

Univerza v Ljubljani

Fakulteta za elektrotehniko

JAN GODEC

**PRIMERJAVA SLIKOVNIH
DESKRIPTORJEV ZA SAMODEJNO
RAZPOZNAVANJE OBRAZOV**

Diplomsko delo univerzitetnega študija

Mentor: doc. dr. Vitomir Štruc, univ. dipl. inž. el.

Ljubljana, 2016

Zahvala

Rad bi se zahvalil mentorju doc. dr. Vitomirju Štrucu, univ. dipl. inž. el. za strokovno pomoč, delitev izkušenj in usmerjanje pri izvedbi eksperimentov in pisanju diplomske naloge.

Posebej bi se rad zahvalil svoji družini za potrpežljivost, vzpodbudo in stalno podporo med študijem in izdelavo zaključnega dela.

Vsebina

1 UVOD	13
1.1 Cilji diplomskega dela.....	14
1.2 Struktura diplomskega dela	15
2 TEORETIČNO OZADJE	17
2.1 Izračun značilnk z analizo glavnih komponent (PCA).....	18
2.2 SURF deskriptor.....	21
2.3 Binarni deskriptorji.....	24
2.3.1 BRISK deskriptor	25
2.3.2 FREAK deskriptor	25
2.4 Prileganje vzorcev in razdalje	27
2.5 Napake in ROC krivulje	28
3 METODOLOGIJA DELA	31
3.1 Oprema in orodja.....	31
3.2 Podatkovne zbirke	31
3.2.1 Zbirka XM2VTS.....	32
3.2.2 Zbirka Extended Yale B	32
3.3 Vrednotenje sistema razpoznavanja	33
4 EKSPERIMENTI	35
4.1 Eksperimenti na zbirki obrazov "XM2VTS"	35
4.1.1 Protokol.....	35
4.1.2 Rezultati in diskusija.....	38

4.2	Eksperimenti na zbirki obrazov "Extended Yale B"	43
4.2.1	Protokol	43
4.2.2	Rezultati in diskusija	44
4.3	Dolžine značilnih vektorjev	51
4.4	Časovna zahtevnost izračuna značilnih vektorjev	52
5	ZAKLJUČEK	53
	LITERATURA	57
	DODATEK	59
A	Kazalo slik	60
B	Kazalo tabel	60

Seznam uporabljenih kratic

XM2VTS	standardna zbirka slik obrazov (angl. Extended Multi Modal Verification for Teleservices and Security applications)
EYB	standardna zbirka slik obrazov (angl. Extended YaleB)
PCA	analiza glavnih komponent (angl. Principal Component Analysis)
SURF	algoritem SURF (angl. Speeded Up Robust Feature)
SURF+PCA	zmanjšanje dimenzij SURF deskriptorjev s PCA postopkom
FREAK	algoritem FREAK (angl. Fast Retina Keypoint)
FREAK+PCA	zmanjšanje dimenzij FREAK deskriptorjev s PCA postopkom
BRISK	algoritem BRISK (angl. Binary Robust Invariant Scalable Keypoints)
BRISK+PCA	zmanjšanje dimenzij BRISK deskriptorjev s PCA postopkom
SIFT	algoritem SIFT (angl. Scale Invariant Feature Transform)
FAST	detektor kotnih točk (angl. Features from Accelerated Segment Test)
AGAST	detektor kotnih točk (angl. Adaptive and Generic Accelerated Segment Test)
FAR	napaka potrditve vsiljivcev (angl. False Accept Rate)
FRR	napaka zavrnitve klientov (angl. False Reject Rate)
ROC	krivulja, ki prikazuje odvisnost med FAR in FRR (angl. Receiver Operating Characteristic)
HTER	srednja vrednost med FAR in FRR napakama (angl. Half Total Error Rate)
EER	napaka HTER, ko je FAR=FRR (angl. Equal Error Rate)

Povzetek

V diplomskem delu je predstavljena analiza računalniškega razpoznavanja obrazov z različnimi postopki. Funkcija, ki jo je opravljal biometrični sistem, je bila verifikacija oseb na podlagi razpoznavanja obrazov. Podano je teoretično ozadje, orodja in metode za izvedbo ter rezultati eksperimentov.

Razpoznavanje je bilo izvedeno na sivinskih slikah obrazov iz dveh standardnih podatkovnih zbirk, in sicer XM2VTS in EYB. V prvi zbirki so bili zajeti obrazi brez večjih motenj v osvetlitvi, v drugi zbirki pa se je postopoma večala zatemnitev obrazov. Na zbirki XM2VTS so bili preizkušeni PCA, SURF, FREAK, BRISK, SURF+PCA, FREAK+PCA in BRISK+PCA postopek. Na zbirki EYB so bili preizkušeni PCA, SURF, FREAK in BRISK postopek. Pri vseh postopkih se je razdalja med značilnimi vektorji merila na tri načine, in sicer z Evklidovo, City Block ter kosinusno razdaljo.

Rezultati razpoznavanja obrazov so prikazani v obliki tabel in grafov. Ovrednoteni so z napakami FAR, FRR, HTER in EER. Soodvisnost med FAR in FRR v različnih delovnih točkah ponazarjajo ROC krivulje. V tabelah so izpostavljeni rezultati verifikacije v treh delovnih točkah. Prikazana je tudi časovna zahtevnost izračuna posameznih značilnih vektorjev.

Ključne besede: razpoznavanje obrazov, verifikacija oseb, PCA, SURF, FREAK, BRISK, deskriptor, XM2VTS, EYB, ROC krivulja, Evklidova razdalja, City Block, kosinusna razdalja

Abstract

The diploma thesis presents an analysis of computer facial recognition with various procedures. The function performed by the biometric system was face verification on the basis of facial recognition. Theoretical background, tools, and methods for the experiments and results thereof are also presented.

Recognition was performed on grey pictures of faces from two standard databases, specifically XM2VTS and EYB. The first database included faces without any significant lighting interference, whereas the pictures in the second database were increasingly darker. PCA, SURF, FREAK, BRISK, SURF+PCA, FREAK+PCA, and BRISK+PCA were tested on the XM2VTS database, whereas PCA, SURF, FREAK, and BRISK were tested on the EYB database. In all procedures, the distance between feature vectors was measured using three methods, specifically the Euclidean, City Block, and cosine distance.

The results of facial recognition are shown in the form of tables and graphs. They are evaluated with FAR, FRR, HTER, and EER values. The correlation of FAR and FRR at various operating points are shown by ROC curves. The tables highlight the verification results in three operating points. The time requirement for the calculation of specific feature vectors is also shown.

Key words: facial recognition, face verification, PCA, SURF, FREAK, BRISK, descriptor, XM2VTS, EYB, ROC curve, Euclidean distance, City Block, cosine distance

1 UVOD

V sodobnem svetu vedno več sistemov za identifikacijo ter verifikacijo oz. avtentikacijo ljudi temelji na razpoznavanju unikatnih biometričnih lastnosti človeka, kot so oblika obraza, dlani, šarenice, mrežnice, prstni odtisi, barva glasu itd.

Razpoznavanje slikovnih vzorcev oz. v dotičnem primeru razpoznavanje obrazov, je podpodročje znanstvene vede, ki ji širše pravimo razpoznavanje vzorcev. Biometrični sistem je v osnovi sistem razpoznavanja vzorcev, ki deluje tako, da s tipali zajame biometrične podatke osebe, izlušči značilnosti zajetih podatkov ter primerja dobljene značilke z značilkami že prej zgrajene podatkovne zbirke sistema. Z izrazom značilka označujemo izluščeno značilnost zajetih podatkov. Ko značilke nanizamo v vektor, mu pravimo značilni vektor oz. vektor značilk. Odvisno od konteksta aplikacije lahko biometrični sistem opravlja funkcijo verifikacije ali identifikacije. Pri verifikaciji se ugotavlja, ali je oseba, ki se prijavlja v sistem, res oseba, za katero se izdaja. Značilke obraza prijavljajoče se osebe se primerjajo samo z značilkami ene osebe iz podatkovne zbirke. Če je podobnost značilk večja od določenega praga, je verifikacija pozitivna in je oseba prepoznana kot klient, če pa je podobnost premajhna, je oseba prepoznana kot vsiljivec in jo sistem zavrne. Pri identifikaciji se značilke obraza osebe, ki jo sistem razpozna, primerjajo z značilkami vseh oseb v podatkovni zbirki obrazov. Oseba na vhodu je identificirana kot tista oseba iz podatkovne zbirke, kjer je podobnost največja, ob pogoju, da je presežen prag. Če prag ni presežen, oseba ni identificirana. [1, 13]

V sistemu razpoznavanja obrazov lahko naletimo na različne težave v različnih fazah sistema. Takoj na začetku, pri zajemu vhodnih podatkov, je pomembno, da so le ti zajeti pri čimbolj podobnih pogojih, kot so bili v fazi registracije v sistem. Relevantni so osvetlitev, nagib in kot obraza pri zajemu slike, mimika, staranje klienta, čustva, ki se odražajo na obrazu itd. Problem lahko predstavlja tudi dimenzija vzorcev. Digitalne dvodimenzionalne sivinske slike so zapisane kot dvorazsežne matrike, katerih elementi predstavljajo nivo sivine istoležne slikovne točke slike. Matriko slike zapišemo kot vektor tako, da nanizamo vse vrstice matrike

eno za drugo [14]. Dobljeni vektor slike je dolg toliko, kolikor je zmnožek števila vrstic in stolpcev matrike. Običajno gre za vektor z več tisoč dimenzijami, kar ob velikem številu slik povzroča računsko zahtevnost in zamudnost. Ena izmed možnosti poenostavitve je preslikava vzorcev v podprostor z manj dimenzijami. Posledica je seveda določena stopnja izgube podatkov zaradi poenostavitve. Naslednji izziv je izbira takšnega algoritma za izpeljavo značilk in izračun razdalj podobnosti med vzorci, da bo sistem razpoznavanja čim bolj učinkovit in robusten.

V svojem diplomskem delu sem se omejil na razpoznavanje obrazov iz digitalnih sivinskih slik obrazov, ki so bile vzete iz dveh različnih podatkovnih zbirk [2, 3]. Prva se imenuje "XM2VTS", druga pa "Extended Yale B". Faza predhodne obdelave slik ni bila potrebna, saj so vse slike že normirane na standardne velikosti. Programa, ki sem ju napisal v programskem okolju Matlab za vsako zbirko posebej, sta sistema razpoznavanja obrazov, ki opravita verifikacijo oseb iz podatkovnih zbirk. Uporabil sem funkcije iz Matlaba ter Matlabove knjižnice orodij imenovane PhD Toolbox (Pretty helpful Development function for face recognition), katere avtor je mentor dr. Vitomir Štruc [4]. Zgoraj omenjeni zbirki slik obrazov sta standardni zbirki, na katerih raziskovalci, ki se ukvarjajo s področjem razpoznavanja obrazov, opravljajo eksperimente [15, 16, 17, 18, 19].

Raziskoval sem učinkovitost oz. uspešnost razpoznavanja obrazov različnih algoritmov za detekcijo in zapis značilk (deskriptorjev) ter različnih postopkov za izračun podobnosti med značilkami na slikah z različno stopnjo variabilnosti v osvetlitvi. V zbirki XM2VTS so bile slike z minimalno spremembo osvetlitve, medtem ko je v EYB obseg zatemnitev naraščal.

1.1 Cilji diplomskega dela

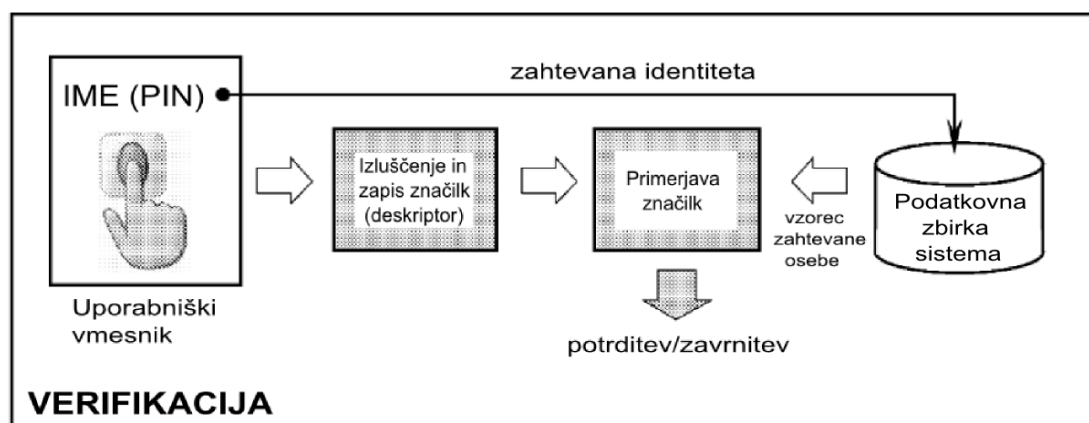
- Analiza in primerjava razpoznavanja obrazov s PCA postopkom ter SURF, BRISK in FREAK deskriptorji.
- Analiza in primerjava razpoznavanja obrazov s kombiniranimi postopki SURF in PCA, BRISK in PCA ter FREAK in PCA.
- Analiza in primerjava razpoznavanja obrazov z omenjenimi postopki na slikah z različno stopnjo osvetlitve.
- Analiza in primerjava različnih postopkov za izračun razdalj med značilnimi vektorji slik.
- Primerjava binarnih in ne binarnih deskriptorjev.

1.2 Struktura diplomskega dela

Diplomsko delo je razdeljeno na pet poglavij. Začetno poglavje je uvod, ki ga berete. Opisuje motivacije, cilje in strukturo diplomskega dela. V poglavju dva je predstavljeno teoretično ozadje in matematični postopki potrebni za izgradnjo sistema razpoznavanja z omenjenimi deskriptorji. Opisan je biometrični sistem, ki opravlja verifikacijo, nato je razložen postopek izračuna značilnega vektorja z analizo glavnih komponent ter SURF, BRISK in FREAK deskriptorja. Opisan je še način prileganja vzorcev, trije načini merjenja razdalj med značilnimi vektorji ter napake, ki nastanejo pri verifikaciji. V tretjem poglavju so na kratko nanizane specifikacije sistema, na katerem so bili eksperimenti izvedeni ter opisane uporabljene podatkovne zbirke in način vrednotenja sistema razpoznavanja. V poglavju štiri so predstavljeni in ovrednoteni rezultati eksperimentov na obeh podatkovnih zbirkah. Na koncu je zaključek z ugotovitvami.

2 TEORETIČNO OZADJE

Shematski prikaz biometričnega sistema, katerega funkcija je verifikacija, je prikazan na spodnji sliki.



Slika 2.1: Biometrični sistem, ki opravlja verifikacijo osebe. [1]

V podatkovni zbirki sistema so shranjene značilke/deskriptorji oz. vzorci vseh oseb, ki imajo dostop do nekega varovanega sistema. Zajeti in shranjeni so bili v fazi registracije v sistem. Oseba, ki želi vstopiti v varovan sistem, mora opraviti verifikacijo oz. potrditi svojo istovetnost. Najprej se predstavi z imenom ali PIN kodo ali na kakšen drug način (kartico, geslo,...). Nato se zahtevana identiteta poišče v podatkovni zbirki in pripravi shranjen vzorec. Biometrično tipalo zajame trenutne podatke, potem algoritem za razpoznavanje vzorcev iz zajetega vzorca izlušči značilke in jih zapiše v deskriptor. Dobljen deskriptor se primerja s shranjenim in algoritem za izračun podobnosti odloči, ali sta vzorca dovolj podobna za potrditev istovetnosti prijavljajoče osebe. Če je podobnost manjša od določenega praga, verifikacija ni uspela in osebi je onemogočen vstop v varovan sistem.

Za izračun in zapis značilnih vektorjev smo uporabili štiri različne postopke, in sicer projekcijo v podprostor (PCA) ter tri lokalne deskriptorje (SURF, FREAK, BRISK). Vektorju, v katerega zapišemo značilnosti, pravimo tudi deskriptor. SURF, FREAK in BRISK algoritem so računali značilnosti okoli slikovnih točk, ki smo jim sami določili koordinate v obliki mreže. Pri postopku projekcije v podprostor pa gre za preslikavo vzorcev v manj razsežen prostor. Ne iščejo se značilne točke tako kot pri lokalnih deskriptorjih, ampak sam transformiran vzorec predstavlja značilke originalnega vzorca. Za izračun razdalj med značilnimi vektorji smo uporabili Evklidovo, kosinusno in City Block razdaljo. Rezultate razpoznavanja smo prikazali v obliki ROC krivulj in tabel.

2.1 Izračun značilk z analizo glavnih komponent (PCA)

Analizi glavnih komponent (angl. Principal Component Analysis) pravimo tudi transformacija Karhunen in Loéveja. Pri računalniškem razpoznavanju obrazov, kjer so slike zapisane v obliki matrik oz. vektorjev z več tisoč vrsticami/stolpci, je za hitrost razpoznavanja dobro, da zmanjšamo dimenzijo vektorjev. PCA postopek omogoča preslikavo v podprostor z manj koordinatnimi osmi. Ker se pri taki transformaciji lahko izgubi nekaj informacije o vzorcu, lahko PCA definiramo kot postopek minimizacije srednje kvadratne napake približnega zapisa vzorcev.

Za vsako polno množico ortonormiranih vektorjev $\{e_j\}$ lahko zapišemo vzorec $\mathbf{y} = (y_1, y_2, \dots, y_r)^T$ kot [5, stran 133-136]:

$$\mathbf{y} = \sum_{j=1}^r c_j \mathbf{e}_j, \quad (2.1)$$

kjer so c_j koeficienti razvrstitve.

Vzemimo sedaj, da želimo približno zapisati vzorec $\mathbf{y}$ tako, da nadomestimo koeficiente razvrstitve $(c_{n+1}, c_{n+2}, \dots, c_r)$ s konstantami b_j . Približni vzorec $\hat{\mathbf{y}}$ zapišemo kot:

$$\hat{\mathbf{y}} = \sum_{j=1}^n c_j \mathbf{e}_j + \sum_{j=n+1}^r b_j \mathbf{e}_j. \quad (2.2)$$

Pri tem smo zagrešili napako približnega zapisa vzorca Δ_y :

$$\Delta_y = \mathbf{y} - \hat{\mathbf{y}} = \sum_{j=n+1}^r (c_j - b_j) \mathbf{e}_j \quad (2.3)$$

oziroma srednjo kvadratno napako [5]:

$$\bar{\Delta}_y^2 = E \left\{ \|\Delta_y\|^2 \right\} = E \{ \Delta_y^T \Delta_y \}, \quad (2.4)$$

kjer je E operator matematičnega upanja.

Če vstavimo enačbo (2.3) v (2.4) in ob upoštevanju, da so vektorji $\mathbf{e}_j$ deterministični in ortonormirani:

$$\mathbf{e}_j^T \mathbf{e}_k = \begin{cases} 1, & \text{če } j = k \\ 0, & \text{če } j \neq k \end{cases} \quad (2.5)$$

lahko srednjo kvadratno napako zapišemo kot:

$$\bar{\Delta}_y^2 = \sum_{j=n+1}^r E \left\{ (c_j - b_j)^2 \right\}. \quad (2.6)$$

Poiščemo takšne konstante b_j in takšne ortonormirane vektorje $\mathbf{e}_j$, da bo srednja kvadratna napaka najmanjša. Enačbo (2.6) odvajamo po b_j in odvod enačimo z 0. Dobimo, da je to pri $b_j = E\{c_j\}$, oziroma, ker je $c_j = \mathbf{e}_j^T \mathbf{y}$, velja:

$$b_j = \mathbf{e}_j^T E\{\mathbf{y}\}. \quad (2.7)$$

Zdaj poiščemo še ortonormirane bazne vektorje $\mathbf{e}_j$ (koordinatne osi). Ker lahko zapišemo $c_j = \mathbf{y}^T \mathbf{e}_j$, je srednja kvadratna napaka [5]:

$$\begin{aligned} \bar{\Delta}_y^2 &= \sum_{j=n+1}^r E \left\{ (c_j - E\{c_j\})^2 \right\} = \sum_{j=n+1}^r \mathbf{e}_j^T E\{(\mathbf{y} - E\{\mathbf{y}\})(\mathbf{y} - E\{\mathbf{y}\})^T\} \mathbf{e}_j \\ &= \sum_{j=n+1}^r \mathbf{e}_j^T \mathbf{K} \mathbf{e}_j, \end{aligned} \quad (2.8)$$

kjer je $\mathbf{K}$ kovariančna matrika vzorcev $\mathbf{y}$.

Minimizirano srednjo kvadratno napako dobimo tako, da Lagrangejevo funkcijo [5]:

$$\mathcal{L}_{\bar{\Delta}_y^2} = \sum_{j=n+1}^r \mathbf{e}_j^T \mathbf{K} \mathbf{e}_j + \sum_{j=n+1}^r \lambda_j (1 - \mathbf{e}_j^T \mathbf{e}_j) , \quad (2.9)$$

kjer so λ_j Lagrangejevi multiplikatorji, odvajamo po baznih vektorjih in odvede izenačimo z 0. Po odvajanju vidimo, da najmanjšo vrednost Lagrangejeve funkcije dobimo z vektorji $\mathbf{e}_j$, ki zadoščajo enačbi:

$$\mathbf{K} \mathbf{e}_j = \lambda_j \mathbf{e}_j . \quad (2.10)$$

Če enačbo (2.10) vstavimo v (2.8) in upoštevamo ortonormiranost vektorjev $\mathbf{e}_j$, dobimo za srednjo kvadratno napako:

$$\bar{\Delta}_y^2 = \sum_{j=n+1}^r \mathbf{e}_j^T \mathbf{K} \mathbf{e}_j = \sum_{j=n+1}^r \mathbf{e}_j^T \lambda_j \mathbf{e}_j = \sum_{j=n+1}^r \lambda_j . \quad (2.11)$$

Opazimo, da zagrešimo najmanjšo srednjo kvadratno napako približnega zapisa vzorca z n koeficienti razvrstitve, če ga zapišemo s koeficienti tistih lastnih vektorjev $\mathbf{e}_j$, ki ustrezajo največjim lastnim vrednostim kovariančne matrike $\mathbf{K}$, preostale koeficiente pa zamenjamo s konstantami b_j , $j = n+1, \dots, r$. Te preostale koeficiente lahko zamenjamo z 0, če je $E\{\mathbf{y}\}=0$. Koeficiente PCA razvrstitve c_j , $j = 1, 2, \dots, n$ ($n \leq r$), zato določimo s transformacijo [5]:

$$\mathbf{c} = \mathbf{A}(\mathbf{y} - E\{\mathbf{y}\}) , \quad (2.12)$$

transformacijsko matriko $\mathbf{A}$ pa sestavimo iz začetnih n vektorjev koordinatnih osi $\mathbf{e}_j$, $j = 1, 2, \dots, n$ ($n \leq r$), ki so urejeni po padajočem vrstnem redu lastnih vrednosti matrike $\mathbf{K}$:

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq \dots \geq \lambda_r . \quad (2.13)$$

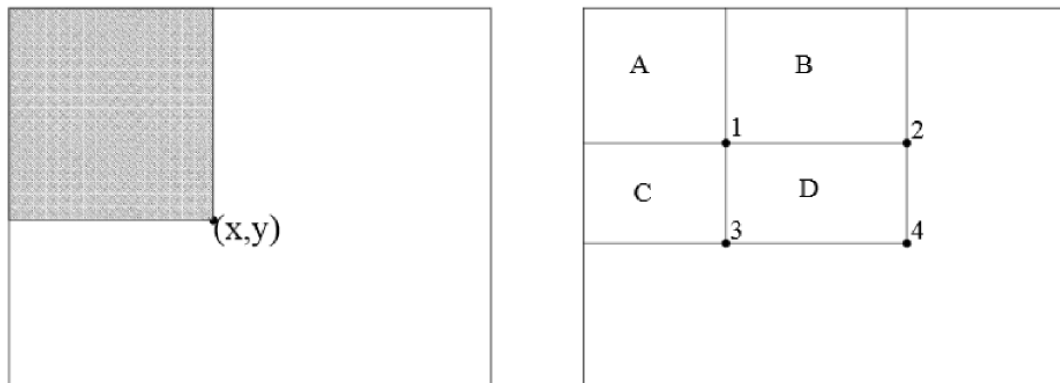
Večja kot je lastna vrednost, več informacije o vzorcu zajame pripadajoči vektor $\mathbf{e}_j$. Tako dobljeni transformirani vzorec $\mathbf{c}$ ima komponente c_j nekorelirane, torej linearno neodvisne. Hkrati vzorec $\mathbf{c}$ predstavlja značilke originalnega vzorca $\mathbf{y}$ (slike) in v tem podprostoru poteka tudi primerjava podobnosti med transformiranimi vzorci za potrebe verifikacije.

2.2 SURF deskriptor

SURF (Speeded Up Robust Feature) [7] deskriptor je invarianten na spremembo velikosti in rotacijo objektov razpoznavanja. Osnovan je na bolj poznanem SIFT (Scale Invariant Feature Transform) [10] deskriptorju, vendar so postopki in principi še bolj poenostavljeni, zato doseže večjo hitrost izračuna ob podobni učinkovitosti [20]. Glavni vzrok za hitrejši izračun je v uporabi integralnih slik (angl. Integral Image [6]). Vhodne slike, katerih značilne točke iščemo, pretvorimo v integralne tako, da jih zapišemo kot:

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y'), \quad (2.14)$$

kjer je $ii(x, y)$ integralna slika in $i(x, y)$ originalna slika. Vrednost integralne slike v točki (x, y) je torej vsota sivinskih nivojev vseh slikovnih točk originalne slike, ki so levo od točke (x, y) in nad njo (Slika 2.2 levo).



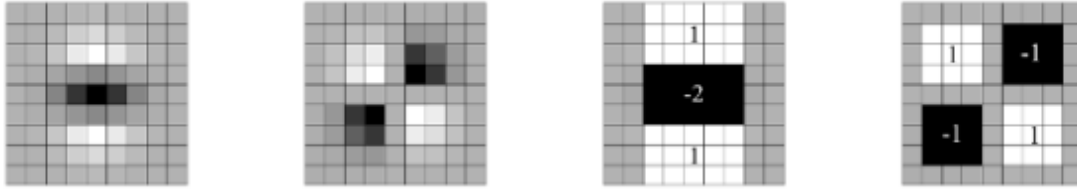
Slika 2.2: Izračun integralne slike (vrednost integralne slike v točki 1 je vsota vseh slikovnih točk pravokotnika A, v točki 2 je A+B, v točki 3 je A+C, v točki 4 je A+B+C+D). [6]

Vsako pravokotno območje na sliki, ne glede na velikost, lahko izračunamo s samo tremi operacijami. Območje D na Sliki 2.2. desno, se izračuna kot $4+1-(2+3)$. Uporaba integralnih slik naredi SURF algoritem tudi do trikrat hitrejšega od SIFT, vendar je manj robusten na osvetlitev in spremembo oz. zasuk obraza pred kamero.

Integralno sliko nato filtriramo s približkom drugega odvoda Gaussovega filtra (Slika 2.3) pri različnih vrednostih parametra σ_i . Gaussova funkcija je določena kot:

$$G_i(x, y, \sigma_i) = \frac{1}{2\pi\sigma_i^2} e^{-(x^2+y^2)/2\sigma_i^2}. \quad (2.15)$$

Dobimo prostor meril (angl. scale-space), kjer so slike zaporedno nanizane v oktave po naraščajoči zamegljenosti.



Slika 2.3: Gaussov filter (levi dve slike prikazujeta diskretiziran in obrezan drugi parcialni odvod Gaussove funkcije v y in xy smeri; desni slike prikazujeta njuna približka v obliki pravokotnega filtra). [7]

Za detekcijo značilnih točk SURF uporablja Hessian matriko. Determinanta Hessian matrike je merilo za lokalne spremembe okoli slikovne točke in če ima točka maksimalno determinanto je prepoznana kot značilna točka. Za dano točko $\mathbf{x} = (x, y)$ slike I v prostoru meril (velikosti) σ , definiramo Hessian matriko [7]:

$$H(\mathbf{x}, \sigma) = \begin{bmatrix} L_{xx}(\mathbf{x}, \sigma) & L_{xy}(\mathbf{x}, \sigma) \\ L_{xy}(\mathbf{x}, \sigma) & L_{yy}(\mathbf{x}, \sigma) \end{bmatrix}, \quad (2.16)$$

kjer je $L_{xx}(\mathbf{x}, \sigma)$ konvolucija drugega odvoda Gaussove funkcije s sivinsko sliko I v točki $\mathbf{x}$. Podobno velja za $L_{xy}(\mathbf{x}, \sigma)$ in $L_{yy}(\mathbf{x}, \sigma)$.

Izračunamo Hessian matriko pri različnih vrednostih σ , ki predstavlja nivo zglajenosti slike. Začetni nivo prostora meril predstavlja vrednost $\sigma = 1.2$. Za ta nivo se uporablja filter velikosti 9×9 (Slika 2.3 desno). Za vsak naslednji nivo, ki je vse bolj zamegljen (zglajen z večjim σ), se uporablja postopoma večji filter (15×15 , 21×21 , 27×27 itd.), velikost slike pa ostane enaka. Za lokalizacijo značilne točke v sliki se izvede dušenje ne-maksimalnih točk v $3 \times 3 \times 3$ okolici. Maksimumi determinante Hessian matrike so potem interpolirani v prostoru meril in slikovnem prostoru.

Namen deskriptorja je unikaten in robusten zapis značilnosti slike z opisovanjem lokalne okolice najdene značilne točke. SURF deskriptor je sestavljen iz dveh delov.

Najprej se določi orientacijo značilne točke. Izračuna se odziv na Haarove funkcije v x in y smeri (Slika 2.4 levo) v krožni okolici z radijem $6s$ (s predstavlja nivo prostora meril, v katerem je bila točka zaznana) okoli značilne točke. Dobljeni odzivi (obteženi z Gaussovo funkcijo) so potem narisani kot točke v xy ravnini, kjer so odzivi v vodoravni smeri na abscisi, odzivi v navpični smeri pa na ordinatni osi. Dominantna orientacija je ocenjena z računanjem vsote vseh odzivov znotraj okna

velikosti $\pi/3$, ki naredi krog okoli koordinatnega središča. Smer tistega okna, kjer je vektor vsote odzivov najdaljši, je določena kot skupna (dominantna) orientacija značilne točke.

Nato na mestu značilne točke določimo kvadratno območje, ki je orientirano v smer, ki smo jo dobili v prejšnjem koraku. Velikost območja je $20s$. Območje je razdeljeno na 4×4 podobmočja. Za vsako podobmočje se izračunajo odzivi na Haarovo funkcijo v horizontalno smer (d_x) in vertikalno smer (d_y), relativno na orientacijo značilne točke. Uporabimo Haarovo funkcijo velikosti 5×5 slikovnih točk. Odzive vsakega podobmočja d_x in d_y seštejemo, prav tako seštejemo tudi njihove absolutne vrednosti, zato da dobimo informacijo o polarnosti sprememb intenzitet. Rezultat vsot podobmočja zapišemo v štiri dimenzionalen vektor $\mathbf{v}$ [7]:

$$\mathbf{v} = (\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y|). \quad (2.17)$$

Ker je podobmočij 16, dobimo 64 dimenzionalen deskriptor značilne točke.



Slika 2.4: Levo: Haarovo okno v x in y smer. Desno: Prikaz SURF deskriptorja na sliki (vidimo orientacijo in velikost kvadrata, ki je odvisna od nivoja prostora meril, v katerem je bila značilna točka najdena). [7]

V primerih, ko je zajem slike oz. kamera vedno v istem, horizontalnem položaju, je izračun orientacije deskriptorja nepotreben. Takrat lahko uporabimo različico deskriptorja, ki mu pravimo U-SURF. Ta ni invarianten na rotacijo, zato je izračun še hitrejši.

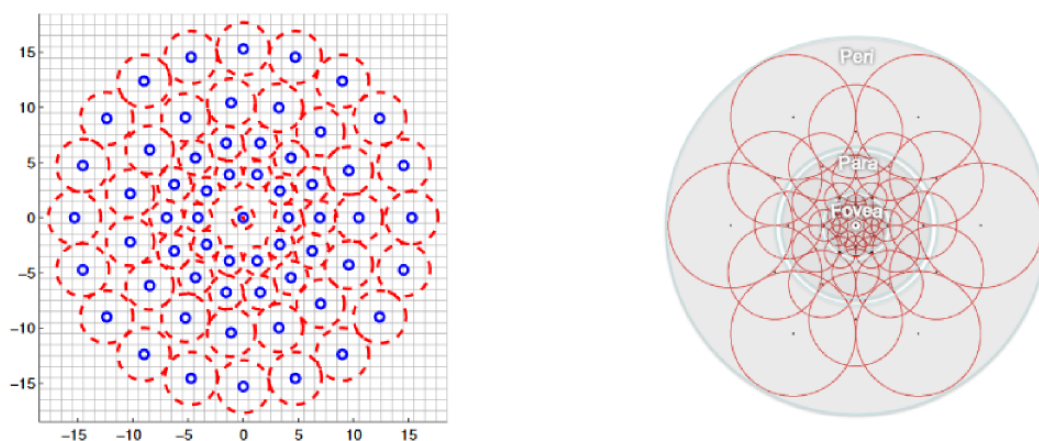
2.3 Binarni deskriptorji

Deskriptorji omenjeni v prejšnjem poglavju (SIFT [10], SURF [7]) temeljijo na računanju gradientov (odvodov v x in y smer) vsake slikovne točke. Kljub temu, da SURF pohitri delovanje z uporabo integralnih slik, so izračuni še vedno lahko prepočasni za določene aplikacije razpoznavanja. Poleg tega sta oba algoritma patentirana, kar oteži prosto uporabo. Ena izmed prednosti binarnih deskriptorjev je, da večinoma niso patentirani.

Ideja binarnih deskriptorjev je, da področja na sliki opišemo kot binarni niz s primerjavo intenzivnosti oz. sivinskih nivojev slikovnih točk. Taki izračuni so hitri in relativno preprosti [21]. V diplomskem delu smo raziskovali učinkovitost binarnih deskriptorjev, imenovanih BRISK (angl. Binary Robust Invariant Scalable Keypoints) [8] in FREAK (angl. Fast Retina Keypoint) [9].

Za detekcijo značilnih točk v prostoru meril oba uporabljata izboljššan FAST detektor [11], imenovan AGAST [12]. Temelji na iskanju kotov v sliki, ki jih najde tako, da primerja svetilnost šestnajstih slikovnih točk razporejenih v krog, s svetilnostjo točke v centru kroga. Če je vsaj devet zaporednih točk na krogu svetlejših (ali temnejših) od centralne slikovne točke, je ta prepoznana kot značilna točka, ki predstavlja nek kot na sceni slike.

V splošnem so binarni deskriptorji sestavljeni iz treh delov, in sicer vzorca za vzorčenje, kompenzacije orientacije in vzorčnih parov.



Slika 2.5: Levo: vzorec za vzorčenje pri BRISK deskriptorju [8]. Desno: vzorec za vzorčenje pri FREAK deskriptorju [9]. Pri obeh vzorcih različno veliki rdeči krogi pomenijo različne standardne deviacije Gaussovega filtra, s katerim gladimo slikovne točke.

2.3.1 BRISK deskriptor

V prvem koraku se poišče značilne točke z algoritmom za iskanje kotov na sliki. Po detekciji značilnih točk sledi vzorčenje z vzorcem, sestavljenim iz 60 različno velikih krogov, ki je prikazan na Sliki 2.5 levo.

Za določitev orientacije značilne točke, BRISK računa lokalne gradiente parov slikovnih točk z veliko medsebojno razdaljo znotraj vzorca iz Slike 2.5. Lokalni gradient med parom slikovnih točk ($\mathbf{p}_i, \mathbf{p}_j$) je:

$$g(\mathbf{p}_i, \mathbf{p}_j) = (\mathbf{p}_j - \mathbf{p}_i) \cdot \frac{I(\mathbf{p}_j, \sigma_j) - I(\mathbf{p}_i, \sigma_i)}{\|\mathbf{p}_j - \mathbf{p}_i\|^2}, \quad (2.18)$$

kjer je I svetilnost posamezne točke, zglajene z Gaussovim filtrom pripadajoče standardne deviacije. Nato se izračuna vsota vseh lokalnih gradientov dolgih parov točk in uporabi funkcija $\alpha = \arctan2(g_y, g_x)$. Dobimo kot orientacije značilne točke. Za ta kot moramo rotirati vzorec za vzorčenje, ko bomo opisovali okolico značilne točke s kratkimi pari točk. S tem je deskriptor postal neobčutljiv (invarianten) na rotacijo.

BRISK sestavi deskriptor s primerjavo svetilnosti vzorčnih parov s kratko medsebojno razdaljo po naslednji dvojiški logiki:

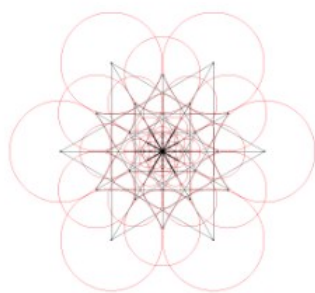
$$b = \begin{cases} 1, & I(\mathbf{p}_j^\alpha, \sigma_j) > I(\mathbf{p}_i^\alpha, \sigma_i) \\ 0, & \text{drugače} \end{cases}. \quad (2.19)$$

To pomeni, da za vsak kratek par, kjer je zglajena svetilnost prve točke v paru večja od druge, zapiše enico. Ko je obratno, pa zapiše ničlo. Če izberemo 512 kratkih parov, dobimo deskriptor, sestavljen iz niza enic in ničel, ki je dolg 512 znakov ter opisuje lokalno okolico značilne točke.

2.3.2 FREAK deskriptor

FREAK ima nekaj podobnosti z BRISK deskriptorjem. Najprej se poišče značilne točke z algoritmom za iskanje kotov na sliki. Prav tako uporablja vnaprej določen vzorec za vzorčenje (Slika 2.5 desno) ter ima podoben mehanizem za določitev orientacije. Vzorec temelji na posnemanju gostote porazdelitve ganglion celic v očesni mrežnici človeka. Gostota točk je največja v središču vzorca in eksponentno pada z razdaljo od središča.

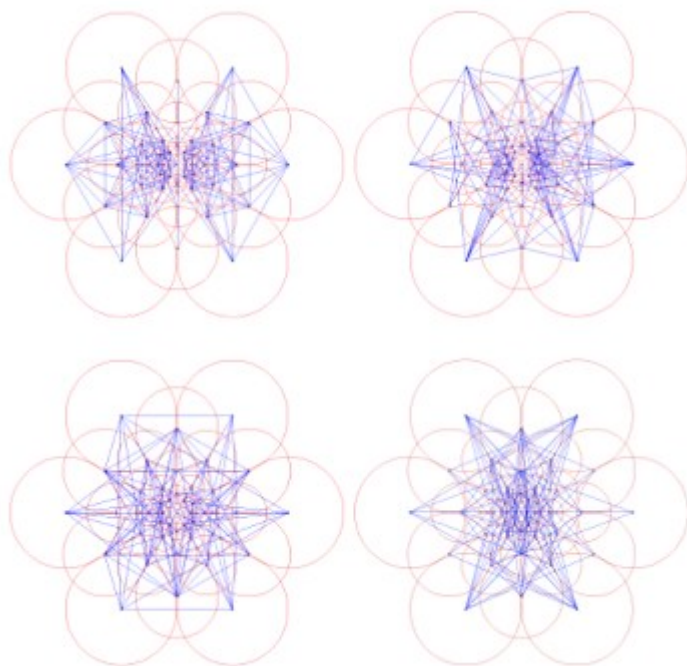
Orientacijo značilne točke se izračuna iz 45 parov točk, ki so določeni.



Slika 2.6: Prikaz 45 določenih parov za izračun orientacije FREAK deskriptorja. [9]

Lokalni gradienti 45 parov se izračunajo po enačbi (2.18), nato se jih sešteje in deli s številom parov. Tako je določena smer značilne točke.

Pri BRISK-u so vzorčni pari za določitev deskriptorja določeni na podlagi kratke prostorske razdalje med točkami. Lahko se zgodi, da so taki pari korelirani in niso diskriminantni, kar ni zaželeno. Zato FREAK, z maksimiziranjem variance parov, poizkusi poiskati najboljše nekorelirane pare iz učne množice. Izkaže se, da obstaja struktura najdenih parov, kot je prikazana na Sliki 2.7. Vzorčenje točk se začne z rojem levo zgoraj, ki vzorči večinoma na obrobni poljih, konča pa se z rojem desno spodaj, ki vzorči središčna polja. Avtorji deskriptorja so eksperimentalno opazili, da je prvih 512 parov najpomembnejših in zato dodajanje parov ne izboljša izvedbe.



Slika 2.7: Vzorčni pari v štirih rojih (128 parov na roj, $4 \cdot 128 = 512$). [9]

Vzorčenje sledi našemu razumevanju delovanja očesne mrežnice pri zaznavanju okolja. Najprej uporabimo obrobna receptorska polja za oceno lokacije objekta, nato poteka validacija z bolj na gosto porazdeljenimi receptorskimi polji v središču mrežice.

Vzorčne pare zavrtimo za kot, dobljen pri oceni orientacije značilne točke in sestavimo deskriptor po enačbi (2.19).

2.4 Prileganje vzorcev in razdalje

Za prileganje vzorcev je nujno, da so le-ti ter njihovi deskriptorji enakih dimenzij. Pri biometričnem sistemu, ki opravlja verifikacijo, se značilke (deskriptorji) prijavljajoče osebe primerjajo z značilkami osebe iz podatkovne zbirke, za katero se prijavitelj izdaja.

Podobnost med deskriptorji, ki so predstavljeni kot vektorji, lahko ugotovljamo na različne načine z računanjem razdalj med njihovimi elementi. Manjša kot je razdalja med njimi, večja je podobnost (princip najbližjega sosedu). V pričujočem zaključnem delu smo uporabili Evklidovo, "City Block" in kosinusno razdaljo za primerjavo učinkovitosti razpoznavanja [5].

Evklidova in "City Block" razdalja sta posebni vrsti razdalje Minkovskega:

$$D(\mathbf{x}_k, \mathbf{x}_l) = \left(\sum_{j=1}^n |x_{k,j} - x_{l,j}|^s \right)^{\frac{1}{s}}, \quad (2.20)$$

kjer je n število značilk vzorca $\mathbf{x}_k$ oz. $\mathbf{x}_l$. Za $s=2$ izračunamo Evklidovo razdaljo po enačbi:

$$D_E(\mathbf{x}_k, \mathbf{x}_l) = \sqrt{\sum_{j=1}^n (x_{k,j} - x_{l,j})^2} = \sqrt{(\mathbf{x}_k - \mathbf{x}_l)^T (\mathbf{x}_k - \mathbf{x}_l)} = \|\mathbf{x}_k - \mathbf{x}_l\|, \quad (2.21)$$

za $s=1$ izračunamo "City Block" ali Manhattan razdaljo [5]:

$$D_{CB}(\mathbf{x}_k, \mathbf{x}_l) = \sum_{j=1}^n |x_{k,j} - x_{l,j}|. \quad (2.22)$$

Kosinusna razdalja med dvema vektorjema je:

$$D_{cos}(\mathbf{x}_k, \mathbf{x}_l) = 1 - \frac{\sum_{j=1}^n x_{k,j} x_{l,j}}{\sqrt{\sum_{j=1}^n x_{k,j}^2} \sqrt{\sum_{j=1}^n x_{l,j}^2}} . \quad (2.23)$$

Princip verifikacije lahko formalno opredelimo kot [1]:

$$(\mathbf{x}_k, \mathbf{x}_l) \in \begin{cases} \text{ista oseba, če } D(\mathbf{x}_k, \mathbf{x}_l) \leq t \\ \text{različni osebi, drugače} \end{cases} . \quad (2.24)$$

Če je razdalja med vzorcema $\mathbf{x}_k$ in $\mathbf{x}_l$ manjša ali enaka pragu, potem vzorca pripadata isti osebi. Če je prag presežen, vzorca pripadata različnim osebam.

2.5 Napake in ROC krivulje

Biometrični sistem verifikacije lahko naredi dve napaki razpoznavanja [1]. Prijavljajočo osebo (vsiljivca) lahko prepozna kot uporabnika sistema, čeprav to ni (angl. false accept). Ali pa prijavljajočo osebo (klienta) zavrne, čeprav je registrirana v sistemu (angl. false reject). Pri večjem številu prijav napaki predstavimo kot razmerji FAR (angl. false accept rate) in FRR (angl. false reject rate). Obe napaki sta funkciji praga sistema in potrebno je izbrati kompromis med obema napakama, odvisno od aplikacije biometričnega sistema. Če prag povečamo, da naredimo sistem bolj toleranten na vhodne variacije in šum, potem se tudi FAR poveča. Na drugi strani, če prag zmanjšamo, da naredimo sistem bolj zanesljiv, potem se FRR poveča (in FAR zmanjša).

Napaki se izračunata po enačbah [16]:

$$FAR = \frac{N_V}{I_V} \cdot 100\% , \quad (2.25)$$

$$FRR = \frac{N_K}{I_K} \cdot 100\% , \quad (2.26)$$

kjer so:

N_V - število sprejetih vsiljivcev v sistem,

I_V - število vseh prijav vsiljivcev v sistem,

N_K - število zavrnjenih klientov v sistem,

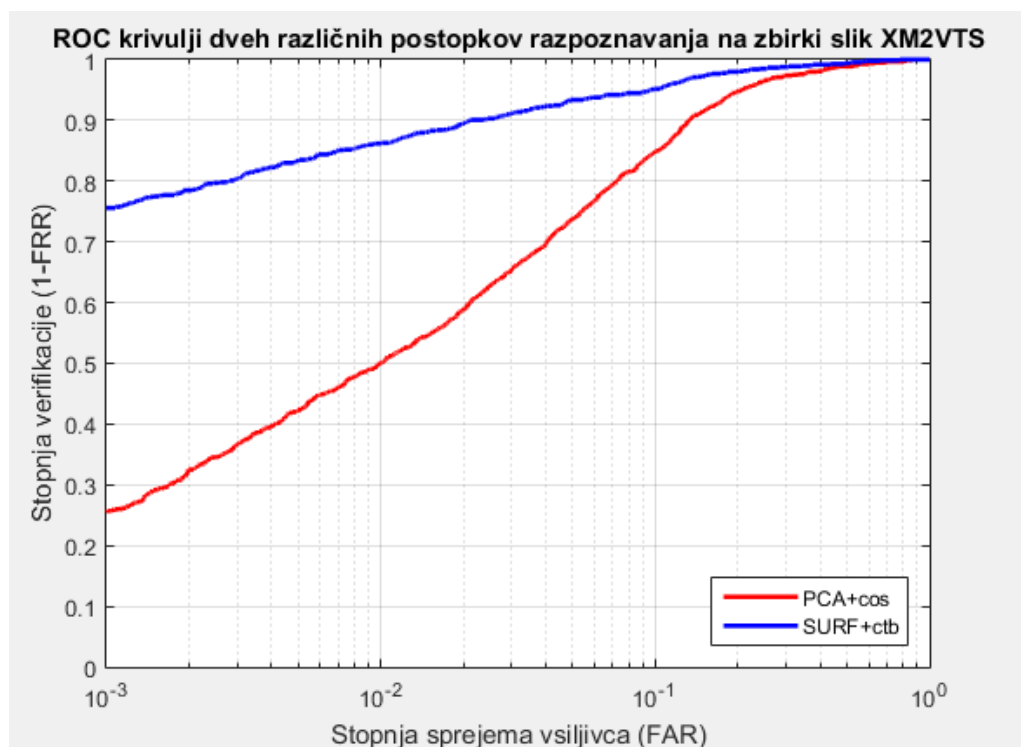
I_K - število vseh prijav klientov v sistem.

Za vrednotenje verifikacije se uporablja tudi HTER (angl. Half Total Error Rate) napaka. Določa srednjo vrednost med zgoraj omenjenima napakama:

$$HTER = 0.5(FAR + FRR) . \quad (2.27)$$

Vrednosti napake HTER v delovni točki ROC krivulje, ko je $FAR=FRR$, pravimo tudi EER (angl. Equal Error Rate). Takrat je $HTER=FAR=FRR$.

Delovanje sistema v različnih delovnih točkah oz. pragovih lahko prikažemo z ROC krivuljo (angl. Receiver Operating Characteristic) [1], ki podaja odvisnost med FAR in FRR (ali $1 - FRR$). Rezultate testiranj razpoznavanja obrazov na podanih podatkovnih zbirkah smo prikazali na ROC krivulji, ki ima na abscisi vrednosti FAR, na ordinatni osi pa vrednosti $1-FRR$. Na Sliki 2.8 je primer krivulj pri razpoznavanju obrazov iz podatkovne zbirke XM2VTS s PCA postopkom (uporabljena kosinusna razdalja) in SURF deskriptorjem (uporabljena City Block razdalja). Abscisna os je v logaritemskem merilu. Vrednosti na oseh pomnožimo za faktor sto, da dobimo stopnje v odstotkih. Tako so tudi prikazane v tabeli rezultatov v poglavju 4.



Slika 2.8: Primer dveh ROC krivulj uporabljenih za prikaz rezultatov verifikacije (podajata vrednosti FAR in stopnje verifikacije ob različnih pragovih). Vidimo, da SURF deskriptor dosega boljše rezultate verifikacije kot PCA postopek.

3 METODOLOGIJA DELA

3.1 Oprema in orodja

Delo je potekalo na osebem računalniku Acer Aspire V3-772G z naslednjimi specifikacijami:

- procesor Intel Core i7-4702MQ CPU 2.2GHz,
- grafika NVIDIA GeForce GTX 850M,
- spomin RAM 12 GB DDR3,
- operacijski sistem 64-bit Windows 8.1.

Sistem razpoznavanja je bil izveden v okolju Matlab R2015a. Za izvedbo protokolov smo uporabili osnovne matlabove funkcije, funkcije matlabove zbirke orodij za strojni vid (Computer Vision System Toolbox) ter funkcije zbirke orodij PhD Toolbox [4]. Zadnjo je razvil mentor dr. Vitomir Štruc in vsebuje implementacije različnih znanih tehnik za razpoznavanje obrazov. S pomočjo matlabovih funkcij in skript omogoča izgradnjo ter analizo sistemov za razpoznavanje.

3.2 Podatkovne zbirke

Analizo in primerjavo različnih deskriptorjev smo opravili na sivinskih slikah bmp formata iz dveh različnih podatkovnih zbirk, in sicer XM2VTS ter Extended Yale B. Za vsako zbirko slik smo napisali svoj protokol, ki je predstavljen v naslednjem poglavju.

Slike obeh zbirk so velikosti 100×100 slikovnih točk, kjer vsaka točka predstavlja nivo sivine v vrednosti od 0 (črna) do 255 (bela). Slike v Matlabu zapišemo kot vektorje z 10000 elementi, in sicer tako, da nanizamo vse vrstice slike eno za drugo.

3.2.1 Zbirka XM2VTS

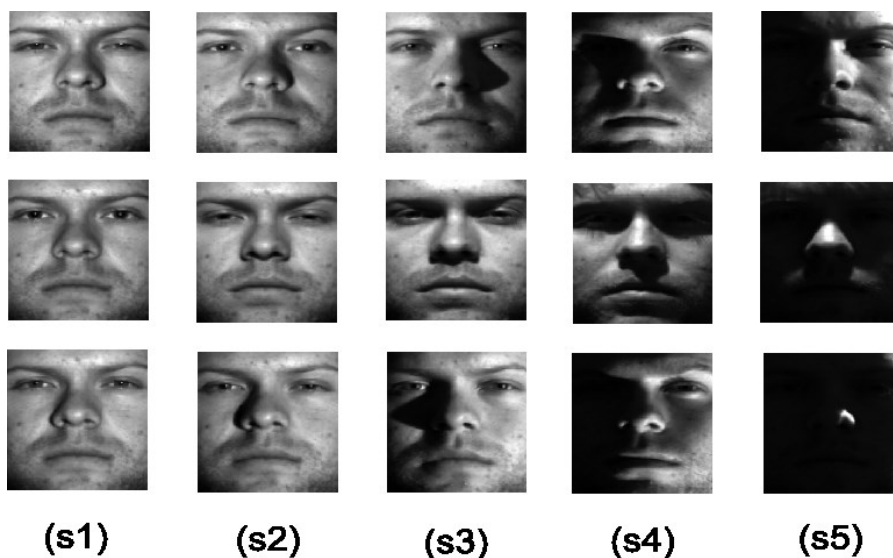
Zbirka obrazov XM2VTS (angl. Extended Multi Modal Verification for Teleservices and Security applications) [2] je bila pridobljena kot del projekta Evropske Unije, ki se ukvarja z nadzorom dostopa s pomočjo identifikacije obrazov ljudi. Zbirka je sestavljena iz sivinskih slik obrazov 295 ljudi, ki so jih fotografirali v obdobju štirih mesecev. Skupno število slik je 2360, saj je vsaka oseba bila slikana osemkrat. Zajem slik je bil realiziran s kamero ob nespremenjenih nastavitvah ter ob isti poziciji in svetilnosti svetlobnih teles. Zbirka je torej dokaj "čista", brez večjih variabilnosti v osvetljenosti obrazov. Izbrali smo jo zato, da analiziramo razpoznavanje z deskriptorji v najboljšem primeru, ko so spremembe v osvetlitvi čim manjše.



Slika 3.1: Primer osmih zajemov obraza enega človeka iz zbirke XM2VTS. Slike v naši zbirki so že obrezane. [2]

3.2.2 Zbirka Extended Yale B

Zbirka obrazov "Extended Yale B" [3] je še ena izmed podatkovnih zbirk, ki se pogosto uporabljajo za izvedbo eksperimentov različnih algoritmov razpoznavanja obrazov. Je razširjena verzija zbirke pridobljene na ameriški univerzi Yale. Izbrali smo jo zato, da raziščemo delovanje deskriptorjev na slikah z veliko variabilnostjo v osvetlitvi. Naša množica vsebuje 2414 sivinskih slik obrazov 38 oseb. Razdeljena je na pet podmnožic (angl. subset), kjer so zajete obrezane slike z različnimi stopnjami osvetlitve. Podmnožica s1 vsebuje 263 slik, s2 jih vsebuje 456, s3 vsebuje 455, s4 vsebuje 526 ter s5 jih vsebuje 714. Podmnožice si sledijo po naraščajoči zatemnitvi obrazov, kakor je prikazano na sliki 3.2. V s5 je osvetlitev najmanjša, zato je razpoznavanje tam daleč najmanj učinkovito. Večina izmed 38 oseb je bila slikana pod 64 različnimi osvetlitvenimi pogoji, nekaj oseb pa pod nekoliko manj, zato je v naši zbirki skupaj 2414 slik.



Slika 3.2: Prikaz nivoja zatemnitve obrazov v posameznih podmnožicah. [3]

3.3 Vrednotenje sistema razpoznavanja

Uspešnost sistema razpoznavanja smo vrednotili z enako stopnjo napake (EER), minimalno polovično skupno stopnjo napake (minHTER) in stopnjo verifikacije v treh različnih delovnih točkah, pri vrednostih FAR=1%, FAR=0.1% ter FAR=0.01%. V ostalih delovnih točkah so vrednosti FAR in FRR določljive iz ROC krivulje, ker je stopnja verifikacije definirana kot $1 - \text{FRR}$. Napaka minHTER določa najmanjšo napako HTER med vsemi delovnimi točkami.

4 EKSPERIMENTI

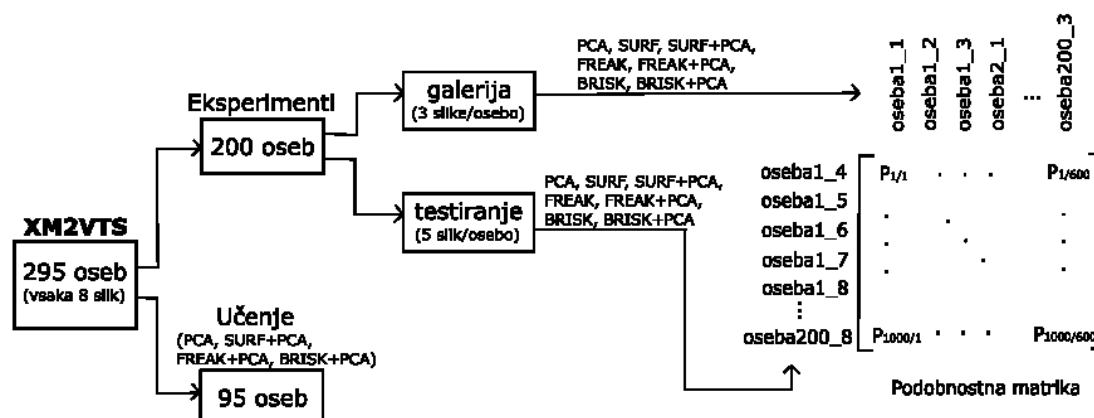
Rezultati testiranja razpoznavanja in diskusija so predstavljeni za vsako podatkovno zbirko obrazov posebej.

4.1 Eksperimenti na zbirki obrazov "XM2VTS"

Na omenjeni zbirki, kjer so slike brez velikih sprememb v osvetljenosti, smo raziskovali in primerjali razpoznavanje s PCA, SURF, SURF+PCA, FREAK, FREAK+PCA, BRISK ter BRISK+PCA postopkom. Pri vseh postopkih smo računali razdaljo med značilnimi vektorji na tri načine, in sicer z Evklidovo, City Block ter kosinusno razdaljo. Uspešnost razpoznavanja smo vrednotili tako, kot je opisano v razdelku 3.3.

4.1.1 Protokol

Bločni diagram protokola je prikazan na Sliki 4.1. V zbirki XM2VTS so slike 295 oseb. Slike 95 oseb uporabimo za izgradnjo manj razsežnega podprostora pri PCA postopku in pri kombiniranih SURF+PCA, FREAK+PCA ter BRISK+PCA postopkih. Slike preostalih 200 oseb (1600 slik) uporabimo za eksperimentiranje na sistemu razpoznavanja obrazov pri vseh sedmih postopkih, in sicer tako, da prve tri slike vsake osebe predstavljajo vzorce v galeriji (oz. podatkovni zbirki sistema razpoznavanja, glej Sliko 2.1), naslednjih pet slik pa predstavlja testne slike, ki so vhod v sistem. V podobnostno matriko se zapišejo rezultati računanja razdalj med vektorji značilk oz. deskriptorji vseh slik iz galerije in deskriptorji vseh testnih slik. V našem primeru se vseh pet deskriptorjev pet različnih slik vsake osebe iz testne množice primerja z vsemi tremi deskriptorji treh različnih slik vsake osebe iz galerije (glej Sliko 4.1). Matrika podobnosti je zato velikosti 1000×600 in opravi 600.000 primerjav. Ko določena oseba predstavlja klienta, ostalih 199 oseb prevzame vlogo vsiljivca.



Slika 4.1: Shema protokola do izračuna podobnostne matrike za zbirko XM2VTS.

PCA:

Najprej uvozimo slike prvih 95 oseb iz zbirke obrazov, ki jih bomo uporabili kot učne slike za izgradnjo PCA podprostora. Skupaj uporabimo 760 slik za učenje. Matrika učnih slik je dimenzij 10000×760 . Rang matrike učnih slik, zmanjšan za ena, določimo kot želeno velikost podprostora. Dobimo PCA podprostor, ki ima 759 lastnih vektorjev. Za izračun vektorjev PCA značilk preslikamo galerijske in testne slike s pomočjo funkcije PhD Toolboxa v PCA podprostor. Dobimo vektorje značilk, ki imajo 759 dimenzij. Nato se izračuna tri podobnostne matrike za tri različne razdalje med značilkami (Evklidovo, City Block, in kosinusno razdaljo). V naslednjem koraku se dobljene matrike ovrednotijo in izrišejo se rezultati razpoznavanja obrazov s PCA postopkom.

DESKRIPTORJI:

Za izračun SURF, FREAK in BRISK deskriptorjev določimo matriko, kjer so zapisane koordinate 196 značilnih točk. Značilne točke smo izbrali sami, in sicer tako, da smo sestavili mrežo, kjer smo vzeli vsako sedmo slikovno točko na sliki. Začeli smo iz leve proti desni ter od gor proti dol. Mreža je velikosti 14×14 točk.

SURF:

S pomočjo funkcije iz Matlabove zbirke orodij za strojni vid izračunamo SURF deskriptorje okoli naših značilnih točk za vse slike iz galerije in testne množice. Vsaka značilna točka je opisana s SURF vektorjem, ki ima 64 dimenzij. Ker je točk 196, ima skupni SURF vektor posamezne slike 12544 elementov. Izračunajo se podobnostne matrike za tri razdalje, nato se ovrednotijo ter izrišejo rezultati.

FREAK:

S pomočjo funkcije iz Matlabove zbirke orodij za strojni vid izračunamo binarne FREAK deskriptorje okoli značilnih točk za vse slike iz galerije in testne

množice. Razlika s SURF deskriptorjem je v številu veljavnih značilnih točk. Ker se za določitev FREAK deskriptorjev uporablja vzorec iz Slike 2.5, se robne značilne točke, katerih okolica sega izven območja slike, ne upoštevajo pri izračunu. Ta problem rešimo tako, da okoli celotne slike dodamo črn pas širine 32 slikovnih točk. Razširjene slike so dimenzij 164×164 . Pri tako razširjeni sliki funkcija upošteva vseh 196 značilnih točk za izračun deskriptorja slike. Vsaka značilna točka je opisana z vektorjem dolžine 64. Vsak element v tem vektorju ima vrednost med 0 in 255, ki predstavlja nivo sivine. Če se ga zapiše z 8 bitnim binarnim številom je deskriptor značilne točke dolg 512 bitov. Skupni FREAK deskriptor posamezne slike je sestavljen iz nanizanih FREAK deskriptorjev značilnih točk, iz česar sledi, da ima skupni vektor 12544 elementov (196×64). Potem se izračunajo podobnostne matrike za tri razdalje (Evklidovo, City Block, kosinusno), nato se še ovrednotijo ter izrišejo rezultati.

BRISK:

Enak protokol kot za FREAK velja tudi za binarni BRISK deskriptor. Da funkcija z BRISK vzorcem (Slika 2.5) upošteva vseh 196 značilnih točk iz naše mreže, tudi tukaj vsaki sliki dodamo črn pas širine 32 slikovnih točk. Vsaka značilna točka je opisana z 64 dimenzionalnim BRISK vektorjem, zato ima BRISK deskriptor posamezne slike 12544 elementov.

KOMBINIRANI POSTOPKI:

SURF+PCA:

Pri tem postopku se najprej izračunajo SURF deskriptorji učnih slik in nato se iz njih zgradi nov PCA podprostor. Matrika učnih slik je v tem primeru velikosti 12544×760 . Dimenzijo podprostora določimo z rangom učne matrike, ki je zmanjšan za ena. Dobimo PCA prostor, ki ima 304 dimenzije. SURF deskriptorji galerijskih in testnih slik se linearno preslikajo v manj razsežni podprostor, kjer dobimo nove SURF+PCA značilke s 304 dimenzijami. SURF deskriptorji slik so se iz vektorjev dolžine 12544 preslikali v vektorje velikosti 304. Te primerjamo med sabo in sestavimo podobnostne matrike za tri razdalje. Ovrednotimo jih in izpišemo rezultate.

FREAK+PCA:

Izračunajo se binarni FREAK deskriptorji učnih slik, ki so razširjene s črnim pasom širine 32 slikovnih točk. Matrika učnih slik je velikosti 12544×760 . Nato se zgradi nov PCA podprostor. Dimenzija podprostora je enaka rang učne matrike zmanjšanemu za ena, tj. 299. S preslikavo FREAK deskriptorjev slik iz testne množice in galerije v linearni podprostor dobimo nove FREAK+PCA značilne

vektorje, ki imajo 299 elementov. Prišlo je do preslikave vektorjev z dolžino 12544 v vektorje z 299 dimenzijami. Dobljene značilne vektorje primerjamo v podobnostnih matrikah s tremi različnimi razdaljami in ovrednotimo rezultate verifikacije.

BRISK+PCA

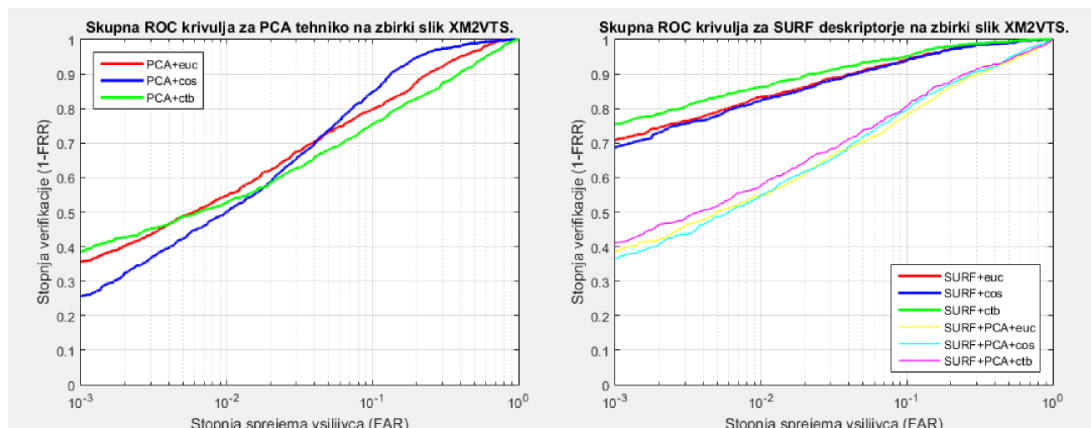
Enak protokol kot pri kombiniranem FREAK+PCA se ponovi tudi pri BRISK+PCA postopku. Najprej se izračunajo BRISK deskriptorji razširjenih učnih slik. Matrika učnih slik je velikosti 12544×760 in ima rang 304. Iz teh deskriptorjev se sestavi nov PCA podprostor, ki ima 303 lastne vektorje. Preslikamo BRISK deskriptorje testnih in galerijskih slik v manj razsežen prostor in dobimo nove BRISK+PCA značilne vektorje. Preslikava je zmanjšala dimenzijo vektorjev iz 12544 na 303. Izračunajo se še matrike podobnosti z vsemi tremi razdaljami in izrišejo se rezultati.

4.1.2 Rezultati in diskusija

V Tabeli 4.1 so prikazani rezultati eksperimentov na zbirki obrazov XM2VTS s tremi načini vrednotenja. Najboljše vrednosti pri posameznih tehnikah razpoznavanja so obarvane. Za vsak postopek so podani grafi, ki prikazujejo ROC krivulje.

PCA:

Pri tehniki analize glavnih komponent so rezultati med posameznimi kategorijami vrednotenja najbolj spremenljivi. Tako opazimo, da se delovni točki, kjer sta EER in minHTER najmanjša, nahajata pri računanju razdalj s kosinusno enačbo. To je edini primer, kjer se je najbolj izkazala kosinusna razdalja. Iz grafa razberemo tudi, da stopnja verifikacije začne v drugem delu krivulje hitreje naraščati z naraščanjem FAR kot v ostalih dveh primerih. Uspešnost razpoznavanja klienta pri napaki FAR=1% je 54.80%, in sicer z Evklidovo razdaljo. Ob zaostritvi pogojev vstopa v sistem, pri FAR=0.1% in FAR=0.01%, je stopnja verifikacije največja pri algoritmu razdalje City Block. Celotne ROC krivulje so vidne na Sliki 4.2 levo.



Slika 4.2: ROC krivulje za PCA (levo) ter SURF in SURF+PCA tehniko (desno).

DESKRIPTORJI:

SURF:

Deskriptorji SURF so se izkazali z daleč najmanjšima napakama EER in minHTER ter največjo stopnjo verifikacije pri vseh treh pragovih iz tabele. Vseh devet vrednosti iz tabele, ki opisujejo stopnjo verifikacije s SURF postopkom, je bilo boljših kakor pri kateremkoli drugem postopku. Najboljši rezultati so bili s City Block razdaljo. Ko je FAR=1%, je uspešnost razpoznavanja 86.20%. To pomeni, da bo klienta pri stotih prijavah sistem povprečno zavrnil približno štirinajstkrat. Ob tem bo vsiljivec prišel v sistem v povprečju na vsak njegov stoti poizkus. ROC krivulje so izrisane na Sliki 4.2 desno.

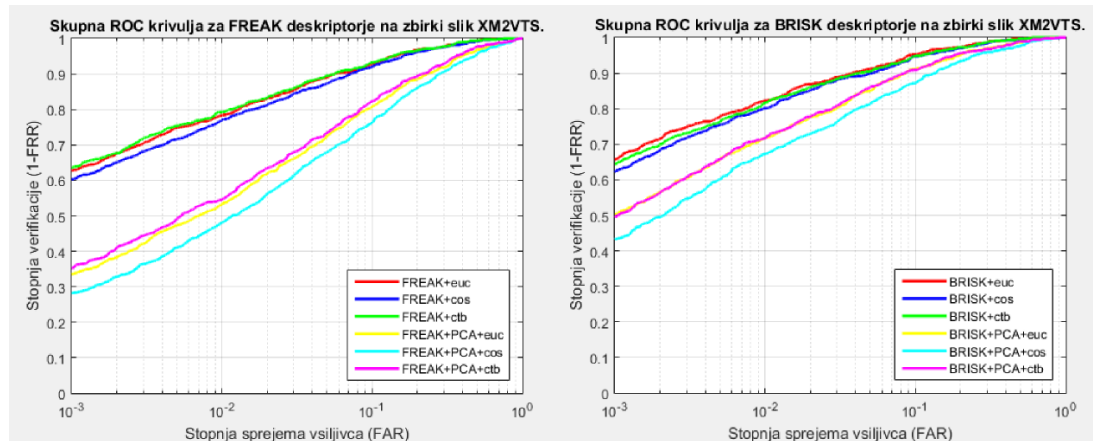
FREAK:

Od dveh obravnavanih binarnih deskriptorjev se je ta obnesel nekoliko slabše, razen v delovnih točkah pri FAR=0.01%, je bila pri vseh treh izračunih razdalje verifikacija malenkost boljša s FREAK kot z BRISK deskriptorji. Razpoznavanje obrazov ob uporabi City Block algoritma je proizvedlo manjše napake v primerjavi z Evklidovo in kosinusno razdaljo. Glede na SURF je FREAK dosegel za nekaj odstotkov slabše rezultate. Verifikacija v ostalih delovnih točkah je razvidna iz Slike 4.3 levo.

BRISK:

Razpoznavanje z binarnimi deskriptorji BRISK je edini primer, kjer je računanje podobnosti z Evklidovo razdaljo napravilo najmanjši napaki EER in minHTER ter največjo stopnjo verifikacije. Kakor je vidno iz tabele, je pri FAR=1% in FAR=0,1% BRISK za nekaj odstotnih točk uspešnejši pri ugotavljanju istovetnosti klientov in zavrnitvi vsiljivcev kot FREAK. Je drugi najuspešnejši postopek od vseh sedmih, ki smo jih testirali. ROC krivulje so izrisane na Sliki 4.3. Pri obeh binarnih

deskriptorjih lahko iz grafov vidimo, da se potek krivulj zelo malo razlikuje pri Evklidski in City Block razdalji. V našem primeru red razdalje Minkovskega (enačba (2.20)) ne vpliva dosti na uspešnost razpoznavanja pri binarnih deskriptorjih.



Slika 4.3: ROC krivulje razpoznavanja z binarnimi deskriptorji ter kombinacijo binarnih deskriptorjev in PCA postopka.

KOMBINIRANI POSTOPKI:

Motivacija kombiniranih postopkov je bila ugotoviti, ali se z implementacijo analize glavnih komponent na SURF, FREAK in BRISK deskriptorjih lahko znebimo nekaj nezaželene variabilnosti v osvetlitvi na slikah. Torej, kako vpliva preslikava deskriptorjev v prostor z manj dimenzijami na uspešnost razpoznavanja. Taki vektorji značilno so dosti krajši od originalnih in nosijo posledično manj informacije o sliki. ROC krivulje kombiniranih postopkov so prikazane na Sliki 4.2 desno in na Sliki 4.3.

SURF+PCA:

Tako kot pri SURF deskriptorju je tudi pri kombiniranem SURF+PCA bila najuspešnejša City Block razdalja. Vidimo, da sta bili tukaj napaki EER in minHTER več kot dvakrat večji v primerjavi s SURF postopkom. Z dodatkom analize glavnih komponent se je dolžina značilnih vektorjev skrajšala iz 12544 elementov na 304 elemente in posledično se je stopnja verifikacije zmanjšala za 30 do 35 odstotkov.

FREAK+PCA:

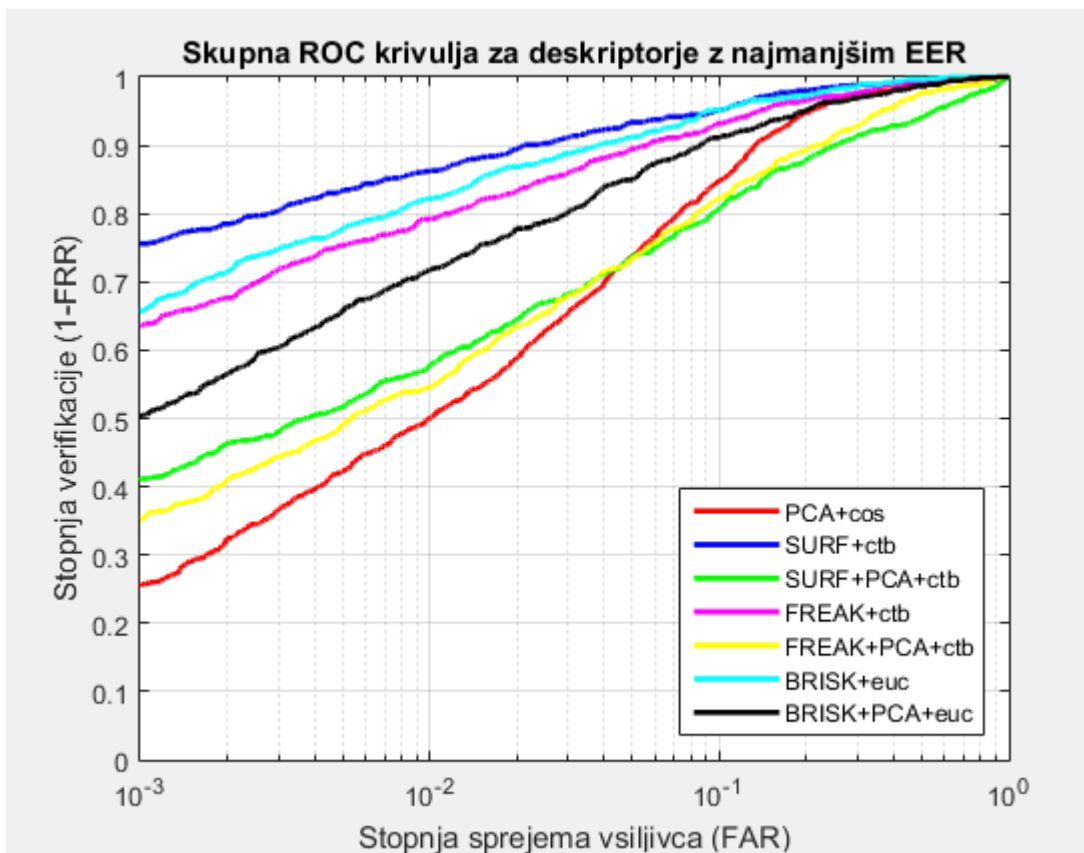
Med kombiniranimi postopki je bil ta glede na stopnjo razpoznavanja najmanj uspešen. Tako kot pri samem lokalnem FREAK deskriptorju je tudi kombinirani postopek največ obrazov pravilno razpoznal s City Block algoritmom. Ob preslikavi v PCA podprostor so se vektorji zmanjšali iz 12544 na 299 elementov, razpoznavanje pa v splošnem za okoli 30%. Čeprav ima postopek v vseh primerih

manjšo stopnjo verifikacije od SURF+PCA pa doseže manjši napaki EER in minHTER pri Evklidovi in City Block razdalji.

BRISK+PCA:

Kakor je razvidno iz Tabele 4.1, je bila med tremi kombiniranimi postopki pri razpoznavanju obrazov najuspešnejša kombinacija med BRISK in PCA. Opazimo tudi, da se stopnja verifikacije ni toliko zmanjšala kot pri ostalih dveh kombiniranih postopkih v primerjavi s samimi lokalnimi deskriptorji. Glede na lokalni deskriptor BRISK se je zmanjšala za cca. 10 do 20 odstotkov, ob zmanjšanju vektorjev iz 12544 elementov na 303 elemente. Prav tako se je najbolj izkazal Evklidov način izračuna razdalje med značilnimi vektorji, vendar mu takoj sledi City Block razdalja, kar se lepo vidi na Sliki 4.3 desno, kjer ROC krivulji skoraj sovpadata.

Slika 4.4 prikazuje ROC krivulje najboljših postopkov, kjer je merilo uspešnosti enaka stopnja napake. Najmanjši EER je imel SURF, na drugem mestu je bil binarni deskriptor BRISK, sledi binarni FREAK, kombiniran BRISK+PCA, PCA, FREAK+PCA in SURF+PCA postopek.



Slika 4.4: Skupen prikaz ROC krivulj postopkov z najmanjšim EER na zbirki XM2VTS.

Sprejem klientov v sistem in zavrnitev nepooblaščenih oseb je bil najbolj pravilen s SURF deskriptorji ob uporabi City Block razdalje, ki se je pokazala kot najbolj učinkovit algoritem za merjenje podobnosti tudi pri večini ostalih postopkov. Iz tabele opazimo, da je bil padec uspešnosti verifikacije med tremi obravnavanimi delovnimi točkami največji pri PCA in kombiniranih postopkih. Pri kombiniranih se je stopnja verifikacije v splošnem zmanjšala za tretjino ali več. Podrobnejši podatki za vse primere so v Tabeli 4.1.

<i>Vrsta značilk/ deskriptorja</i>	<i>Način izračuna razdalje</i>	<i>EER [%]</i>	<i>minHTER [%]</i>	<i>Stopnja verifikacije (FAR=1%) [%]</i>	<i>Stopnja verifikacije (FAR=0.1%) [%]</i>	<i>Stopnja verifikacije (FAR=0.01%) [%]</i>
PCA	Evklidova	15.83	14.92	54.80	35.50	23.20
	City Block	17.88	17.15	52.80	38.50	28.20
	kosinusna	12.18	11.50	50.10	25.50	13.90
SURF	Evklidova	7.32	6.87	83.40	70.80	57.00
	City Block	6.20	5.75	86.20	75.60	62.80
	kosinusna	7.40	7.11	82.20	68.70	55.00
SURF+PCA	Evklidova	16.22	15.78	54.80	38.60	23.90
	City Block	14.65	14.25	57.70	40.90	26.40
	kosinusna	15.34	14.77	54.90	36.50	22.30
FREAK	Evklidova	8.32	7.77	78.30	62.60	49.50
	City Block	8.20	7.62	79.20	63.50	50.70
	kosinusna	8.68	8.54	77.00	60.20	46.60
FREAK+PCA	Evklidova	14.70	14.49	53.20	33.50	20.20
	City Block	14.09	13.62	54.50	35.20	22.40
	kosinusna	16.43	16.20	48.10	28.10	17.40
BRISK	Evklidova	7.19	6.81	82.20	65.60	49.30
	City Block	7.29	7.03	81.50	64.40	48.70
	kosinusna	7.52	7.27	80.10	62.20	45.40
BRISK+PCA	Evklidova	9.13	9.10	71.70	50.30	30.40
	City Block	9.29	9.12	71.70	49.40	29.80
	kosinusna	11.30	11.06	67.20	43.00	25.00

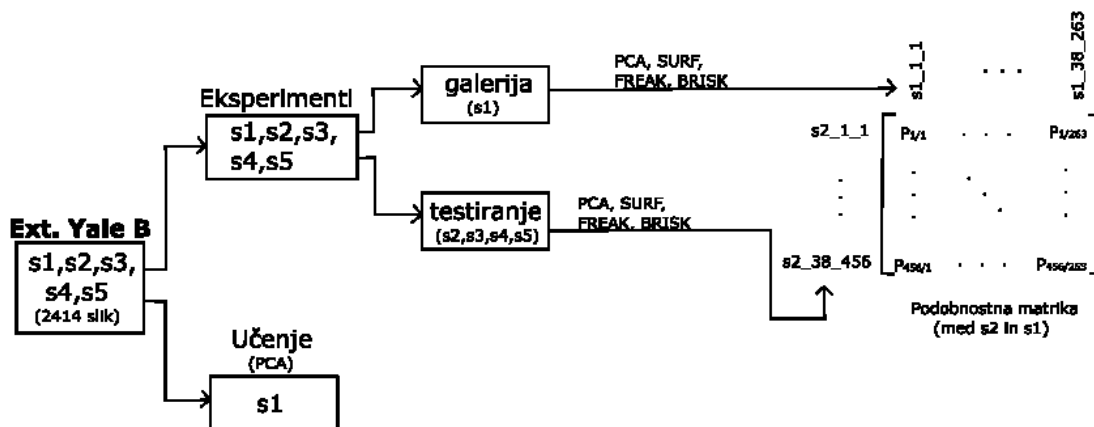
Tabela 4.1: Rezultati razpoznavanja obrazov iz zbirke XM2VTS podani v odstotkih.

4.2 Eksperimenti na zbirki obrazov "Extended Yale B"

Na omenjeni zbirki, kjer so slike z različnimi spremembami v osvetljenosti, smo raziskovali in primerjali razpoznavanje s PCA, SURF, FREAK in BRISK postopkom. Eksperimentov s kombiniranimi postopki nismo naredili, ker smo pri prejšnji zbirki prišli do zaključka, da ne prinesejo dodane vrednosti. Pri vseh štirih postopkih smo računali razdaljo med značilnimi vektorji na tri načine, in sicer z Evklidovo, City Block ter kosinusno razdaljo. Uspešnost razpoznavanja smo vrednotili tako, kot je opisano v razdelku 3.3.

4.2.1 Protokol

Bločni diagram protokola je prikazan na Sliki 4.5. V zbirki EYB je 2414 slik razdeljenih na pet podmnožic. Princip izvedbe eksperimentov je podoben kot pri prejšnji podatkovni zbirki. Podmnožico s_1 uporabimo za učenje PCA postopka in kot galerijo pri vseh štirih algoritmih. Podmnožice slik s_2 , s_3 , s_4 in s_5 uporabimo kot testne slike. V podobnostni matriki se primerjajo deskriptorji slik med s_1 in s_2 , s_1 in s_3 , s_1 in s_4 ter s_1 in s_5 . To naredimo za vse štiri algoritme in vse tri načine izračuna razdalj med značilnimi vektorji. Matrike podobnosti so različnih velikosti, odvisno od števila slik v posamezni podmnožici. Kadar določena oseba predstavlja klienta, slike ostalih predstavljajo vsiljivce v sistem.



Slika 4.5: Shema protokola za izračun podobnostne matrike med galerijo in testno podmnožico s_2 .

PCA:

Najprej uvozimo 263 slik iz podmnožice s_1 in jih uporabimo kot učne slike za izgradnjo PCA podprostora. Matrika učnih slik je dimenzij 10000×263 . Rang matrike učnih slik, zmanjšan za ena, določimo kot želeno velikost podprostora. Dobimo PCA podprostor, ki ima 262 lastnih vrednosti oz. vektorjev. PCA značilne

vektorje dobimo s preslikavo slik iz galerije in testnih množic v manj razsežen prostor, ki ima 262 dimenzij. Nato se izračunajo podobnostne matrike za vse štiri kombinacije množic slik (s1 in s2, s1 in s3, s1 in s4, s1 in s5). Vsaka izmed kombinacij se izračuna po treh različnih metodah za določitev razdalje med vektorji, in sicer Evklidovo, City block ter kosinusno. V naslednjem koraku se dobljene matrike ovrednotijo in izrišejo se rezultati razpoznavanja obrazov s PCA postopkom.

DESKRIPTORJI:

Matrika koordinat 196 značilnih točk za določitev SURF, FREAK in BRISK deskriptorjev je enaka kot pri eksperimentih na zbirki XM2VTS.

SURF

Matlabova funkcija opiše značilne točke slik iz galerije in testnih množic s SURF vektorji. Ker je vsaka značilnost posamezne slike opisana s 64 dimenzionalnim lokalnim SURF deskriptorjem, je skupen SURF deskriptor dolg 12544 elementov. Nato se sestavijo matrike podobnosti za tri razdalje in prikažejo rezultati.

FREAK

Binarne FREAK deskriptorje slik vseh petih podmnožic izračunamo enako, kot je že opisano v razdelku 4.1.1. Tudi tukaj dodamo okoli slik črn pas, zato da algoritem upošteva vseh 196 značilnih točk. Lokalni FREAK deskriptor ima 64 elementov, celoten FREAK deskriptor posamezne slike pa 12544. Naslednji korak je izračun matrik, izpis napak in rezultatov verifikacije.

BRISK

Enako kot zgoraj pri FREAK deskriptorjih naredimo tudi pri binarnih BRISK deskriptorjih. Slike na robovih razširimo s črnim pasom širine 32 slikovnih točk. Vsako značilno točko opiše lokalni BRISK deskriptor s 64 dimenzijami, zato je celoten BRISK deskriptor slike dolg 12544 elementov. Po izračunu matrik podobnosti se izrišejo rezultati razpoznavanja.

4.2.2 Rezultati in diskusija

Rezultati razpoznavanja za podmnožico s2:

Od vseh štirih testnih množic vsebuje s2 slike z najmanj zatemnitve. Zato je uspešnost prepoznavne klientov in vsiljivcev mnogo večja kot pri ostalih podmnožicah. Manjša območja zatemnitve so večinoma na sredini slike ob robu nosu.

Glede na rezultate analize deskriptorjev iz podpoglavja 4.1 se je tudi v tem primeru generalno gledano najbolje odrezal SURF postopek, vendar je bila prednost pred ostalimi postopki manjša kot pri zbirki XM2VTS. V delovni točki, kjer je FAR=1%, pa se je zgodilo, da je FREAK deskriptor s kosinusno razdaljo dosegel celo malenkost boljše razpoznavanje kot SURF. Vendar je nato z manjšanjem razmerja FAR pravilnejše razpoznavanje izvedel SURF. Verifikacija s SURF deskriptorjem je bila približno 98%, kar pomeni, da bo od stotih prijav klienta povprečno samo dvakrat zavrnilo. Ob tem bo primerov, ko bo vsiljivca sprejelo v sistem, 1%. Postopek izračuna razdalje med značilnimi vektorji, ki je v povprečju izvedel najbolj pravilno razpoznavanje, je bil City Block.

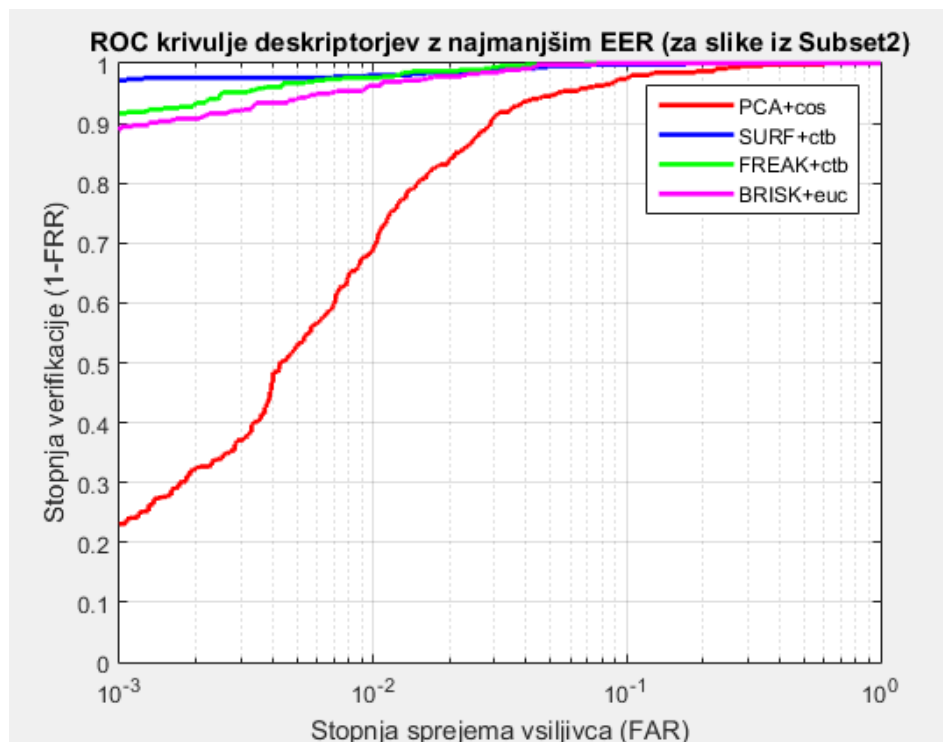
Če je pri zbirki XM2VTS med binarnimi deskriptorji bolje razpoznaval BRISK, je tukaj vodilno vlogo prevzel FREAK. Razlika ni velika, vendar se večja z manjšanjem stopnje sprejema vsiljivca.

Opazimo, da je PCA postopek najbolj občutljiv na zmanjšanje napake FAR, saj je tam padec uspešnosti verifikacije najbolj strm. Še posebej izstopa primer s kosinusnim izračunom razdalje.

Podrobnejši podatki so v Tabeli 4.2 in Sliki 4.6.

<i>Vrsta značilk/ deskriptorja</i>	<i>Način izračuna razdalje</i>	<i>EER [%]</i>	<i>minHTER [%]</i>	<i>Stopnja verifikacije (FAR=1%) [%]</i>	<i>Stopnja verifikacije (FAR= 0.1%) [%]</i>	<i>Stopnja verifikacije (FAR=0.01%) [%]</i>
PCA	Evklidova	11.85	11.20	66.67	43.42	30.04
	City Block	8.11	7.49	76.97	56.58	35.75
	kosinusna	5.25	5.07	69.08	23.03	6.36
SURF	Evklidova	1.76	1.38	97.81	96.05	90.13
	City Block	1.59	1.27	98.03	97.15	93.86
	kosinusna	1.76	1.44	97.81	95.83	89.69
FREAK	Evklidova	1.75	1.46	97.81	91.45	85.53
	City Block	1.37	1.37	97.59	91.67	83.99
	kosinusna	1.54	1.30	98.25	91.45	84.87
BRISK	Evklidova	2.19	1.96	96.27	88.82	80.04
	City Block	2.23	2.12	95.83	87.28	77.63
	kosinusna	2.42	2.25	95.39	85.96	71.71

Tabela 4.2: Rezultati razpoznavanja na podmnožici s2.



Slika 4.6: ROC krivulje verifikacije na podmnžici s2 postopkov z najmanjšim EER.

Rezultati razpoznavanja za podmnžico s3:

V tej podmnžici je obseg zatemnitve obrazov nekoliko večji kot pri s2, kar pa že občutno vpliva na rezultate razpoznavanja. Deli obraza, ki so zatemnjeni, so predvsem območje ličnic ter usta in nos ob strani.

Stopnja verifikacije s SURF je bila ob FAR=1% za tretjino manjša, ob FAR=0.1% za več kot polovico in ob FAR=0.01% kar trikrat manjša kot pri prejšnji množici. Delovna točka, kjer sta napaki EER in minHTER najmanjši, se je tokrat nahajala v Evklidovi matriki podobnosti. Napaki sta bili sicer približno šestkrat in sedemkrat večji kot pri najboljšem primeru iz s2.

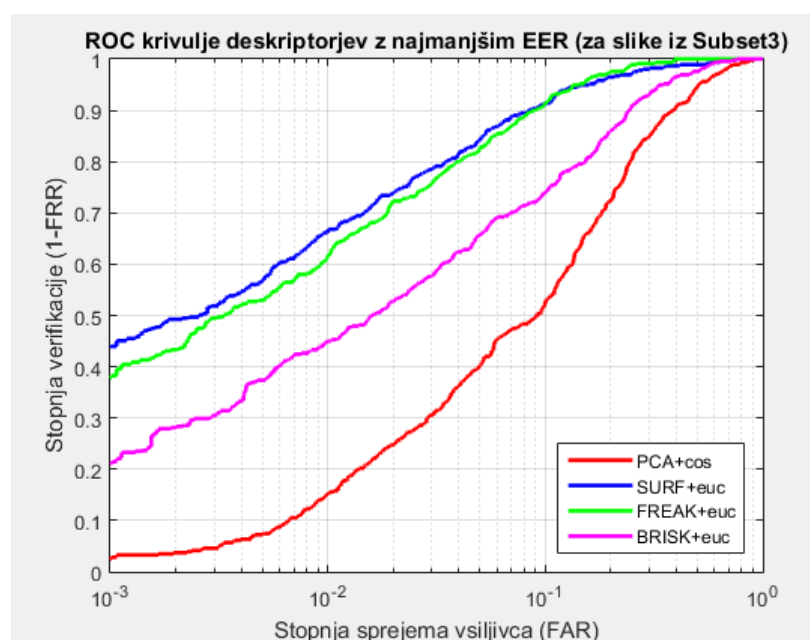
Med samima binarnima algoritmoma deskripcije (FREAK, BRISK) je nastopilo že nekoliko večje odstopanje. Iz Slike 4.7 vidimo, da je FREAK občutno bolje razpoznaval zatemnjene obraze kot BRISK.

PCA postopek je proizvedel največje napake in najnižje verifikacije, saj so značilni vektorji najkrajši (262 elementov) in zato nosijo najmanj informacije o obrazu na sliki. Ker množica s3 vsebuje 455 slik, predstavlja 0.44% 2 sliki. Od 455 prijav klientov v sistem sta samo 2 prijavi pravilno sprejeti s PCA postopkom in uporabo Evklidove metode ob FAR=0.01%. Razpoznavanje obrazov na slikah z opazno zatemnitvijo je s PCA postopkom bolj ali manj odpovedalo.

Spodnja slika prikazuje primere postopkov, kjer je bila napaka EER najmanjša. Opazimo, da prednjačita SURF in FREAK deskriptor. Več je razvidno tudi iz Tabele 4.3.

Vrsta značilk/ deskriptorja	Način izračuna razdalje	EER [%]	minHTER [%]	Stopnja verifikacije (FAR=1%) [%]	Stopnja verifikacije (FAR=0.1%) [%]	Stopnja verifikacije (FAR=0.01%) [%]
PCA	Evklidova	32.25	31.42	9.45	3.08	0.44
	City Block	27.27	26.85	25.27	11.21	3.30
	kosinusna	23.03	21.72	15.16	2.42	1.54
SURF	Evklidova	9.25	8.94	66.15	43.96	29.89
	City Block	10.54	10.31	61.32	44.62	31.43
	kosinusna	9.68	9.20	65.71	43.30	28.57
FREAK	Evklidova	9.45	8.86	61.32	37.80	19.78
	City Block	10.11	9.94	60.00	36.92	19.56
	kosinusna	9.49	8.92	60.88	38.68	23.08
BRISK	Evklidova	17.36	16.88	44.84	21.10	9.89
	City Block	17.48	17.08	44.18	20.66	8.57
	kosinusna	18.66	18.21	42.20	19.34	7.25

Tabela 4.3: Rezultati razpoznavanja na podmnožici s3.



Slika 4.7: ROC krivulje verifikacije na podmnožici s3 postopkov z najmanjšim EER.

Rezultati razpoznavanja za podmnožico s4:

Tukaj zatemnitve obsegajo že večji del obraza, kar se odraža v zanemarljivi prepoznavi.

Edino SURF algoritem je v delovni točki FAR=1% prepoznal malo več kot četrtno klientov. Prvič se je zgodilo, da so bili rezultati verifikacije v vseh treh obravnavanih delovnih točkah iz tabele pri dveh različnih razdaljah enaki. Pri SURF so bili rezultati enaki z Evklidovo in kosinusno razdaljo.

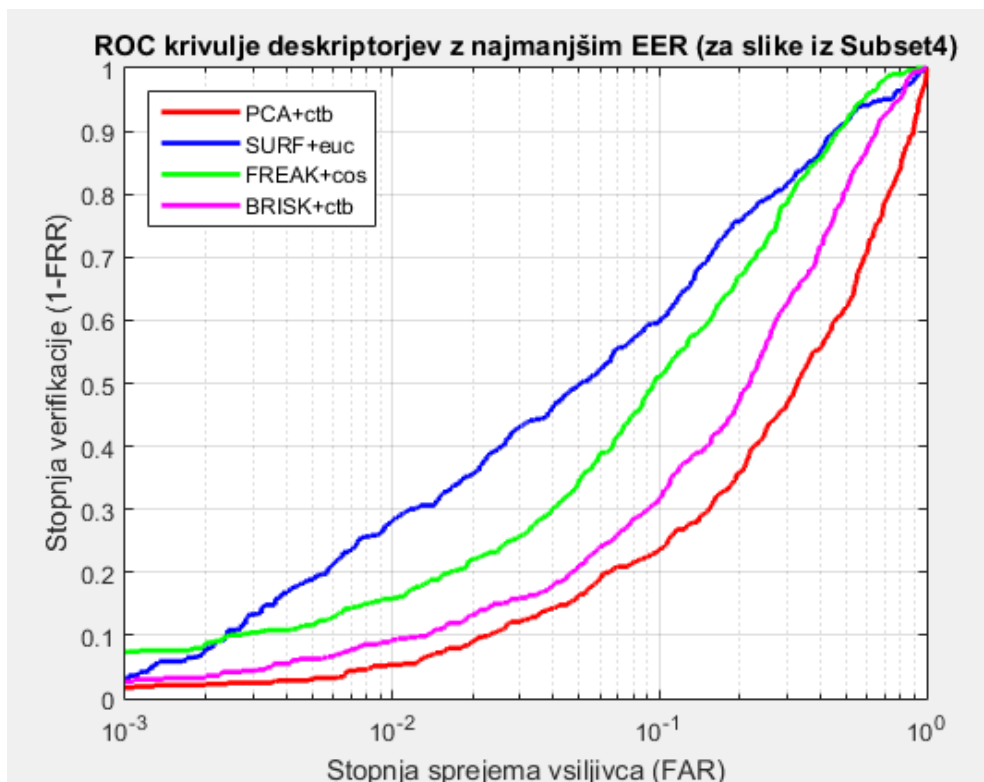
Od binarnih deskriptorjev je FREAK ponovno dosegel pravilnejše razpoznavanje, in sicer s kosinusno razdaljo. Iz spodnjega grafa ROC krivulj vidimo, da je FREAK pri velikem in majhnem FAR izvedel boljšo verifikacijo od SURF deskriptorja.

Ker ima s4 množica 526 slik, predstavlja verifikacija 0.19% samo eno pravilno sprejeto sliko osebe. To se je zgodilo v treh primerih s PCA postopkom.

Tabela 4.4 in ROC krivulje iz Slike 4.8 ponazarjajo podrobnejše stanje.

<i>Vrsta značilk/ deskriptorja</i>	<i>Način izračuna razdalje</i>	<i>EER [%]</i>	<i>minHTER [%]</i>	<i>Stopnja verifikacije (FAR=1%) [%]</i>	<i>Stopnja verifikacije (FAR=0.1%) [%]</i>	<i>Stopnja verifikacije (FAR=0.01%) [%]</i>
PCA	Evklidova	42.79	42.05	1.71	0.19	0.19
	City Block	42.59	41.01	5.32	1.71	0.57
	kosinusna	42.97	42.31	2.28	0.95	0.19
SURF	Evklidova	22.45	21.66	28.14	3.04	0.57
	City Block	23.03	22.75	27.19	2.66	0.57
	kosinusna	22.45	21.91	28.14	3.04	0.57
FREAK	Evklidova	29.20	29.20	11.98	4.37	1.14
	City Block	29.29	29.10	11.41	4.56	1.52
	kosinusna	26.06	25.19	15.78	7.22	3.42
BRISK	Evklidova	34.83	34.19	8.17	2.85	1.90
	City Block	33.83	33.48	9.13	2.47	1.71
	kosinusna	35.18	34.81	8.56	2.28	1.14

Tabela 4.4: Rezultati razpoznavanja na podmnožici s4.



Slika 4.8: ROC krivulje verifikacije na podmnožici s4 postopkov z najmanjšim EER.

Rezultati razpoznavanja za podmnožico s5:

V množici s5 so osvetljeni samo še posamezni deli obraza, kakor je razvidno iz Slike 3.2.

Ko smo računali razdaljo med SURF vektorji s kateri kolim od treh načinov, je bila ob stopnji sprejema vsiljivca 1% napaka FRR približno 88%. V primeru Evklidove in kosinusne razdalje je bila uspešnost verifikacije enaka za vse tri vrednosti FAR iz tabele.

Oba binarna deskriptorja sta večinoma dosegla višjo stopnjo verifikacije kot SURF približno do vrednosti FAR=3% (razberemo iz Slike 4.9 in Tabele 4.5). Prvič med vsemi opravljenimi eksperimenti se je zgodilo, da je katerikoli postopek dosegel boljše razpoznavanje od SURF deskriptorja na tako širokem obsegu FAR vrednosti, kot sta ga FREAK in BRISK na podmnožici s5. Razpoznavanje obrazov s FREAK deskriptorjem je bilo najuspešnejše ob uporabi kosinusne razdalje, in sicer je prepoznal 16.67% klientov, ko je bila stopnja sprejema vsiljivcev 1%. V nekaterih primerih iz tabele pri uporabi drugih razdalj je BRISK izvedel boljšo verifikacijo kot FREAK z istimi razdaljami. Zanimivo je tudi, da sta oba binarna deskriptorja v določenih primerih opravila pravilnejšo verifikacijo in imela malenkost manjše

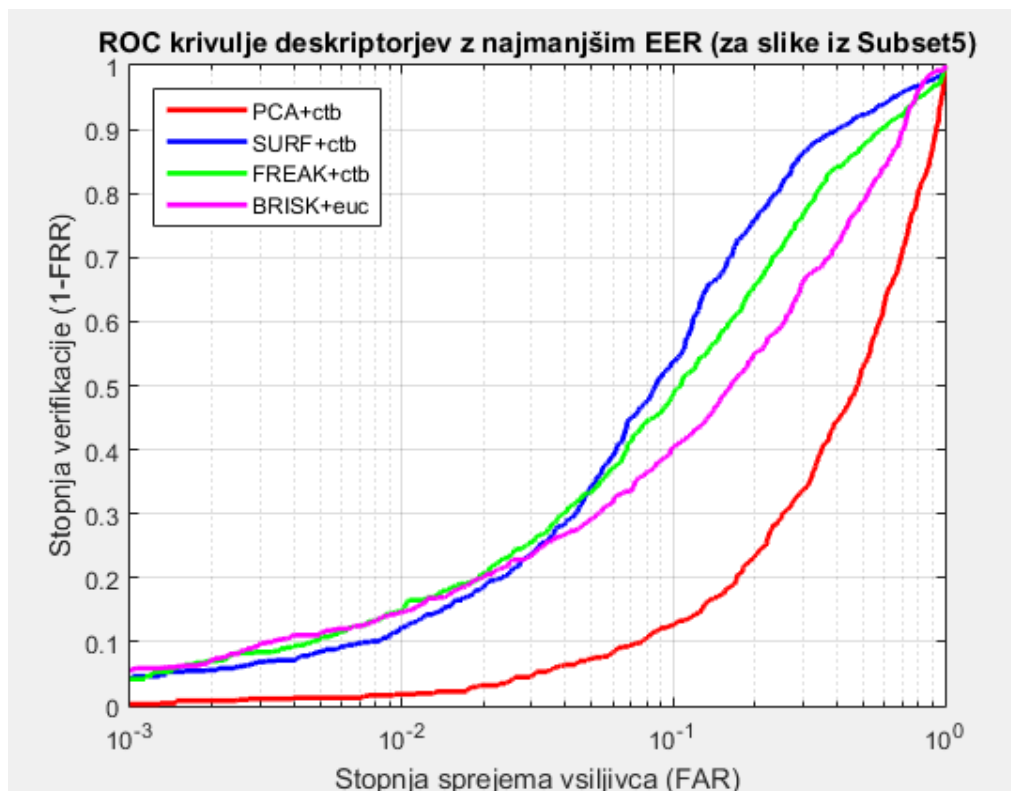
napake pri množici s5 kot pri s4, kjer so zatemnitve manjše, vendar niso toliko koncentrične.

Analiza glavnih komponent je komaj prepoznala kakšnega klienta. V tabeli vrednost 0.14% pomeni, da je bila prepoznana slika enega klienta, ker je vseh testnih slik 714.

Kljub velikemu obsegu zatemnitev v testni množici s5 je z vsakim postopkom bila prepoznana vsaj ena slika klienta. V spodnji tabeli in grafu so rezultati verifikacije opravljene na podmnožici s5.

<i>Vrsta značilk/ deskriptorja</i>	<i>Način izračuna razdalje</i>	<i>EER [%]</i>	<i>minHTER [%]</i>	<i>Stopnja verifikacije (FAR=1%) [%]</i>	<i>Stopnja verifikacije (FAR=0.1%) [%]</i>	<i>Stopnja verifikacije (FAR=0.01%) [%]</i>
PCA	Evklidova	48.84	48.35	0.70	0.14	0.14
	City Block	48.60	47.47	1.82	0.28	0.14
	kosinusna	49.58	47.77	1.54	0.70	0.14
SURF	Evklidova	24.72	24.20	12.04	4.48	0.98
	City Block	22.02	21.55	12.04	3.78	0.98
	kosinusna	24.76	24.21	12.04	4.48	0.98
FREAK	Evklidova	28.67	28.39	14.85	4.34	0.98
	City Block	26.65	26.37	14.85	3.64	0.84
	kosinusna	29.41	28.96	16.67	6.02	1.96
BRISK	Evklidova	32.36	31.88	14.57	5.18	1.82
	City Block	32.64	31.97	15.55	4.90	1.40
	kosinusna	32.68	31.76	15.27	4.62	1.40

Tabela 4.5: Rezultati razpoznavanja na podmnožici s5.



Slika 4.9: ROC krivulje verifikacije na podmnožici s5 postopkov z najmanjšim EER.

4.3 Dolžine značilnih vektorjev

V Tabeli 4.6 so prikazane dolžine značilnih vektorjev posameznih postopkov razpoznavanja, kakor je že razloženo v opisu protokolov (razdelek 4.1.1 in 4.2.1).

Dolžina značilnih vektorjev pri PCA in kombiniranih postopkih je bila odvisna od vrednosti ranga matrike učnih slik. V našem primeru smo to vrednost zmanjšali za ena in tako določili dolžino vektorja.

SURF, FREAK in BRISK deskriptorji slik so bili v eksperimentih na obeh zbirkah enako dolgi. Njihova dolžina je bila odvisna od števila značilnih točk na sliki, ki jih je v našem primeru bilo 196. Vsaka značilna točka je bila opisana z lokalnim deskriptorjem dolgim 64 elementov, zato je bil deskriptor posamezne slike dolg 12544 elementov.

Značilni vektor/ deskriptor	Dolžina značilnega vektorja slike iz zbirke	Dolžina značilnega vektorja slike iz zbirke
	XM2VTS	EYB
PCA	759	262
SURF	12544	12544
FREAK	12544	12544
BRISK	12544	12544
SURF+PCA	304	
FREAK+PCA	299	
BRISK+PCA	303	

Tabela 4.6: Tabela s primerjalnimi dolžinami značilnih vektorjev.

4.4 Časovna zahtevnost izračuna značilnih vektorjev

Tabela 4.7 prikazuje povprečne čase izračuna posameznih značilnih vektorjev z opremo in orodji opisanimi v razdelku 3.1.

Izmed deskriptorjev se je v naših eksperimentih najhitreje izračunal binarni BRISK, in sicer v 9.2 milisekunde. Avtorji binarnega FREAK deskriptorja v svojem članku navajajo [9], da je hitrejši od SURF in BRISK, vendar se v naših meritvah to ni izkazalo. Hitrost izračuna je odvisna od same implementacije in v Matlabu je v našem primeru FREAK deskriptor potreboval približno dvakrat več časa kot SURF in petkrat več kot BRISK.

Pri PCA postopku smo čas izračuna značilnega vektorja merili drugače. Zapis časa v tabeli je v dveh delih, ker je potrebno najprej zgraditi PCA podprostor in nato preslikati sliko v podprostor, da dobimo značilni vektor. Vrednost v oklepaju je čas, ki je bil potreben za izgradnjo podprostora, druga vrednost pa je čas preslikave slike v podprostor.

Značilni vektor/ deskriptor	Povprečen čas izračuna značilnega vektorja/ deskriptorja ene slike [ms]
PCA	(2330) + 0.2
SURF	19.7
FREAK	45.6
BRISK	9.2

Tabela 4.7: Tabela časovne zahtevnosti izračuna značilnih vektorjev.

5 ZAKLJUČEK

Preizkušali smo uspešnost razpoznavanja obrazov z različnimi deskriptorji v biometričnem sistemu, ki opravlja funkcijo verifikacije oseb. Primerjali smo PCA postopek, SURF, FREAK in BRISK deskriptorje ter kombinirane postopke SURF+PCA, FREAK+PCA in BRISK+PCA. Podobnost med posameznimi značilnimi vektorji smo računali na tri načine, in sicer z Evklidovo, City Block ter kosinusno razdaljo. Analizirali smo učinkovitost binarnih in ne binarnih deskriptorjev.

Rezultate razpoznavanja obrazov smo prikazali v obliki tabel in grafov. Ovrednotili smo jih z napakami FAR, FRR, HTER in EER. Soodvisnost med FAR in FRR v različnih delovnih točkah oz. pragovih so ponazarjale ROC krivulje. V tabelah so bili izpostavljeni rezultati verifikacije v treh delovnih točkah.

PCA postopek je transformiral slike v manj razsežen prostor, kjer se je opravila primerjava podobnosti. Za deskripcijo slik s SURF, FREAK in BRISK algoritmom smo sami določili lokacijo značilnih točk na slikah, in sicer s fiksno mrežo kvadratne oblike, ki je imela 196 točk. SURF postopek je vsako od 196 značilnih točk slike opisal s 64 dimenzionalnim vektorjem. Celoten SURF deskriptor slike je bil dolg 12544 elementov. Zaradi narave binarnih deskriptorjev, ker uporabljata določen vzorec za deskripcijo značilnih točk, sta FREAK in BRISK na začetku upoštevala manj značilnih točk. Tiste, ki so bile preblizu roba slike, sta ignorirala. Ta problem smo rešili tako, da smo okoli celotne slike dodali črn pas širine 32 slikovnih točk. Razširjena slika je bila dimenzij 164×164 . Pri tako razširjeni sliki je funkcija upoštevala vseh 196 značilnih točk za izračun binarnega deskriptorja slike. Vsaka značilna točka je bila opisana z vektorjem dolžine 64, slika pa z binarnim deskriptorjem dolžine 12544. Pri kombiniranih postopkih smo najprej izračunali SURF, FREAK in BRISK deskriptorje učnih slik in nato iz njih zgradili podprostore. Potem smo SURF, FREAK in BRISK deskriptorje slik iz testne množice in galerije preslikali v podprostore in jih primerjali.

Eksperimente smo opravili na dveh standardnih podatkovnih zbirkah, in sicer na XM2VTS in EYB. Na prvi zbirki smo naredili teste s PCA, SURF, FREAK, BRISK, SURF+PCA, FREAK+PCA in BRISK+PCA postopkom. Na drugi zbirki smo preizkusili PCA, SURF, FREAK in BRISK postopek. Sivinske slike so bile velikosti 100×100 slikovnih točk in zapisane v bmp formatu.

Zbirka XM2VTS je vsebovala slike 295 oseb z minimalno variabilnostjo v osvetlitvi. Vseh slik je bilo 2360. Izmed testiranih deskriptorjev je najvišjo stopnjo verifikacije in najmanjše napake dosegel SURF algoritem. Naslednji je bil binarni BRISK, ki je pri FAR=1% dosegel za 4 odstotke nižjo prepoznavo klientov. Binarni FREAK deskriptor je imel nižje stopnje verifikacije kot BRISK, razen pri delovni točki FAR=0.01%. Opazili smo tudi, da se z zaostrovanjem pogojev vstopa v sistem prepoznavanja klientov hitreje manjša pri binarnih deskriptorjih. PCA postopek je dosegel najslabše razpoznavanje, saj ima tudi najkrajše vektorje. Poleg omenjenih štirih algoritmov razpoznavanja smo na tej zbirki testirali še kombinirane postopke SURF+PCA, FREAK+PCA in BRISK+PCA. Zanimalo nas je, ali se lahko znebimo kaj variabilnosti v osvetlitvi slik. Izkazalo se je, da se stopnja verifikacije zmanjša za nekaj deset odstotkov, odvisno od delovne točke. V delovni točki pri FAR=1% je SURF+PCA postopek dosegel za 30% slabše razpoznavanje v primerjavi s SURF, FREAK+PCA za 25% slabše v primerjavi s FREAK in BRISK+PCA za 10% slabše v primerjavi z BRISK. S tem načinom so se značilni vektorji skrajšali iz 12544 elementov na dolžino 304, 299 in 303 elemente. Način izračuna podobnosti med vektorji, ki je zagotovil najvišje stopnje verifikacije, je v splošnem bil s City Block razdaljo.

Zbirka "Extended Yale B" je vsebovala sivinske slike 38 oseb, ki so bile razdeljene na pet podmnožic z različno stopnjo zatemnitve. Skupaj je bilo 2414 slik. Uspešnost prepoznave obrazov je bila pri množici s2 večja kot pri čisti XM2VTS, ker je bilo v zbirki EYB manj razredov vzorcev (oseb). V nekaj primerih v podmnožicah s2, s4 in s5 je FREAK deskriptor dosegal celo boljše razpoznavanje od SURF, vendar je stopnja verifikacije hitreje padala z manjšanjem napake FAR. Med obema binarnima deskriptorjema je bil FREAK skoraj v vseh primerih uspešnejši od BRISK. Zanimivo je tudi, da sta nekajkrat binarna deskriptorja boljše prepoznala kliente sistema pri množici s5 kot pri s4. Opazili smo, da je ob zatemnitvah na slikah FREAK bolje prepoznaval obraze, ko zatemnitev ni (zbirka XM2VTS) pa je bil boljši BRISK. PCA postopek je ob večanju zatemnitev zelo hitro izgubljal učinkovitost. Ugotovili smo, da je ob povečevanju zatemnitve uspešnost razpoznavanja med posameznimi deskriptorji manj predvidljiva. Od mer podobnosti je bila City Block najučinkovitejša, ampak ni tako prevladovala kot pri XM2VTS.

Nekajkrat se je zgodilo, da je bila stopnja verifikacije enaka pri različnih razdaljah. Spoznali smo pomembnost kvalitetnega zajema vhodnih slik in vpliv, ki ga ima na učinkovitost verifikacije. Že manjše motnje v osvetlitvi lahko občutno poslabšajo delovanje sistema.

Najhitrejši je bil izračun BRISK deskriptorja, in sicer približno dvakrat hitrejši od SURF in petkrat od FREAK. V sistemih razpoznavanja obrazov, kjer je ključna hitrost delovanja, bi bilo priporočljivo izbrati BRISK ali SURF deskriptor. V sistemih, kjer je pričakovati veliko motenj v osvetlitvi pri zajemu slike, bi bilo smotrno izbrati SURF ali FREAK deskriptor.

Potrebno je poudariti, da je bilo razpoznavanje obrazov s SURF, FREAK in BRISK deskriptorji realizirano na določeni mreži značilnih točk v obliki kvadrata. Ob drugačni mreži bi bili tudi rezultati razpoznavanja drugačni. V prihodnje bi bilo smiselno raziskati, kakšna je uspešnost verifikacije na omenjenih zbirkah še z drugimi algoritmi detekcije in deskripcije značilnih točk.

LITERATURA

- [1] A. K. Jain, A. Ross, S. Prabhakar, "An Introduction to Biometric Recognition". IEEE Transactions on circuits and systems for video technology, vol. 14, No.1, January 2004
- [2] XM2VTS database, dosegljivo:
<http://www.ee.surrey.ac.uk/CVSSP/xm2vtsdb/>. [Dostopano: 20.12. 2015]
- [3] Extended Yale B database, dosegljivo:
<http://vision.ucsd.edu/~leekc/ExtYaleDatabase/ExtYaleB.html>.
[Dostopano 20. 12. 2015]
- [4] V. Štruc, "The PhD Toolbox". Dosegljivo: http://luks.fe.uni-lj.si/sl/osebje/vitomir/face_tools/PhDface/. [Dostopano: 19. 12. 2015]
- [5] N. Pavešić, "Razpoznavanje vzorcev: Uvod v analizo in razumevanje vidnih in slušnih signalov, 1. zvezek". Fakulteta za elektrotehniko, Ljubljana, 2012
- [6] P. Viola, M. J. Jones, "Robust Real-Time Face Detection". International Journal of Computer Vision 57(2), pp. 137–154, 2004
- [7] H. Bay, A. Ess, T. Tuytelaars, L. Van Gool, "SURF: Speeded Up Robust Features". Computer Vision and Image Understanding (CVIU), Vol. 110, No. 3, pp. 346-359, 2008
- [8] S. Leutenegger, M. Chli, R. Y. Siegwart, "BRISK: Binary Robust Invariant Scalable Keypoints". IEEE Proceedings of the 2011 International Conference on Computer Vision (ICCV), pp. 2548-2555, Barcelona, 2011
- [9] A. Alahi, R. Ortiz, P. Vandergheynst, "FREAK: Fast Retina Keypoint". IEEE Proceedings of the 2012 Conference on Computer Vision and Pattern Recognition (CVPR), pp. 510-517, 2012
- [10] D. G. Lowe, "Distinctive image features from scale-invariant keypoints". International Journal of Computer Vision, 60, 2 (2004), pp. 91-110

- [11] E. Rosten, T. Drummond, "Machine learning for high-speed corner detection". European Conference on Computer Vision 1, pp. 430–443, 2006
- [12] E. Mair, G. D. Hager, D. Burschka, M. Suppa, G. Hirzinger, "Adaptive and generic corner detection based on the accelerated segment test". Proceedings of the European Conference on Computer Vision (ECCV'10), September 2010
- [13] S.Z. Li, A.K. Jain, "Handbook of Face Recognition", pp. 13-133, Springer, New York, USA, 2005
- [14] M.A. Turk, A.P. Pentland, "Face Recognition Using Eigenfaces". IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 586-591, Maui, USA, 1991
- [15] N. Pavešić, I. Fratrić, S. Ribarić, "Degradation of the XM2VTS database face images". Biometrics on the Internet: fundamentals, advances and applications – proceedings, pp. 17-21, Vigo (Spain): University of Vigo, 2004
- [16] J. Luetttinn, G. Maitre, "Evaluation protocol for the extended M2VTS database (XM2VTSDB)". IDIAP-COM 98-05, Julij, 1998
- [17] A. Y. Yang, J. Wright, Y. Ma, S. S. Sastry, "Feature Selection in Face Recognition: A Sparse Representation Perspective". University of California at Berkeley, Electrical Engineering and Computer Sciences, Technical Report No. UCB/EECS-2007-99, August 2007
- [18] V. M. Patel, T. Wu, S. Biswas, P. J. Phillips, R. Chellappa, "Dictionary-based Face Recognition Under Variable Lighting and Pose". IEEE Transactions on Information Forensics and Security, Volume:7, Issue: 3, pp. 954-965, 2012
- [19] V. Štruc, "Ovrednotenje sistemov za samodejno razpoznavanje obrazov – diplomsko delo". Fakulteta za elektrotehniko, Ljubljana, 2005
- [20] P. M. Panchal, S. R. Panchal, S. K. Shah, "A Comparison of SIFT and SURF". International Journal of Innovative Research in Computer and Communication Engineering, Vol. 1, Issue 2, April 2013
- [21] O. Miksik, K. Mikolajczyk, "Evaluation of local detectors and descriptors for fast feature matching". IEEE 21st International Conference on Pattern Recognition (ICPR), pp. 2681-2684, Tsukuba, 2012

DODATEK

A Kazalo slik

Slika 2.1: Biometrični sistem, ki opravlja verifikacijo osebe.....	17
Slika 2.2: Izračun integralne slike.....	21
Slika 2.3: Gaussov filter.....	22
Slika 2.4: Levo: Haarovo okno. Desno: Prikaz SURF deskriptorja na sliki...	23
Slika 2.5: Vzorca za vzorčenje pri BRISK in FREAK deskriptorju.....	24
Slika 2.6: Prikaz 45 parov za izračun orientacije FREAK deskriptorja.....	26
Slika 2.7: Vzorčni pari v štirih rojih.....	26
Slika 2.8: Primer dveh ROC krivulj za prikaz rezultatov verifikacije.....	29
Slika 3.1: Primer osmih zajemov obraza enega človeka iz zbirke XM2VTS.....	32
Slika 3.2: Prikaz nivoja zatemnitve obrazov v posameznih podmnožicah.....	33
Slika 4.1: Shema protokola za zbirko XM2VTS.....	36
Slika 4.2: ROC krivulje za PCA ter SURF in SURF+PCA tehniko.....	39
Slika 4.3: ROC krivulje razpoznavanja z binarnimi deskriptorji.....	40
Slika 4.4: ROC krivulje postopkov z najmanjšim EER na zbirki XM2VTS.....	41
Slika 4.5: Shema protokola za izračun matrike med galerijo in s2.....	43
Slika 4.6: ROC krivulje verifikacije na podmnožici s2.....	46
Slika 4.7: ROC krivulje verifikacije na podmnožici s3.....	47
Slika 4.8: ROC krivulje verifikacije na podmnožici s4.....	49
Slika 4.9: ROC krivulje verifikacije na podmnožici s5.....	51

B Kazalo tabel

Tabela 4.1: Rezultati razpoznavanja obrazov iz zbirke XM2VTS.....	42
Tabela 4.2: Rezultati razpoznavanja na podmnožici s2.....	45
Tabela 4.3: Rezultati razpoznavanja na podmnožici s3.....	47
Tabela 4.4: Rezultati razpoznavanja na podmnožici s4.....	48
Tabela 4.5: Rezultati razpoznavanja na podmnožici s5.....	50
Tabela 4.6: Tabela s primerjalnimi dolžinami značilnih vektorjev.....	52
Tabela 4.7: Tabela časovne zahtevnosti izračuna značilnih vektorjev.....	52