazov, sledenje osebam ter klasifikacijo oseb po spolu. Na izhodu iz aplikacije nas zanimajo naslednje aktivnosti: ali se je oseba pred promocijo ustavila, jo je le opazovala ob mimohodu, ali pa je promocija sploh ni zanimala. Seveda nas to zanima za oba spola. Aplikacijo testirajte in poročajte o rezultatih.

Mentor:

  
doc. dr. Peter Peer



Dekan:

  
prof. dr. Nikolaj Zimic

## IZJAVA O AVTORSTVU DIPLOMSKEGA DELA

Spodaj podpisani Tomaž Kunst, z vpisno številko **63060173**, sem avtor diplomskega dela z naslovom:

*Aplikacija za analizo promocij na osnovi računalniškega vida*

S svojim podpisom zagotavljam, da:

- sem diplomsko delo izdelal samostojno pod mentorstvom doc. dr. Petra Peera,
- so elektronska oblika diplomskega dela, naslov (slov., angl.), povzetek (slov., angl.) ter ključne besede (slov., angl.) identični s tiskano obliko diplomskega dela
- soglašam z javno objavo elektronske oblike diplomskega dela v zbirki "Dela FRI".

V Ljubljani, dne 17. septembra 2012

Podpis avtorja:

*Najprej bi se rad zahvalil mojemu mentorju doc. dr. Petru Peeru, za vso pomoč, razumevanje in za dobro sodelovanje pri izdelavi mojega diplomskega dela.*

*Iskrena hvala mojemu šefu dr. Petru Dolinarju za spodbudo in nasvete.*

*Hvala tudi mojim prijateljem in dekletu, ki so mi nudili moralno podporo in so mi priskočili na pomoč tudi pri izdelavi diplomske naloge.*

*Največja zahvala za to, gre seveda moji družini, ki me je podpirala med tekom študija. Hvala vam za vse!*

# Kazalo

## Seznam uporabljenih kratic

<b>Povzetek</b>	<b>1</b>
<b>Abstract</b>	<b>3</b>
<b>1 Uvod</b>	<b>5</b>
1.1 Specifikacija problema . . . . .	5
1.2 Cilj: analiza promocij . . . . .	6
1.3 Struktura naloge . . . . .	7
<b>2 Sorodna dela</b>	<b>9</b>
2.1 Strokovno znanstvena dela . . . . .	9
2.2 Produkti . . . . .	10
<b>3 Razvoj produkta</b>	<b>13</b>
3.1 Uporabljene tehnologije in orodja . . . . .	13
3.2 Načrt sistema . . . . .	14
3.3 Zaznava osebe na video posnetku . . . . .	14
3.4 Sledenje oseb . . . . .	20
3.5 Klasifikacija oseb po spolu . . . . .	24
3.5.1 Redukcija dimenzij . . . . .	25
3.5.2 Klasifikator . . . . .	26
3.6 Priprava izhodnih podatkov . . . . .	27
3.6.1 Združevanje podatkov . . . . .	28

## KAZALO

3.6.2 Določanje iskanih aktivnosti . . . . .	29
<b>4 Uporaba produkta</b>	<b>33</b>
<b>5 Rezultati</b>	<b>37</b>
5.1 Klasifikacija spola . . . . .	37
5.2 Pridobitev z uvedbo ROI . . . . .	38
5.3 Končne zmogljivosti obdelanih slik . . . . .	38
<b>6 Zaključek</b>	<b>41</b>
<b>Slike</b>	<b>43</b>
<b>Tabele</b>	<b>45</b>
<b>Literatura</b>	<b>47</b>

# Seznam uporabljenih kratic

**XML** – Označevalni jezik za lokalno shranjevanje podatkov in za spletno komunikacijo (angl. Extensible Markup Language)

**SVM** – Metoda za strojno učenje na osnovi podpornih vektorjev (angl. Support Vector Machines)

**PCA** – Metoda za redukcijo dimenzij na temelju glavnih komponent (angl. Principal Component Analysis)

**ROI** – Zmanjšanje opazovanega območja (angl. Region of Interest)

**GPS** – Sistem globalnega določanja geografske lege (angl. Global Positioning System)

# Povzetek

V diplomskem delu je predstavljen produkt za avtomatično analizo oseb v okolici promocijske točke, ki deluje na osnovi računalniškega vida. Za izvajanje potrebujemo računalnik, na katerem teče program, in kamero za zajem slike v realnem času. Uporablja se lahko kamera, ki je fizično priklopljena na računalnik preko USB priključka, ali pa oddaljena kamera z omrežno povezavo. Z metodo Viola-Jones se išče obraze in se jih nato klasificira po spolu. Za klasifikacijo se uporablja metoda podpornih vektorjev (SVM) v kombinaciji z metodo glavnih komponent (PCA) za določanje značilnic, ki najboljše predstavljajo obraze. Pri učenju klasifikatorja se uporablja baza slik študentov iz univerze Essex (VB). Dosežena klasifikacijska točnost za spol je 84,75 %. Kadar obrazov ni mogoče zaznati, se osebam sledi z metodo Lucas-Kanade. Tako so podatki o številu oseb natančnejši. Spremlja se odziv potrošnikov na promocijo, končno poročilo vsebuje podatke o vnaprej določenih odzivih potrošnikov in klasificirane podatke o njih.

## **Ključne besede:**

iskanje obrazov, klasifikacija spola, produkt, analiza potrošnikov

# Abstract

In our work we present product based on computer vision for automatic analysis of persons in the vicinity of a promotion point. For execution we need a computer with camera and program for capturing real time video. We can use camera that is plugged in via USB port or remote network enabled camera. Firstly, we use Viola-Jones method for face detection. Later, we classify them based on gender. Support vector machine (SVM) in combination with principle component analysis (PCA) is used for classification and face features recognition. The classifier is built using image database of students pictures from University of Essex (GB). Achieved gender classification score is 84.75 %. When faces cannot be detected, we use Lucas-Kanade method for following people. This approach helps us gain accurate data for number of persons who walked by. We monitor customer feedback on promotion. Final report includes data of predefined customer responses and classified data about them.

**Key words:**

face detection, gender classification, product, customers analysis

# Poglavje 1

## Uvod

### 1.1 Specifikacija problema

Pri razvoju produkta se ukvarjamo z iskanjem, klasifikacijo in sledenjem ljudem v trgovinah, ki temelji na iskanju obrazov. V trgovskih centrih se ob določenih urah nahaja ogromno ljudi, zaradi česar je velika verjetnost medsebojnega prekrivanja obrazov. Zaradi zelo raznolikega ozadja pa lahko pride tudi do napačnih zaznav obrazov. Zato je bil naš namen narediti robusten produkt, ki bo čim manj odvisen od okolja in nam bo dajal realne podatke.

V diplomskem delu je predstavljen razvoj in končni produkt za analizo promocijskih mest. Produkt s pomočjo zaznave obrazov sledi osebam in jim določi spol in njihovo zainteresiranost do promocije. Pri razvoju smo uporabljali že preizkušene rešitve. V manjši meri je bila potrebna razširitev posameznih modulov in nadgradnja obstoječih rešitev. V poglavju 3 so predstavljene implementacije funkcionalnosti in njihove predelave.

Z izdelavo produkta smo imeli namen avtomatizirati analizo odziva potrošnikov na promocije in s tem tudi zainteresiranost do promoviranega izdelka. S tem bi lahko preprečili prodajne neuspehe. Z analizo odziva in klasifikacijo oseb bi lahko prodajalci bolje načrtali prodajno strategijo.

## 1.2 Cilj: analiza promocij

Trgovci za uspešno poslovanje izvajajo analize prodaje. V prodajalnah razporedijo police z različnimi izdelki tako, da je skupaj čim več izdelkov, ki bi zanimali določenega potrošnika. To lahko nadzirajo na tak način, da preverijo vsak posamezen nakup in vidijo, kateri izdelki so bili kupljeni skupaj. Pri tej analizi lahko preverimo samo povezave med izdelki, nimamo pa nobenih podatkov o potrošniku. Podatki o potrošniku nam lahko povedo veliko, saj lahko na tak način bolje sklepamo, kaj si želi posamezna ciljna skupina ljudi. Če bi spremljali dogajanje direktno pri polici, bi lahko videli, kako dolgo se je določen potrošnik odločeval za nakup izdelka, oz. če se je sploh odločil za nakup. Lahko tudi nadziramo, kateri potrošniki so se odločili za katerega proizvajalca in kako se spreminjajo potrošniški trendi. Ta način je vsekakor najbolj primeren za dobavitelje, ki lahko z minimalnimi stroški preverijo odziv potrošnikov na določeno blagovno znamko ali izdelek. Dobavitelj tako lahko sestavi svojo strategijo, ki jo poižkuša uskladiti s strategijo trgovca, kar je ključno za uspešen posel.

S pomočjo analize promocij dobimo okvirno sliko kupnih navad populacije. S podrobnejšo analizo lahko spremljamo tudi druge dejavnike, ki vplivajo na odločitev nakupa kot npr. spol, starost itd. Kupne navade se spreminjajo iz leta v leto, zato morajo ponudniki vseskozi analizirati trg, da lahko ostanejo konkurenčni.

V trgovini se pri promocijski mizi giblje veliko ljudi, tudi več naenkrat. Zato je skoraj nemogoče, da bi en človek lahko klasificiral vse osebe in poleg tega še preverjal, kaj katera oseba počne. Za rešitev tega problema je potrebno narediti produkt, ki bo nadziral dogajanje okoli promocijske točke in bo sposoben slediti osebam, jim določati spol in prepoznati, kaj te osebe počnejo na posnetku. Ker nas zanima predvsem odziv potrošnikov na promocijo, štejemo, koliko potrošnikov je zanimala promocija. Zato opazujemo dolžine pogledov proti promocijski točki, in če se je katera oseba približala promocijski točki.

## 1.3 Struktura naloge

Produkt je sestavljena iz petih sklopov. Določili smo jih že pred razvojem in nato razvijali vsakega posamezno, dokler ni dajal zadovoljivih rezultatov. Na začetku smo na posnetku s pomočjo metode Viola-Jones [22] iskali obraze. Pri uporabi metode lahko izbiraš, kako učinkovita naj bo pri zaznavi, zato smo potrebovali nekaj primerov posnetkov končne uporabe produkta, da smo lahko parametre nastavili optimalno. V drugem sklopu je bilo potrebno slediti obrazom [11], ki jih ni bilo mogoče zaznati, da smo lahko na koncu imeli realne podatke o številu prisotnih oseb na posnetku. Ker vemo, kje na sliki se nahaja določen obraz, lahko zanjo določimo še spol. Za klasifikacijo spola smo uporabili PCA redukcijo dimenzij [21] in nato SVM klasifikator [4]. V četrtem sklopu smo zbrane podatke združili za vsako osebo posebej, tako da se shrani vsa zgodovina premikanja in načina sledenja. Na koncu je bilo potrebno te podatke razčleniti na iskane odzive potrošnikov in jih vizualizirati.

V poglavju 2 si bomo pogledali sorodna dela, v poglavju 3 razvoj našega produkta, uporabljene metode in njihove optimizacije, v poglavju 4 uporabo produkta, v poglavju 5 rezultate testiranja, v poglavju 6 pa bomo zapisali bistvene sklepe naloge.



# Poglavje 2

## Sorodna dela

### 2.1 Strokovno znanstvena dela

Robert Ravnik je v diplomskem delu [19] predstavljal interaktivno oglasno mesto, ki se dinamično odziva na dogajanje pred oglasnim mestom. Na posnetku je zaznaval osebe in jim določal spol. Za klasifikacijo spola je testiral štiri različne klasifikatorje. Za učenje klasifikatorjev je uporabljal orodje Orange [18]. Najboljšo klasifikacijsko točnost 85 % je dosegel z uporabo K-najbližjih sosedov in z učno množico velikosti 25 slik. Za učno množico je uporabljal FERET bazo slik [10], ki vsebuje 2413 slik. Za učenje je uporabljal različno velike množice slik do največ 400 slik za spol.

Tadej Jagodnik je v diplomskem delu [12] raziskoval primernost uporabe različnih algoritmov za zaznavo obrazov v avtomobilu. Predstavil je rešitve pri zaznavi v prostoru, kjer so lahko obrazi delno prekriti in nastavitve algoritmov, da so delovali optimalno pri različnih osvetlitvah. Pri zaznavi se je najbolje odrezal tudi najbolj uveljavljen algoritem Viola-Jones [22]. Preizkusil je tudi algoritma Nilsson et. al. in Kienzle et. al. [16, 14].

Matija Mlinar je v diplomskem delu [15] predstavil metode za avtomatično klasifikacijo oseb v starostne razrede. Za redukcijo dimenzij slik obrazov je uporabljal metode PCA, LDA, 2DLDA in SAM. Za klasifikacijo oseb v različne starostne razrede je uporabil Gaussove modele, polinomske modele,

najbližje sosede, SVM, odločitveno drevo. Najboljše rezultate je dobil z uporabo SAM in SVM. Klasifikacijska točnost je bila 72 %. Preizkusil je tudi metodo AGES. Z njo je dosegel klasifikacijsko točnost 73 %.

## 2.2 Produkti

- Medijske analize (Comland – Comsight [6]): produkt za analizo gledanja video programov. Šteje gledalce in analizira njihove emocije pri gledanju programa. S pomočjo računalniškega vida poda celostno informacijo o spremljanju in odzivu na predvajano vsebino.
- Adaptivno oglaševanje (Comland – Comsight [6]): produkt izbira oglase primerne za osebo, ki je pred oglasnim ekranom. Preverja spol in starost.
- Analize potrošnikov (Comland – Comsight [6]): produkt s pomočjo računalniškega vida zazna odziv osebe na določeno vsebino, ki ga podzavestno izrazi preko obrazne mimike. Zaznava pozornost in emocionalen odziv (nevtralno, nasmeh, presenečenje in zaskrbljenost).
- RetailNext [20]: ameriško podjetje, ki razvija istoimenski produkt za optimizacijo trgovanja. Produkt preko sistema kamer opazuje potrošnike in njihove reakcije. Spremlja njihovo gibanje in označi gostoto gibanja v trgovini. Zazna, kje se je kdo ustavil, pogledal in kaj je kdo kupil. Omogoča tudi optimizacijo delovne sile. Vse podatke prikazuje v nadzorni plošči.
- ACIC [1]: belgijsko podjetje, ki razvija celovite rešitve na področju računalniškega vida. Tržijo specializirane računalnike in strežnike s programsko opremo za računalniški vid. Na videoposnetkih predvsem štejejo objekte v gibanju. Produkt McTraffic meri hitrost zaznanim vozilom na posnetku. MvActivityDetection zaznava gibanja na posnetku in sproži alarm, kadar se objekt nahaja v določenem območju. MvPeopleCounting šteje ljudi, ki so prečkali določeno mejo. MvStationary

---

ugotavlja trajne spremembe na sceni, preverja, če je kakšen predmet postal statičen, ali pa je bil odstranjen iz scene. MvFlowMonitor spremlja premikanje oseb in sproži alarm, če se oseba premika v napačno smer.

- VideoMining [23]: produkt za analiziranje potrošnikov. VideoMining z lastnimi tehnologijami avtomatično pridobiva podatke o obnašanju potrošnikov iz video kamer v trgovinah. Pridobljeni podatki se analizirajo in interpretirajo za vpogled v dejansko dogajanje v trgovinah. Podatki se lahko kombinirajo z bazami za zvestobo potrošnikov in bazo nakupov.



# Poglavje 3

## Razvoj produkta

### 3.1 Uporabljene tehnologije in orodja

Pri razvoju produkta smo uporabljali programski jezik C++, razvojno okolje Eclipse CDT (C/C++ Development Tool) [8], knjižnico OpenCV [17] in knjižnico Boost [5].

Za programski jezik C++ smo se odločili predvsem zaradi možnosti objektnega programiranja in hitrejšega izvajanja v primerjavi z ostalimi objektno usmerjenimi programskimi jeziki.

OpenCV (Open source Computer Vision library) [2, 17] je odprtokodna programska knjižnica namenjena za realno časovno računalniško zaznavanje iz slikovnega materiala. Licencirana je z BSD licenco, kar pomeni, da jo je dovoljeno prosto uporabljati in spreminjati. Knjižnica je napisana v jeziku C++, a jo je s pomočjo ovojnice mogoče uporabljati z več različnimi programskimi jeziki kot so: C#, Ch, Python, Ruby, Java.

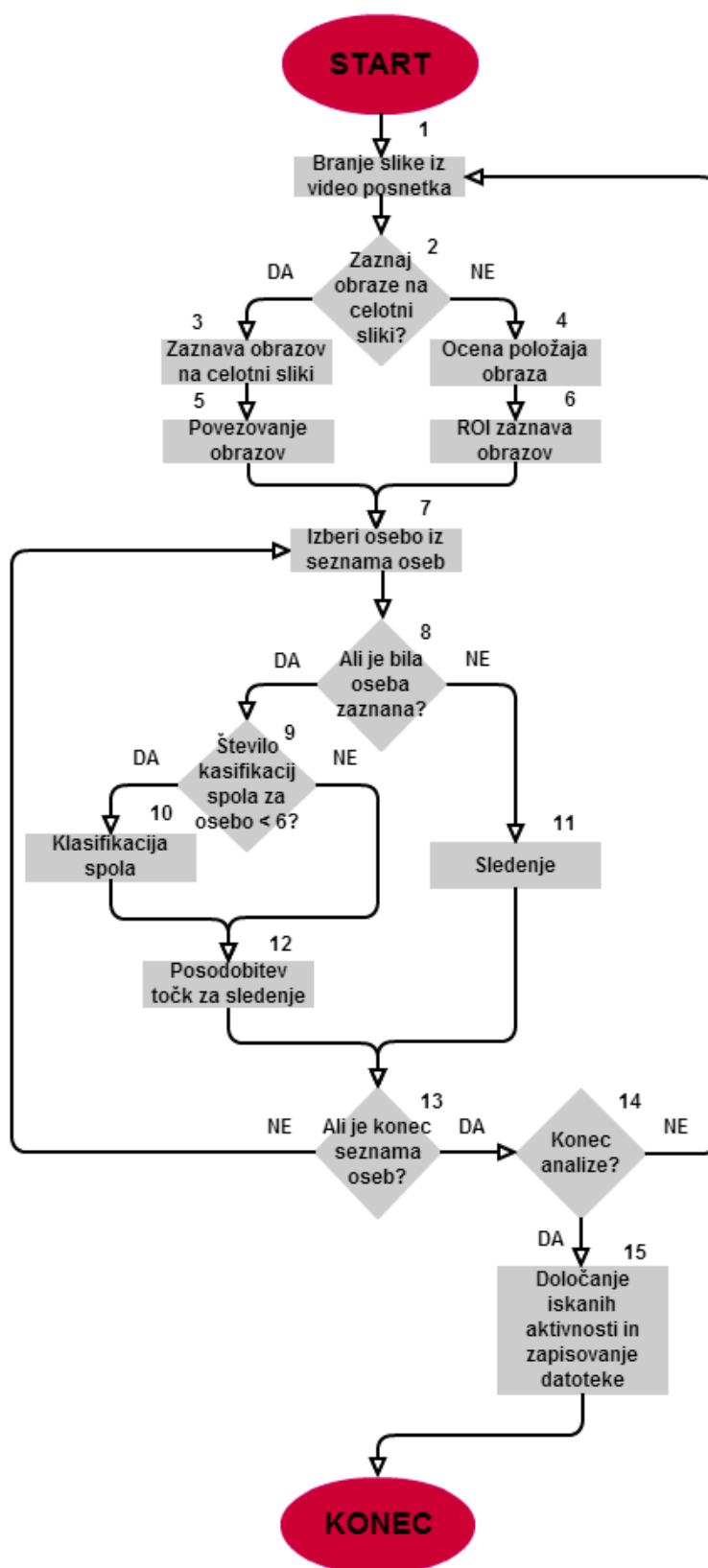
Boost [5] je odprtokodna programska knjižnica, licencirana z lastno istoimensko licenco, ki se zgleduje po BSD in MIT licenci. Knjižnica vsebuje podatkovne strukture, metode za nitenje, matematične operacije itd. Vse metode so optimizirane, da porabijo čim manj procesorskega časa. Podatkovne strukture iz knjižnice Boost je smotrno uporabljati tudi zaradi optimalne alokacije pomnilnika.

## 3.2 Načrt sistema

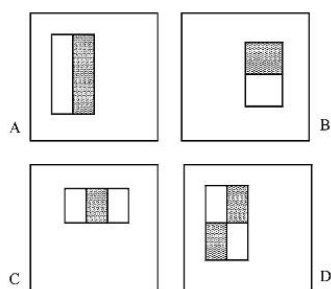
Na sliki 3.1 je predstavljen diagram poteka programa, ki je ključni del našega produkta. Koraki 1, 2, 3, 4, 5 in 6 predstavljajo zaznavo obrazov na posnetku, ki je podrobneje predstavljena v poglavju 3.3. Korak 1 predstavlja branje trenutne slike iz videokamere. Korak 2 predstavlja pogoj, ki se uporablja za optimizacijo detekcije obrazov in izbira tip zaznave obrazov. Korak 3 predstavlja osnovno detekcijo obrazov na sliki in korak 5 povezovanje obrazov na zaporednih slikah video posnetka. Zaradi kompleksnosti zaznavanja obrazov in posledično velike porabe procesorskega časa je bilo potrebno optimizirati ta postopek. Za optimizacijo smo si izbrali metodo zmanjšanja opazovanega območja ROI, ki je v koraku 6. Pred tem se v koraku 4 oceni položaj obraza, ki ga iščemo. Nato sledi zanka od koraka 7 do koraka 13. V koraku 8 za vsako osebo preverimo, če je bil zaznan njen obraz na zadnji sliki videoposnetka. Če ni bil, se v koraku 11 sledi osebi kot je opisano v poglavju 3.4. Če pa je bila oseba zaznana, se v primeru prvih petih zaznav obraza (korak 9) osebi v koraku 10 klasificira spol (poglavje 3.5). V koraku 12 se posodobijo točke za sledenje, ki je opisano v poglavju 3.4. Pogoj v koraku 14 predstavlja izhod iz glavne zanke. Pred zaključkom aplikacije se izvede korak 15, ki predstavlja zapis pridobljenih podatkov v datoteko in je podrobneje opisan v poglavju 3.6.

## 3.3 Zaznava osebe na video posnetku

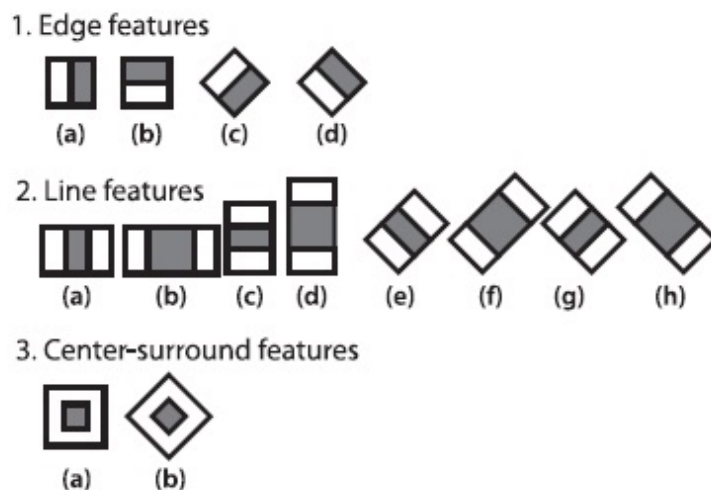
Za zaznavo obrazov smo uporabljal metodo Viola-Jones [22], ki je v delu [12] pokazala najboljše rezultate pri zaznavi obrazov. Metoda namesto slikovnih točk preverja značilnice Haar. Viola-Jones preišče integralno sliko s pravokotnimi regijami (slika 3.2 in slika 3.3) za iskanje Haar značilnic. Značilnice se preverjajo z izračunom razlike povprečnih vrednosti različnih regij (kvadrati različnih barv na sliki 3.2 in sliki 3.3). Slika 3.2 predstavlja tri različne značilnice Haar. Ločijo se po številu regij. Poznamo značilnice z dvema, tremi in štirimi regijami. Knjižnica OpenCV, ki jo uporabljamo pri izgradnji



Slika 3.1: Diagram poteka delovanja sistema.



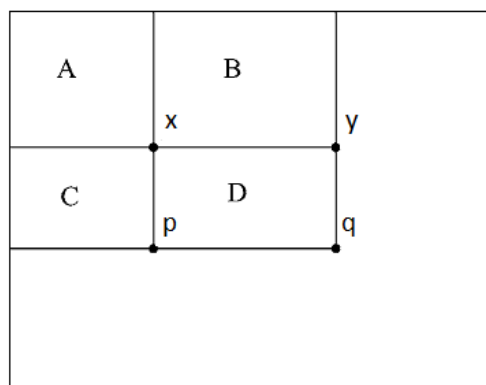
Slika 3.2: Trije tipi značilnic Haar glede na število regij.



Slika 3.3: Značilnice Haar v OpenCV.

našega produkta, uporablja večje število značilnic. Na sliki 3.3 je predstavljena množica značilnic, ki se uporablja pri Viola-Jones zaznavi objektov v knjižnici OpenCV. V fazi učenja se določijo vrednosti, ki so primerne pri iskanju določenega predmeta, v našem primeru obraza. Za učenje se uporablja metoda AdaBoost [25], ki uporablja množico šibkih klasifikatorjev in iz njih zgradi močan klasifikator.

Integralna slika vsebuje podatke slike primerne za izračun vsot regij na sliki. Vsaka točka vsebuje vsoto vseh točk, ki se nahajajo levo in višje od te točke. Tako za izračun vsote točk ene regije potrebujemo le vrednosti štirih točk in ni potrebno sešteti cele regije, kot se bi to izvajalo na originalni sliki.

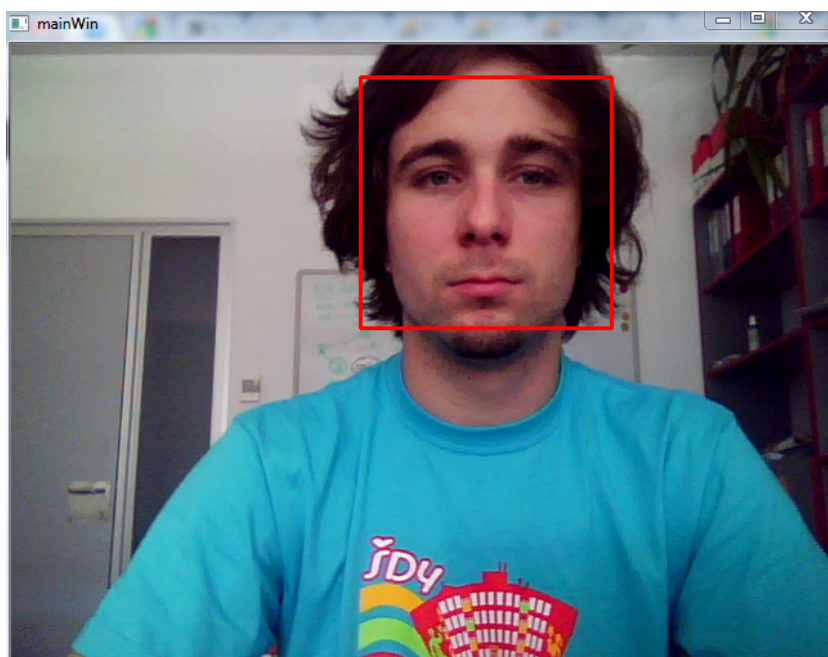


Slika 3.4: Integralna slika.

Na sliki 3.4 lahko razberemo, kako je sestavljena integralna slika. Na lokaciji  $x$  je vsota vrednosti slikovnih točk iz regije A. Vrednost na lokaciji  $y$  je vsota regij A in B, na lokaciji  $p$  je vsota regij A in C in na lokaciji  $q$  je vsota regij A, B, C in D. Vsoto regije D lahko nato izračunamo, če odštejemo vsoti regij  $y$  in  $p$  od vsot regij  $q$  in  $x$  ( $(q+x)-(y+p)$ ).

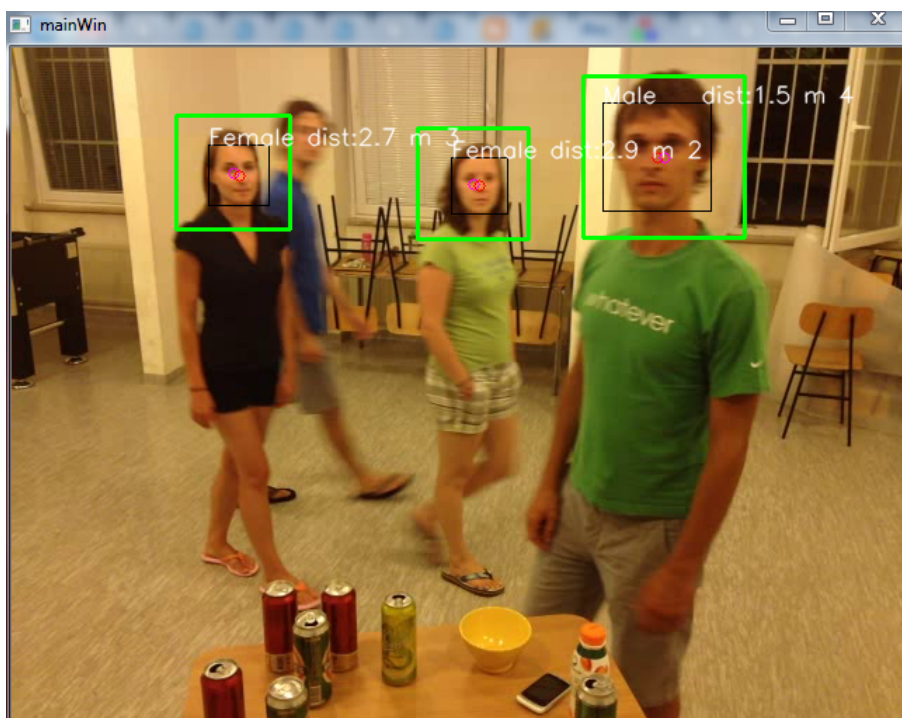
Metoda Viola-Jones deluje tako, da preišče celotno sliko z iskalno regijo Haar značilnic. Vnaprej se določi minimalna in maksimalna velikost iskanega objekta. Z optimalno nastavitvijo tega parametra lahko prihranimo nekaj časa, če vemo, v kakšnem območju se bodo gibali ljudje in lahko sklepamo o velikosti obrazov. Pomemben parameter pri zaznavi je tudi število minimalno zaznanih sosedov. S tem parametrom uravnavamo učinkovitost zaznave. Kadar imamo zelo nizko vrednost, nam metoda zazna obraze tudi tam, kjer ti niso prisotni, če pa je vrednost prevelika, pa lahko omenjena metoda izpusti slabše vidne obraze. Slika 3.5 predstavlja primer zaznave obraza z metodo Viola-Jones.

Metoda Viola-Jones zaradi svoje kompleksnosti porabi večino procesorskega časa, zato je bilo potrebno delovanje metode nujno optimizirati. Ker je videoposnetek zaporedje slik in nas zanimajo obrazi na videoposnetku, bi morali za vsako sliko na posnetku izvajati zaznavo obrazov in nato še preverjati, kateri obraz pripada kateri osebi. Pri optimizaciji smo se odločili za uporabo metode ROI. Ta metoda optimizacije deluje tako, da preveri samo okolico



Slika 3.5: Zaznava obraza z algoritmom Viola-Jones.

območja na sliki, kjer se je na prejšnji sliki nahajal obraz. Iskanje obrazov z metodo ROI se uporablja v kombinaciji s klasičnim iskanjem obrazov po celotni sliki, ki zaznava obraze oseb, ki so prišli v kader kasneje. Če celotno sliko preiščemo enkrat na sekundo, pri kakovosti podatkov ne izgubimo skoraj nič, pri zmogljivosti pa pridobimo dovolj, da lahko izvajamo zaznavo v realnem času. V primeru, da izvedemo zaznavo obrazov na celotni sliki vsakih pet zaporednih slik video posnetka, smo povečali število obdelanih slik na sekundo za približno trikratno vrednost kot pri obdelavi brez metode ROI. Pridobitev seveda ni konstantna, saj je pri zaznavi večjih obrazov tudi število ROI večje, s tem pa se poveča tudi število možnih lokacij obraza. Slika 3.6 prikazuje primer zaznave obrazov s pomočjo metode ROI. Zeleni okvirji predstavljajo ROI območje, črni pa dejanski obraz, ki je bil zaznan v tem območju. V središču obrazov sta še dva kroga, ki predstavljata smer premika obraza. Rdeči krog je središče trenutno zaznanega obraza, rožnati pa je središče predhodno zaznanega obraza. Poleg tega je na sliki vidna ocena spola za posamezen obraz (ki bo podrobno opisana v poglavju 3.5), ocenjena



Slika 3.6: Prikaz detekcije obrazov in območja ROI.

razdalja od kamere in zaporedna številka zaznanega obraza.

Ker se ljudje lahko premikamo z neko maksimalno hitrostjo, lahko določimo območje, kjer se lahko nahaja obraz iste osebe v naslednji sliki videoposnetka. Če predpostavimo, da se lahko človek v normalnih pogojih premika s hitrostjo 1,5 m/s, dobimo pri množenju s časom med dvema zaporednima slikama na video posnetku razdaljo, ki nam pove, za koliko se je maksimalno premaknil obraz na sliki:

$$\Delta t * 1,5m/s = \Delta d.$$

Na sliki razdalje niso podane v metrih, zato moramo pred izračunom razdalje definirati, kakšno razdaljo predstavlja ena slikovna točka na sliki. Upoštevati moramo tudi razdaljo osebe od kamere, ki pa jo je brez stereo kamere težko natančno določiti. Razdalja pa se lahko oceni glede na velikost obraza in vidni kot kamere, podobno, kot je opisano v delu [19]. Ko imamo okvirno razdaljo osebe, lahko tudi določimo maksimalen premik v vse strani. Majhen

prihranek časa lahko ustvarimo, če spremljamo zgodovino premikov osebe. Na tak način lahko predvidevamo pozicijo osebe na naslednji sliki s pomočjo Kalmanovega filtra [24].

Kalmanov filter [24] je razvil R. E. Kalman leta 1960. Namen razvoja je bil izboljšati natančnost vodenja vesoljskih raket v programu Apollo. Filter odstranjuje šum iz meritev in je uporaben za različne namene. Uporablja se predvsem pri navigaciji in odstranjuje šum pri sprejemu podatkov iz GPS satelitov. Kalmanov filter je sestavljen iz več matematičnih enačb, ki omogočajo optimalno obdelavo podatkov. Podatki se obdelujejo rekurzivno. Vse predhodne izračune shranjuje v sistemski matriki, ki jih obnavlja ob vsakem opazovanju. Algoritem oceni dejansko lokacijo iz uteženega povprečja napovedanih in izmerjenih vrednosti. Utež je odvisna od standardnih napak teh vrednosti. V našem produktu s Kalmanovim filtrom ocenjujemo položaj in velikost iskalnega okna pri zaznavi obrazov z metodo ROI. Z njim ocenjujemo tudi položaj obraza, kadar je temu nemogoče slediti, kar je podrobneje opisano v naslednjem poglavju.

### 3.4 Sledenje oseb

Sledenje obrazom oseb je v našem produktu ključnega pomena za pravilno analiziranje podatkov. V primeru, da je na posnetku več oseb, je potrebno natančno vedeti, katera oseba je kje.

Testirali smo tri postopke.

Na začetku smo napisali preprosto metodo za povezovanje obrazov na dveh zaporednih slikah. Vsaka posamezna oseba v kadru predstavlja svoj objekt v podatkovni strukturi aktivnih oseb. V njej so shranjeni tudi podatki o zadnji lokaciji obraza in čas zadnje osvežitve podatkov. Ob zaznavi novega obraza, preverimo lokacije vseh obrazov in če obstaja obraz s primerno lokacijo, se ta osveži s trenutno lokacijo. Če se ne najde obraza s primerno lokacijo, se v po-

datkovno strukturo aktivnih oseb doda nova oseba. Algoritem 1 predstavlja osnovno zamisel povezovanja obrazov na posnetku.

---

**Algoritem 1:** Povezovanje zaznanih obrazov na sliki s predhodno zaznanimi obrazi.

---

```
seznamNovihObrazov =
zaznavaObrazovNaSliki(trenutnaSlikaNaVideoPosnetku);
foreach zaznaniObraz in seznamNovihObrazov do
    obrazJeNajden = false;
    foreach aktivniObraz in seznamAktivnihObrazov do
        if (zaznaniObraz.položaj < aktivniObraz +
            maxOdmik*(zaznaniObraz.čas-aktivniObraz.čas)) &&
            (zaznaniObraz.položaj > aktivniObraz - maxOdmikNaSekundo*(
                zaznaniObraz.čas-aktivniObraz.čas)) then
            aktivniObraz.posodobiZ(zaznaniObraz);
            obrazJeNajden = true;
            break;
        end
    end
    if !obrazJeNajden then
        seznamAktivnihObrazov.push(zaznaniObraz);
    end
end
```

---

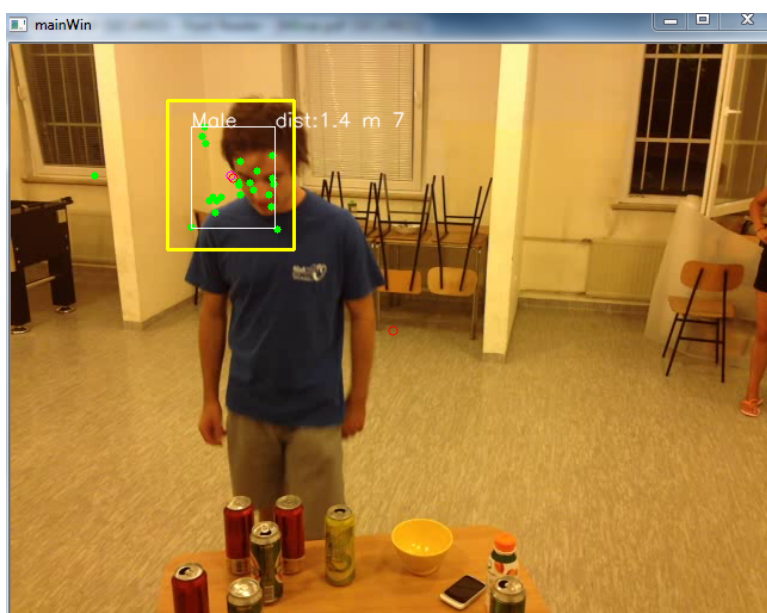
Metoda Camshift [3] je izpeljana iz metode Meanshift in pomeni: Meanshift, ki se vseskozi prilagaja. Camshift metoda poleg sledenja preverja velikost objekta. Objektom sledi z iskanjem območja z enakim histogramom, kot ga je imel objekt na prejšnji sliki.

Metoda Lucas-Kanade (LK) [11] je namenjena sledenju objektov na sliki. Objektom sledi z opazovanjem optičnega toka. Optični tok si lahko predstavljamo z množico vektorjev na sliki, ki predstavljajo premike slikovnih točk. Pri uporabi se najprej določijo točke, ki jim bo metoda sledila. V knjižnici OpenCV je funkcija, ki določi, katere točke so primerne za sledenje, to so navadno robovi ali točke, ki izstopajo iz slike. Na obrazu so to oči, vrh nosu,

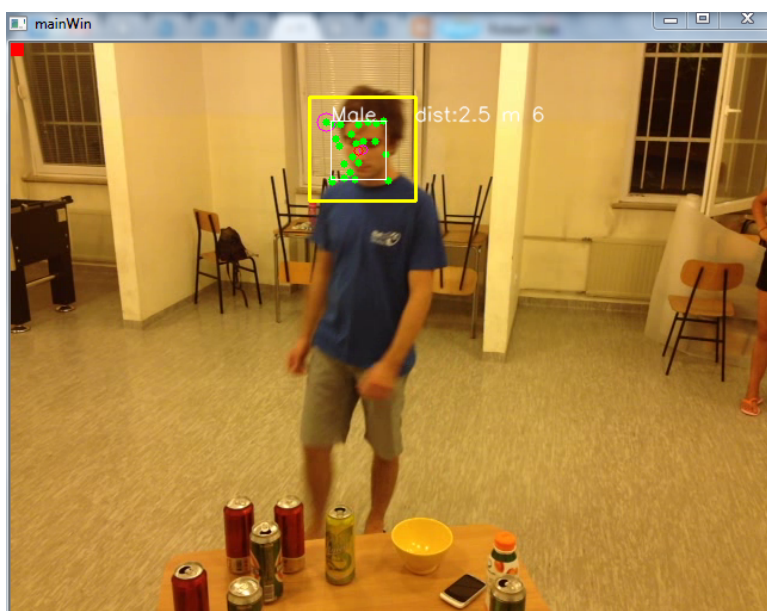
rob ustnic, rob lasišča itd. To metodo smo vgradili v naš produkt. Njena uporaba je v nadaljevanju podrobneje opisana.

Sprva smo ob zaznavi obraza preverjali, če je bol morda v prejšnji sliki kakšen obraz blizu najdenega. Stvar dobro deluje v primeru, da se obraza nahajata daleč narazen, ob prekrivanju pa lahko pride do zamenjave obrazov. Z implementacijo ROI se je ta težava odpravila, kajti pri iskanju obraza na omejenem območju iščemo točno določen obraz. Problem nastane, ko zaradi različnih dejavnikov obraza ni mogoče več zaznati. V knjižnici OpenCV je več rešitev za analizo premikanja in sledenje objektov. V našem primeru se je najbolje odrezalo sledenje z optičnim tokom Lucas-Kanade. Preizkusili smo tudi Camshift metodo, a smo to metodo zaradi slabih začetnih rezultatov kmalu opustili. Pri metodi Lucas-Kanade najprej določiš točke, ki jih želiš slediti. Metodi podaš prejšnjo sliko, kjer so izbrane točke in trenutno opazovano sliko na videoposnetku. Kot rezultat vrne množico točk, ki predstavljajo te iste točke na trenutni sliki. Točke smo določili samo v območju obraza, saj nas zanima le premik obraza. Metoda sledenja nam je vrnila le točke, kjer naj bi se obraz nahajal po premiku, torej na naslednji sliki. Okrog vseh točk smo narisali navidezen kvadrat in tako dobili kvadrat, kjer se nahaja obraz. Problem je nastal v primeru, ko vse točke niso sledile obrazu in v takem primeru je bil okvir, ki naj bi predstavljal obraz, prevelik. Primer prevelikega okvirja predstavlja beli okvir okoli obraza na sliki 3.7. S pomočjo konstante, ki predstavlja maksimalno hitrost človeka, smo določili, kako hitro se lahko ta okvir premika, širi in manjša. Določili smo območje, kje se lahko obraz nahaja. Če se je katera točka nahajala izven tega območja, smo jo izbrisali. Zelene točke na sliki 3.7 in na sliki 3.8 prikazujejo točke za sledenje. Beli okvir je navidezno območje obraza, ki ga predvidimo s pomočjo točk za sledenje. Rožnat krog okoli zelene točke (levi zgornji kot belega okvira na sliki 3.8) nam predstavlja nepravilno zaznano točko (angl. outlier), ki opisuje napačen podatek in se zaradi tega ne upošteva. Rumeni okvir je območje ROI.

Razširjena metoda sledenja zadovoljivo sledi obrazom, kadar jih ni mogoče zaznavati z Viola-Jones metodo. A metoda sledenja deluje samo v primeru,



Slika 3.7: Prikaz sledenja obraza z LK metodo brez odstranjevanja nepravilno zaznanih točk.



Slika 3.8: Prikaz sledenja obraza z LK metodo z odstranjevanjem nepravilno zaznanih točk.

kadar je sleden objekt viden na sliki. Mi pa imamo produkt, ki sledi obrazom oseb v trgovini, kjer je na zelo majhnem prostoru veliko ljudi. Zato prihaja do prekrivanja obrazov in je bilo potrebno implementirati zaznavanje prekrivanja. Prekrivanja smo ločili na dva tipa: obraz-obraz in obraz-telo. Zaznavanje prekrivanja obrazov je enostavno za implementacijo, saj vemo, kje se nahajajo obrazi in tudi okvirna razdalja od kamere je znana. Zato smo osebo, ki je dlje od kamere, dali v posebno strukturo, kjer so osebe, ki jim ni mogoče slediti. Ob ponovnem zaznavanju smo preverili, če se je kje v bližini našel obraz. S pomočjo Kalmanovega filtra [24] smo napovedali, kje se lahko nahaja obraz, da smo v primeru, če se je na istem območju zakrilo več obrazov, povezali pravilne obraze. Ugotoviti, kdaj se obraz nahaja za telesom druge osebe, je bolj kompleksen problem. Z našo rešitvijo ne zaznavamo telesa. Širino telesa bi lahko izračunali z množenjem širine obraza z določeno konstanto, a bi ta rešitev delovala le v primeru, če bi imeli ljudje isto telesno strukturo. Odštevanje ozadja od trenutne slike tudi ni uporabno, saj je v trgovinah dokaj nekonstantna svetloba in bi bila izhodna slika neuporabna. Zato smo se osredotočili le na spremljanje točk za sledenje z optičnim tokom. Spremljali smo, če se je zgodilo kaj nepredvidljivega, kot npr.: hitra sprememba smeri, izginjanje točk ipd. V teh primerih smo označili sledenje kot nepravilno in s Kalmanovim filtrom začeli napovedovati, kje se nahaja oseba.

### 3.5 Klasifikacija oseb po spolu

Za klasifikacijo spola je potrebno pridobiti učno množico slik, s katerimi se klasifikator nauči ločevati spola oz. določi razred, v katerega sodi določena oseba. Uporabili smo bazo slik iz Univerze Essex (VB) [9]. Baza vsebuje 7900 slik 395 posameznikov. Osebe so predvsem moškega spola in starosti 18-20 let. Iz baze smo izbrali 40 slik za posamezen spol. Polovico smo jih uporabili za učno množico, polovico pa za testno množico. Za tako veliko učno množico smo se odločili na podlagi rezultatov iz dela [19], kjer je dosežena največja klasifikacijska točnost pri 25 učnih slikah za spol. Sicer se klasifikacijska

točnost ni dosti razlikovala od množice velikosti 15 slik za spol. Odvisno od klasifikatorja se je ta razlikovala od 1 % do 4 %. Ker smo imeli težave pri zbiranju slik obrazov oseb ženskega spola, smo jih lahko pridobili le 40. Učna množica 20 slik za spol nam je dala zadovoljivo klasifikacijsko točnost, primerljivo z rezultati iz dela [19]. Slike smo iz RGB pretvorili v črno-bele ter uredili (zasukali in jim spremenili velikost) tako, da so imele osebe oči na isti višini in širini. Potem smo jih obrezali, da je bil na sliki le obraz. Nato smo te slike normalizirali in pretvorili v enodimenzionalen vektor [19]. V postopku normalizacije se vrednost vsake slikovne točke v posamezni sliki deli z vrednostjo maksimalne točke na sliki. Vrstice slike postavimo zaporedno v vektor. Na izhodu imamo torej enodimenzionalen vektor, ki ima največjo vrednost 1. Ko imamo vse slike normalizirane in vsako v svojem enodimenzionalnem vektorju, jih združimo v matriko dimenzij  $(K \times (M*N))$ , kjer je  $K$  število slik in  $M*N$  število slikovnih elementov na sliki. S pomočjo redukcije dimenzij dobimo značilnice, ki najboljše predstavljajo obraze. V poglavju 3.5.1 je podrobneje prikazana redukcija dimenzij in nato v 3.5.2 učenje klasifikatorja. Ko imamo naučen klasifikator, lahko začnemo z oceno spola. Spol ocenjujemo ob zaznavi novih obrazov, in sicer ob prvih petih zaznavah obraza, nato osebi dodelimo spol, ki ji je bil največkrat ocenjen. Z več ocenami spola se tako izognemo nepravilnim ocenam zaradi šuma na posnetku. Da spol ocenjujemo le petkrat, pa smo se odločili zaradi optimizacije programa. Ker je spol konstantna lastnost oseb, tudi ni potrebno ocenjevati spola skozi vse zaznave obraza osebe.

### 3.5.1 Redukcija dimenzij

Redukcija dimenzij se uporablja za zmanjševanje števila dimenzij mnogodimenzionalnih problemov. Izbrali smo metodo PCA [21]. Pred obdelavo smo slike pretvorili v vektorje. Prvotna slika je dimenzije  $M*N$ , tako dobimo vektor:

$$VektorSlike = (x_1, x_2, x_3 \dots x_{M*N})$$



Slika 3.9: Povprečen obraz.

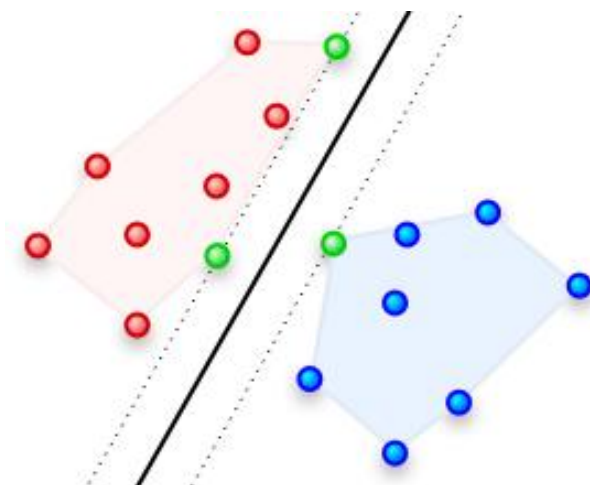
V učni množici imamo  $K$  slik. Vsaka slika je torej v svojem vektorju dolžine  $M \cdot N$  elementov. Nato je potrebno vse vektorje slik združiti v matriko:

$$MatrikaSlik = \begin{pmatrix} VektorSlike1 \\ VektorSlike2 \\ VektorSlike3 \\ \dots \\ VektorSlikeK \end{pmatrix}$$

Ta je vhodni parameter v PCA metodo. Metoda izračuna povprečne vrednosti za vse dimenzije, torej za vsako slikovno točko na vseh slikah (slika 3.9), ter nas pripelje do kovariančne matrike dimenzije  $K \times K$ . Iz kovariančne matrike nato izračuna lastne vrednosti in lastne vektorje [13], ki jih nato uporabljamo pri učenju klasifikatorja.

### 3.5.2 Klasifikator

Knjižnica OpenCV vsebuje tudi implementacijo klasifikatorja SVM [4], ki smo ga uporabili pri razvoju našega produkta. Vsak obraz je predstavljen z vektorjem v vektorskem prostoru. Metoda podpornih vektorjev nato med vektorji (pike na sliki 3.10) različnih razredov določi hiper-ravnino. Vektorjem, ki se nahajajo najbližje hiper-ravnini (zelene pike na sliki 3.10), se razdalja do hiper-ravnine poveča na maksimum. To naredimo tako, da hiper-ravnino postavimo v tak položaj, da so vsi ti vektorji obeh razredov



Slika 3.10: Prikaz vektorjev in hiper-ravnine.

enako oddaljeni od nje. Te vektorje tudi imenujemo podporni vektorji. S tem povečamo širino med razredoma in povečamo zanesljivejšo klasifikacijo. SVM je opisan z množico učnih primerov, kjer  $x_i$  predstavlja sliko in  $y_i$  spol:

$$\{(x_1, y_1), \dots, (x_m, y_m)\}, x_i \in X, y_i \in \{-1, 1\}$$

Pred učenjem iz reduciranih podatkov, ki smo jih pridobili s PCA metodo, je potrebno pripraviti vektor, ki vsebuje informacije o tem, v kateri razred sodi katera slika in SVM parametre za učenje. Pri učenju smo uporabili prizete parametre, ki vsebujejo standardni SVM klasifikator in linearno jedro. Standardni klasifikator se uporablja za klasifikacijo problemov, ki vsebujejo dva razreda. Nato smo izvedli učenje.

## 3.6 Priprava izhodnih podatkov

Namen našega produkta je, da bi z njim čim bolj optimalno sledili ljudem na sceni in jih klasificirali v razrede ter zaznali njihove odzive. Zbrane podatke pa shranili na način, da bo berljiv in v obliki za nemoteno nadaljnje procesiranje. Odločili smo se, da podatke pripravljamo v dveh stopnjah. V prvi stopnji združimo podatke po osebah, da se bodo kasneje lahko dodatno

```
<root>
  <person id="0">
    <data>
      <line id="0">
        <dist>2.448979591836735</dist>
        <time>9683</time>
        <x>86</x>
        <y>122</y>
        <type>0</type>
        <gender>1</gender>
      </line>
      <line id="1">
        <dist>2.448979591836735</dist>
        <time>10026</time>
        <x>86</x>
        <y>122</y>
        <type>0</type>
        <gender>1</gender>
      </line>
      ...
    </data>
  </person>
  <person id="1">
    ...
  </person>
</root>
```

Slika 3.11: Primer izpisa podatkov v XML formatu.

obdelali po želji naročnika. V drugi fazi pa iščemo že vnaprej določene odzive potrošnikov.

### 3.6.1 Združevanje podatkov

V prvi fazi zbiramo podatke o vsaki osebi posebej. Shranjujejo se podatki o položaju, čas zaznave, razdalja, spol in tip zaznave. Tip zaznave nam pove, na kakšen način smo določili koordinate položaja osebe na sliki. V našem primeru imamo zaznavo obraza, sledenje in prekrivanje. Podatke osvežujemo v vsakem ciklu v glavni zanki. Shranjujemo jih v strukturo Property Tree iz knjižnice Boost, ki ima zmožnost generiranja XML kode. Podatki o vsaki osebi se zapišejo, kadar gre oseba izven kadra in se izbriše iz seznama aktivnih oseb. Na sliki 3.11 je prikazan primer XML kode.

### 3.6.2 Določanje iskanih aktivnosti

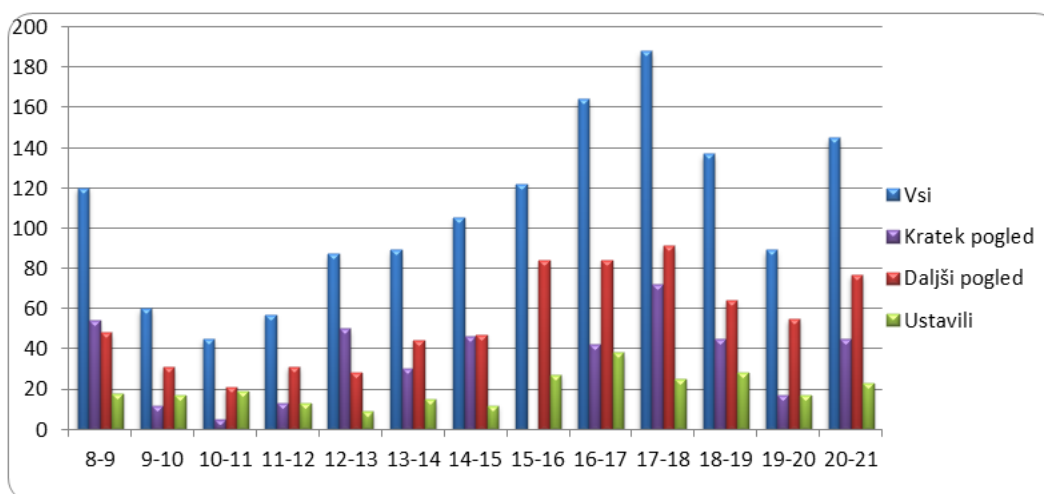
Ob zaključku delovanja programa se v naslednji fazi podatki združijo po že vnaprej določenih scenarijih. Predvideli smo tri scenarije oziroma aktivnosti: kratek pogled, daljši pogled in oseba se ustavi pri promocijski točki. Kratek pogled predstavlja zaznavo obraza krajšo od dveh sekund in ne predstavlja zanimanja za promocijo. Daljši pogled je pogled, ki traja več od dveh sekund, v tem času pa zaznava ne sme biti prekinjena za več kot pol sekunde. Da se je oseba ustavila pri promocijski točki ugotovimo tako, da zaznavamo obraz več kot pet sekund, ob pogoju, da je oseba bližje od dveh metrov in v tem časovnem intervalu zaznava obraz ne sme biti prekinjena za več kot dve sekundi. Podatke preverjamo za vsako osebo posebej. Na koncu te podatke zapišemo v datoteko. Primer zapisa take datoteke je prikazan na sliki 3.11. Podatki iz leve proti desni predstavljajo: zaporedno številko osebe, čas začetka aktivnosti v milisekundah, čas trajanja aktivnosti v milisekundah, oddaljenost od kamere v metrih, spol in aktivnost. Čas začetka aktivnosti predstavlja čas, ki je pretekel od začetka analize. Podatki o pogledih se zbirajo ločeno od podatkov, ko se je oseba ustavila pri promocijski točki. Pri izpisu so najprej osebe, ki so pogledale proti kameri, razvrščene po zaporedni številki. Ker osebi sledimo, kadar njenega obraza ni mogoče zaznati, lahko pride do primera, ko se pojavi oseba z isto zaporedno številko večkrat. V času, ko obraza te osebe nismo zaznali, pa se lahko zazna nova oseba z višjo zaporedno številko in s časom začetka aktivnosti manjšim od osebe z nižjo zaporedno številko. Zato ni nujno, da je pri izpisu čas začetka aktivnosti naraščajoč. Nato se zberejo podatki o osebah, ki so se ustavile pri promocijski točki. Podatke iz take datoteke lahko enostavno uvozimo v program za delo z razpredelnicami in jih tako prikažemo na grafih. Pred risanjem grafa se tudi preveri, če se je oseba z isto zaporedno številko pojavila večkrat. V tem primeru se ji dodeli aktivnost z največjo utežjo. Ostale manj pomembne aktivnosti te osebe se zanemari. Največjo utež predstavlja aktivnost, pri kateri se je oseba ustavila pri promocijski točki. Manjšo utež ima daljši pogled, kratek pogled pa ima izmed naštetih najmanjšo utež. Na

```

0,34469,4400,2.448980,1,2
1,60506,0,2.142857,0,1
2,66807,3249,2.666667,0,2
3,75375,0,1.621622,0,1
4,80270,5149,2.553191,0,2
5,97119,3495,2.352941,1,2
4,84747,9097,1.967213,0,3

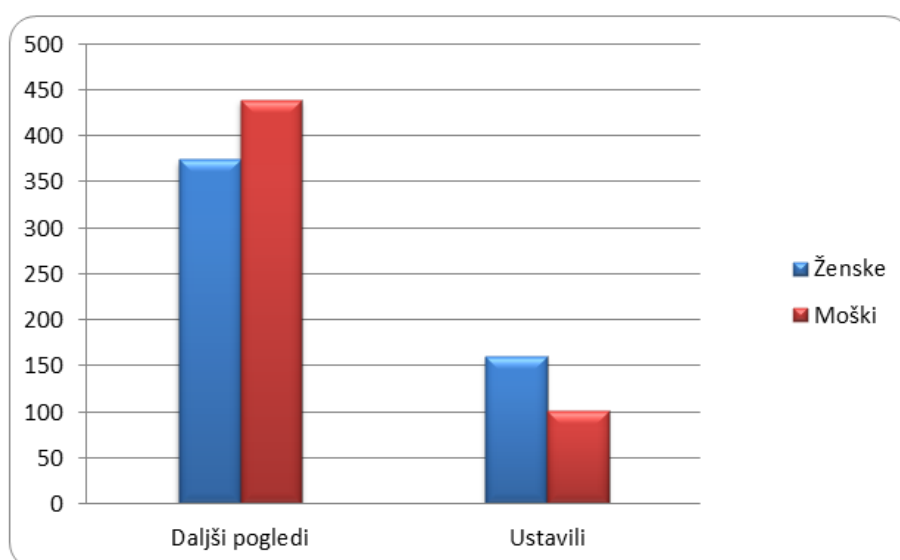
```

Slika 3.12: Primer zapisa podatkov v izhodno datoteko.



Slika 3.13: Prikaz količine potrošnikov in njihovih odzivov na promocijo ob določenih urah.

sliki 3.13 je primer grafa, ki prikazuje količino ljudi z določenimi odzivi ob določenih uri. Graf na sliki 3.14 pa prikazuje odziv na promocijo glede na spol. Graf ne vsebuje kratkih pogledov, ker se pri njih klasifikacija izvede premlokrat, da bi lahko potrdili določen spol. (Grafa na sliki 3.13 in sliki 3.14 ne prikazujeta podatkov iz realnega okolja.)



Slika 3.14: Prikaz odziva na promocijo glede na spol.

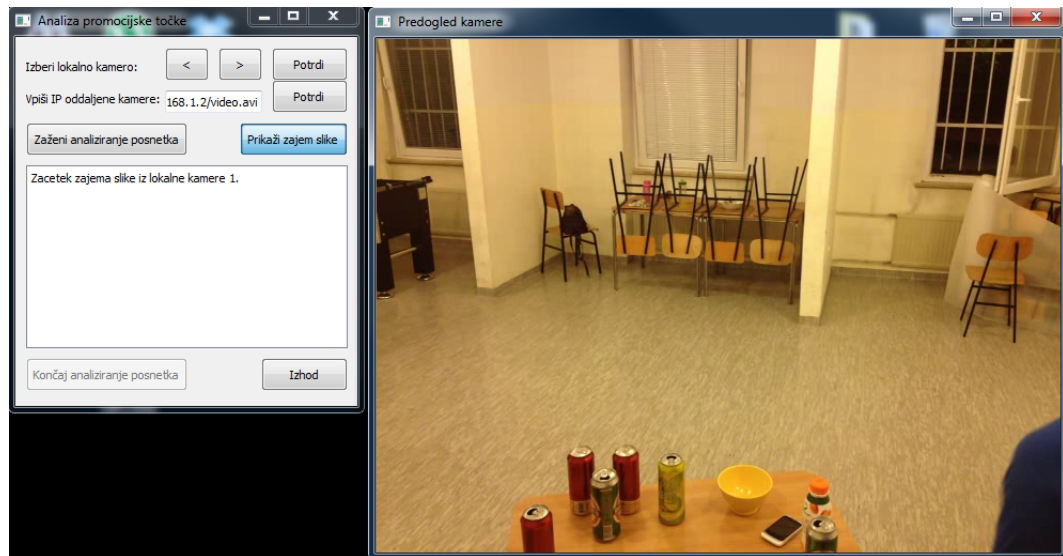


## Poglavje 4

# Uporaba produkta

Na sliki 4.1 je prikazan grafični uporabniški vmesnik. Ob zagonu programa se odpre okno, ki je levo na sliki. Če se za zajem videa uporablja kamera, ki je priklopljena na računalnik preko USB priključka, ali pa je vgrajena v računalniku, se ta izbere s smernima gumboma na vrhu aplikacije. Za tak način izbire kamere smo se odločili, ker knjižnica OpenCV ne omogoča identifikacije naprave. Pri izbiri kamere se izbere indeks kamere. S smernima gumboma se spreminja vrednost indeksa kamere. Za uporabo IP kamere se v polje za vnos naslova vpiše naslov, na katerem se nahaja kamera. To funkcionalnost podpira OpenCV. Običajno se poleg IP naslova vpiše tudi uporabniško ime in geslo, ki omejujeta dostop do kamere. Z gumbom Potrdi začnemo zajemati video z izbrane kamere. Delovanje kamere lahko preverimo z gumbom Prikaži zajem slike. Ob pritisku na gumb se nam odpre okno s trenutno sliko iz izbrane kamere. Če imamo izbrano napačno kamero, lahko spremenimo kamero za zajem videa. S ponovnim pritiskom na gumb se okno, ki prikazuje zajem posnetka, zapre. Ko imamo izbrano kamero pravilno nastavljeno, lahko začnemo analizo s pritiskom na gumb Zaženi analiziranje posnetka. Med analizo so onemogočeni vsi nepotrebni gumbi. Za zaustavitev analize se pritisne gumb Končaj analiziranje posnetka in nato gumb Izhod. V velikem oknu za tekst se izpisuje trenutno stanje programa.

Produkt je namenjen analiziranju kupnih navad potrošnikov. Tipični pro-



Slika 4.1: Uporabniški vmesnik.

stor za njegovo uporabo je trgovski center, kjer se giba večje število ljudi. Ker opazujemo dogajanje okoli promocijske mize, je potrebno smiselno postaviti kamero, da ta posname območje, ki ga želimo opazovati. Pogosto poleg promocijske mize stoji tudi promotor, ki lahko prekriva zajeto območje kamere. Zato je najbolj primerno, da se kamera postavi nekoliko višje nad promotorja in cca. 1 meter za njim, da se zajame tudi dogajanje pred mizo. Paziti moramo, da se kamera ne postavi previsoko, kajti v tem primeru se ne vidi obrazov dovolj dobro, da bi se jih lahko optimalno zaznavalo. Slika 4.2 predstavlja slabo postavitev kamere v primeru, da se za promocijsko mizo nahaja tudi promotor. V takem primeru se lahko kamera postavi nekoliko na eno od strani promocijske mize, promotor pa na drugo stran. Postavljena bi morala biti tudi nekoliko višje.



Slika 4.2: Primer slabega zajema slike za promotorjem.



# Poglavje 5

## Rezultati

### 5.1 Klasifikacija spola

Za klasifikacijo spola smo uporabljali bazo slik študentov iz Univerze Essex (VB) [9]. Med urejanjem slik za učno množico smo pripravil tudi nekaj testnih slik iz iste baze, da smo lahko takoj videli, kako uspešen je naš način klasifikacije spola. Pri testni množici 43 slik (12 ženskih in 31 moških) smo dosegli uspešnost 90,7 %. Ker se bo produkt vsaj na začetku poizkušal tržiti v Sloveniji, smo se odločili, da preverimo uspešnost klasifikacije na slovenski bazi slik. Odločil smo se za bazo slik CVL Face Database [7], ki je bila narejena v Laboratoriju za računalniški vid na Fakulteti za računalništvo in informatiko v Ljubljani. Baza vsebuje 114 barvnih slik (108 moških in 6 žensk). Ker tudi ta baza vsebuje pretežno moške obraze, smo dodali osem naključno izbranih slik prijateljic na socialnih omrežjih, tako je bila končna baza slik velika 122 slik (108 moških in 14 žensk). Da bi se izognili urejanju slik, smo napisali razred, ki avtomatično preverja uspešnost klasifikacije. Iz določene mape bere slike in na njih zaznava obraze. Območje obraza izreže in mu spremeni velikost na velikost slik v učni množici. Pri tej testni množici smo dosegli klasifikacijsko uspešnost 84,75 %. Količino slik in uspešnost glede na spol vidimo v tabeli 5.1. Razliko v rezultatu pripisujemo predvsem temu, da je bil klasifikator naučen nad podmnožico prve baze.

Aktualen spol \ Ocenjen spol	M	Ž
M	<b>87</b>	17
Ž	1	<b>13</b>

Tabela 5.1: Rezultati klasifikacije spola na testni bazi slik CVL FDB.

Št. oseb na sliki	Obdelava celotne slike			ROI		
	1	2	3	1	2	3
Iskanje obrazov	200 ms	200 ms	200 ms	45 ms	50 ms	55 ms
Klasifikacija spola	25 ms	50 ms	75 ms	25 ms	50 ms	75 ms

Tabela 5.2: Časovna pridobitev z implementacijo ROI.

## 5.2 Pridobitev z uvedbo ROI

V tabeli 5.2 so predstavljeni časi izvajanja detekcije obrazov pred in po uvedbi metode ROI. Pri zajemu slike smo uporabljali video posnetek ločljivosti  $640 \times 480$  slikovnih točk. Velikost enega ROI pa je bila  $100 \times 100$  slikovnih točk. Pri ROI metodi smo izvedli zaznavo obrazov na celi sliki vsako peto zaporedno sliko v video posnetku. Rezultat meritev pri ROI zaznavi je povprečen čas pri petih zaporednih slikah, saj ta rezultat najbolje prikazuje dejansko pridobitev z uvedbo metode. Pri zaznavi enega obraza smo tako pri času izvajanja pridobili več kot 400 %. Ta odstotek se niža z velikostjo vsote ROI območij. Z uvedbo ROI smo izenačili čas izvajanja iskanja obrazov in klasifikacije spola pri dveh osebah na sliki. Pri večjem številu oseb na posnetku pa se čas iskanja obrazov zmanjša pod čas klasifikacije spola.

## 5.3 Končne zmogljivosti obdelanih slik

Hitrost obdelave slike je pomembna iz razloga, ker se mora aplikacija izvajati v realnem času. Tabela 5.3 predstavlja nihanje obdelanih slik na sekundo glede na obdelave, ki jih izvaja program. Pri merjenju sledenja obrazom in klasifikacije spola je prisotno tudi iskanje obrazov. Ker produkta še ni-

Korak \ Št. oseb na sliki	1	2	3
Iskanje obrazov	12	11	11
Sledenje obrazom	10	9	8
Klasifikacija spola	9	6	4

Tabela 5.3: Število obdelanih slik na sekundo. Stolpci predstavljajo število obdelanih slik glede na število obrazov na sliki pri uporabi različnih funkcionalnosti programa. Zadnja dva koraka v meritvah vsebujeta tudi funkcionalnost iskanja obrazov.

smo preizkusili v realnem okolju, kjer se bo ta uporabljal, nimamo podatka, kakšna količina obdelanih slik bo zadostovala za uspešno delovanje. Sicer menimo, da bi zadostovale že 2 do 3 slike na sekundo, saj se potrošniki tipično gibljejo nekoliko počasneje kot povprečen človek med hojo. Ker osebam spol ocenjujemo le ob prihodu v območje, ki ga zajema kamera, bi lahko z lahkoto obdelovali tri klasifikacije sočasno. Med izvajanjem analize promocijske točke se največkrat uporablja iskanje obrazov, ki je aktivno tudi takrat kadar oseb ni na posnetku. Zato je bilo ključnega pomena, da smo to funkcionalnost dodatno optimizirali. Sledenje obrazom in klasifikacija spola pa se izvajata le po potrebi in njuna zahtevnost ne predstavlja pretirane obremenitve sistema.



## Poglavje 6

### Zaključek

V diplomskem delu smo predstavili razvoj produkta za analizo oseb v trgovskih centrih. Produkt deluje na osnovi računalniškega vida in s pomočjo kamere spremlja dogajanje v prostoru, zaznava osebe in jih klasificira v razele. Produkt je namenjen predvsem distributerjem, ki izvajajo promocije v trgovskih centrih in tako spremljajo odziv potrošnikov na nek nov izdelek. Pri spremljanju odziva potrošnikov trenutno spremljajo primarno število prodanih izdelkov. Pri tem ne štejejo vseh mimoidočih. Tako ne dobimo objektivnega rezultata prodanih izdelkov glede na mimoidoče potrošnike. Z našim produktom se bo lahko spremljalo širše dogajanje okoli promocijske točke. Osebam, ki se bodo odzvale na promocijo, se bo določalo spol in njihov odziv na promocijo. S tem bo promotor dobil širšo sliko odziva potrošnikov. Razvit produkt je sicer še v prototipni različici, a bi bil vseeno že primeren za predstavitevno uporabo. Z nekaj popravki pri napovedi in z izgradnjo grafičnega vmesnika za napredni prikaz zbranih podatkov, bi lahko z njim nastopili na trgu. Pri razvoju smo imeli največ težav pri sledenju oseb. Že pri izbiri algoritma, ki bi v našem okolju deloval optimalno, smo zapravili veliko časa. Določene algoritme smo preizkušali celo večkrat. Ker dejansko nismo dosegli pričakovanih rezultatov, smo se nato pri preizkusu drugih algoritmov naučili novih pristopov, ki so izboljšali delovanje ostalih algoritmov. V prihodnje imamo namen v produkt vključiti klasifikacijo oseb po starosti in preizkusiti

ostale metode za klasifikacijo spola, ki bi lahko dvignile odstotek uspešnih klasifikacij. Zgradili bomo tudi spletni grafični vmesnik po zgledu spletnih analitik, kjer lahko uporabnik v kratkem času spremeni način združevanja podatkov iz baze podatkov in tako spremeni način vizualizacije.

# Slike

3.1	Diagram poteka delovanja sistema. . . . .	15
3.2	Trije tipi značilnic Haar glede na število regij. . . . .	16
3.3	Značilnice Haar v OpenCV. . . . .	16
3.4	Integralna slika. . . . .	17
3.5	Zaznava obraza z algoritmom Viola-Jones. . . . .	18
3.6	Prikaz detekcije obrazov in območja ROI. . . . .	19
3.7	Prikaz sledenja obraza z LK metodo brez odstranjevanja nepravilno zaznanih točk. . . . .	23
3.8	Prikaz sledenja obraza z LK metodo z odstranjevanjem nepravilno zaznanih točk. . . . .	23
3.9	Povprečen obraz. . . . .	26
3.10	Prikaz vektorjev in hiper-ravnine. . . . .	27
3.11	Primer izpisa podatkov v XML formatu. . . . .	28
3.12	Primer zapisa podatkov v izhodno datoteko. . . . .	30
3.13	Prikaz količine potrošnikov in njihovih odzivov na promocijo ob določenih urah. . . . .	30
3.14	Prikaz odziva na promocijo glede na spol. . . . .	31
4.1	Uporabniški vmesnik. . . . .	34
4.2	Primer slabega zajema slike za promotorjem. . . . .	35



# Tabele

5.1	Rezultati klasifikacije spola na testni bazi slik CVL FDB. . . . .	38
5.2	Časovna pridobitev z implementacijo ROI. . . . .	38
5.3	Število obdelanih slik na sekundo. Stolpci predstavljajo število obdelanih slik glede na število obrazov na sliki pri uporabi različnih funkcionalnosti programa. Zadnja dva koraka v meritvah vsebujeta tudi funkcionalnost iskanja obrazov. . . . .	39



# Literatura

- [1] (2012) Produkti podjetja ACIC. Dostopno na:  
<http://www.acic.biz/en/products/overview.html>
- [2] G. Bradski, A. Kaehler, *Learning Open CV*, Sebastopol: O'Reilly, 2008.
- [3] G. Bradski, *Computer Vision Face Tracking For Use in a Perceptual User Interface*, *Intel Technology Journal*, 2(2), str. 12-21, 1998.
- [4] C. J. C. Burges, A Tutorial on Support Vector Machines for Pattern Recognition, *Data Mining and Knowledge Discovery*, 2(2), str. 121-167, 1998.
- [5] (2012) Knjižnica Boost. Dostopno na:  
<http://www.boost.org/>
- [6] (2012) Comland - Comsight. Dostopno na:  
<http://www.comsight.si/>
- [7] (2012) Baza slik CVL FDB. Dostopno na:  
<http://www.lrv.fri.uni-lj.si/facedb.html>
- [8] (2012) Eclipse CDT. Dostopno na:  
<http://www.eclipse.org/cdt/>
- [9] (2012) Baza slik Univerze Essex. Dostopno na:  
<http://cswww.essex.ac.uk/mv/allfaces/index.html>

- [10] (2012) Baza slik FERET. Dostopno na:  
[http://www.itl.nist.gov/iad/humanid/feret/feret\\_master.html](http://www.itl.nist.gov/iad/humanid/feret/feret_master.html)
- [11] A. Giachetti, Matching techniques to compute image motion, *Image and Vision Computing*, 18(3), str. 247-260, 2000.
- [12] T. Jagodnik, *Detekcija potnikov v avtomobilu*, diplomska naloga, Fakulteta za računalništvo in informatiko Univerze v Ljubljani, Slovenija, 2012. Dostopno na:  
<http://eprints.fri.uni-lj.si/1764/1/Jagodnik-1.pdf>
- [13] A. Košir, *Lastne vrednosti in lastni vektorji*, skripta, Fakulteta za matematiko in fiziko Univerze v Ljubljani, 2009. Dostopno na:  
<http://www.fmf.uni-lj.si/~kosir/poucevanje/skripta/lastne.pdf>
- [14] W. Kienzle, G. Bakir, M. Franz, B. Scholkopf, Face Detection – Efficient and Rank Deficient, *Advances in Neural Information Processing Systems*, zbornik 17, str. 673-680, 2005.
- [15] M. Mlinar, *Kategorizacija starosti iz slik obrazov*, diplomska naloga, Fakulteta za računalništvo in informatiko Univerze v Ljubljani, Slovenija, 2009. Dostopno na:  
[http://eprints.fri.uni-lj.si/876/1/Diplomsko\\_delo-Mlinar.pdf](http://eprints.fri.uni-lj.si/876/1/Diplomsko_delo-Mlinar.pdf)
- [16] M. Nilsson, J. Nordberg, I. Cleasson, Face Detection Using Local SMQT Features and Split Up SNoW Classifier, *Acoustics, Speech and Signal Processing (ICASSP)*, zbornik 2, str. 589-592, 2007.
- [17] (2012) Knjižnica OpenCV. Dostopno na:  
<http://opencv.willowgarage.com/wiki/>
- [18] (2012) Orange. Dostopno na:  
<http://orange.biolab.si/>
- [19] R. Ravnik, *Digitalna karakterizacija z uporabo računalniškega vida v realnem času*, diplomska naloga, Fakulteta za računalništvo in informatiko

- Univerze v Ljubljani, Slovenija, 2009. Dostopno na:  
[http://eprints.fri.uni-lj.si/865/1/Ravnik\\_R\\_UN.pdf](http://eprints.fri.uni-lj.si/865/1/Ravnik_R_UN.pdf)
- [20] (2012) Podjetje RetailNext. Dostopno na:  
<http://www.bviretailnext.com/>
- [21] (2012) L. I. Smith, A tutorial on Principal Components Analysis, 2002.  
Dostopno na:  
[http://www.sccg.sk/~haladova/principal\\_components.pdf](http://www.sccg.sk/~haladova/principal_components.pdf)
- [22] P. Viola, M. J. Jones, Robust Real-Time Face Detection, *International Journal of Computer Vision*, 57(2), str. 137–154, 2004.
- [23] (2012) Produkt VideoMining. Dostopno na:  
<http://www.videominig.com/>
- [24] G. Welch, G. Bishop, An Introduction to the Kalman Filter, *tehnično poročilo*, Department of Computer Science University of North Carolina at Chapel Hill, 2006.
- [25] M. Yang, J. Crenshaw, B. Augustine, R. Mareachen, Y. Wu, AdaBoost-based face detection for embedded systems, *Computer Vision and Image Understanding*, 114(11), str. 1116-1125, 2010.