



# **VISOKA POSLOVNA ŠKOLA STRUKOVNIH STUDIJA ČAČAK**

## **SEMINARSKI RAD**

### **Detekcija lica Gaborovim filterom**

Mentor: \_\_\_\_\_  
Profesor: \_\_\_\_\_

Student: \_\_\_\_\_  
Br.Indeksa: \_\_\_\_\_

# Sadržaj:

## 1. Uvod

## 2. Gaborov filter

|                                                         |   |
|---------------------------------------------------------|---|
| 2.1. Kompleksni Gaborov filter u prostornoj domeni..... | 3 |
| 2.2. Kompleksni sinusoidni nositelj.....                | 3 |
| 2.3. Gaussova ovojnica.....                             | 4 |
| 2.4. Kompleksna Gaborova funkcija.....                  | 5 |
| 2.5. Implementacija Gaborovog filtra.....               | 7 |

## 3. Učenje

|                                    |   |
|------------------------------------|---|
| 3.1. Baza ulaznih uzorka.....      | 8 |
| 3.2. Učenje neuronskom mrežom..... | 8 |

## 4. Testiranje

|                      |    |
|----------------------|----|
| 4.1. Prva faza.....  | 11 |
| 4.2. Druga faza..... | 11 |
| 4.3. Treća faza..... | 12 |

## 5. Rezultati i problemi

|                     |    |
|---------------------|----|
| 5.1. Rezultati..... | 12 |
| 5.2. Problemi.....  | 14 |

## 6. Literatura.....15

# **1. Uvod**

Motivacija za ovaj rad nađena je u činjenici da je Gaborov filter matematička aproksimacija prostornog receptivnog polja jednostavne ćelije u primarnoj vizualnoj moždanoj kori nekih sisavaca (npr. mačke). Kako je detekcija lica vrlo popularan problem na području raspoznavanja uzoraka te je mnogo algoritama u optičaju, pokušaj je da se ovim radom približi računalna detekcija lica onoj biološkoj. Pri tome je Gaborov filter sredstvo kojim se ulazna slika prostorno i frekvencijski lokalizira tako da ulazni uzorak bude što je moguće bolje vremensko-frekvencijski određen tako da se može u  $n$ -dimenzionalnom prostoru uzoraka razlikovati od ostalih elemenata u slici odnosno da se može ispravno klasificirati.

S obzirom da je kvalitetna ekstrakcija značajki ključ za klasifikaciju algoritam je dobio ime upravo prema filteru koji se koristi za ekstrakciju, a ne prema načinu učenja (uči se neuronskom mrežom).

Algoritam je implementiran i testiran u Matlabu 7 te više služi kao testna nego praktična realizacija ponajviše zbog svoje vremenske (ne)efikasnosti. Kada bi algoritam htjeli izvesti u nekoj praktičnoj realizaciji koristili bi komercijalne DSP procesore ili kontrolere.

## 2. Gaborov filter

### 2.1. Kompleksni Gaborov filter u prostornoj domeni

Za Gaborov filter u prostornoj domeni vrijedi slijedeća jednačba:

$$g(x, y) = s(x, y)w_r(x, y) \quad (1)$$

gdje je  $s(x, y)$  kompleksna sinusoida, poznata i kao **nositelj** (engl. **carrier**),  $w_r(x, y)$  je 2-D funkcija Gaussovog oblika, poznata i kao **ovojnica** (engl. **envelope**).

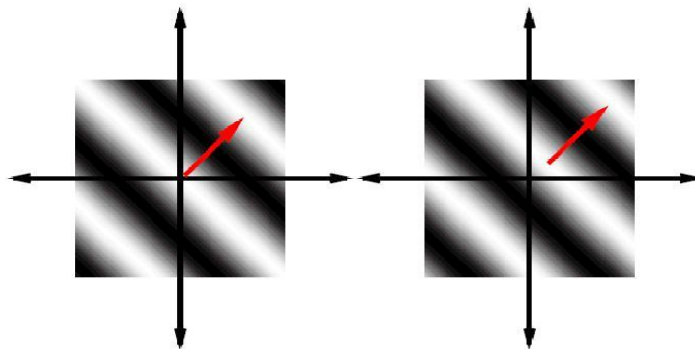
### 2.2. Kompleksni sinusoidni nositelj

Kompleksna sinusoida je definirana na način:

$$s(x, y) = e^{j(2\pi(u_0x+v_0y)+P)} \quad (2)$$

gdje  $(u_0, v_0)$  i  $P$  definiraju prostornu frekvenciju i fazu sinusoide, respektivno.

Sinusoidu možemo zamisliti kao dvije različite realne funkcije, prigodno smještene na realnom i imaginarnom dijelu kompleksne funkcije (slika 1).



Slika 1: Realni i imaginarni dio kompleksne sinusoide. Slike su 128x128 piksela. Parametri su  $u_0 = v_0 = 1/80$  ciklusa/piksel,  $P = 0^\circ$

Realni i imaginarni dio sinusoide su:

$$\begin{aligned}\operatorname{Re}(s(x, y)) &= \cos(2\pi(u_0x + v_0y) + P) \\ \operatorname{Im}(s(x, y)) &= \sin(2\pi(u_0x + v_0y) + P)\end{aligned}\quad (3)$$

Parametri  $u_0$  i  $v_0$  definiraju prostorne frekvencije sinusoide u kartezijevom sustavu. Prostorna frekvencija također može biti prikazana u polarnim koordinatama magnitudom  $F_0$  i smjerom  $w_0$ :

$$\begin{aligned}F_0 &= \sqrt{u_0^2 + v_0^2} \\ w_0 &= \arctg \frac{v_0}{u_0}\end{aligned}\quad (4)$$

te je:

$$\begin{aligned}u_0 &= F_0 \cos(w_0) \\ v_0 &= F_0 \sin(w_0)\end{aligned}\quad (5)$$

Koristeći ovu reprezentaciju, kompleksnu sinusoidu možemo zapisati u obliku:

$$s(x, y) = e^{j(2\pi F_0(x \cos w_0 + y \sin w_0) + P)} \quad (6)$$

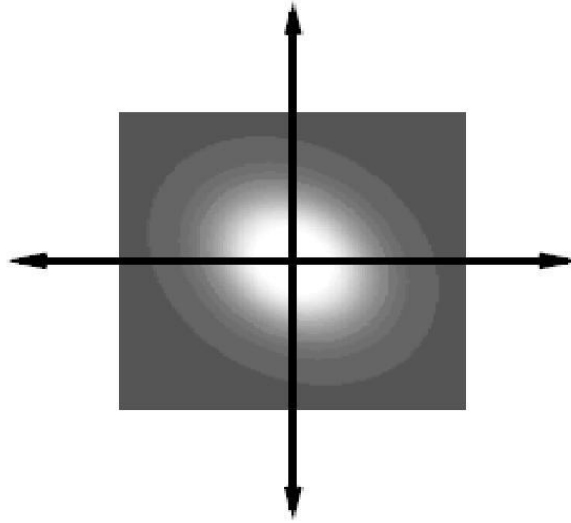
### 2.3. Gaussova ovojnica

Gaussova ovojnica je funkcija oblika:

$$w_r(x, y) = Ke^{-\pi(a^2(x-x_0)_r^2 + b^2(y-y_0)_r^2)} \quad (7)$$

gdje su  $(x_0, y_0)$  maksimumi funkcije, a i b su parametri skale Gaussove funkcije( povećanjem a i b ovojnica se smanjuje u prostornoj domeni), a  $r$  označava operaciju rotacije(u smjeru kazaljke na satu) takvu da :

$$\begin{aligned}(x-x_0)_r &= (x-x_0)\cos\theta + (y-y_0)\sin\theta \\ (y-y_0)_r &= -(x-x_0)\sin\theta + (y-y_0)\cos\theta\end{aligned}\quad (8)$$



Slika 2 : Gaussova ovojnica. Slika je 128x128 piksela. Parametri su :  $x_0 = y_0 = 0$ ,  $a = 1/50$ ,  $b = 1/40$ ,  $\Theta = -45^\circ$

## 2.4. Kompleksna Gaborova funkcija

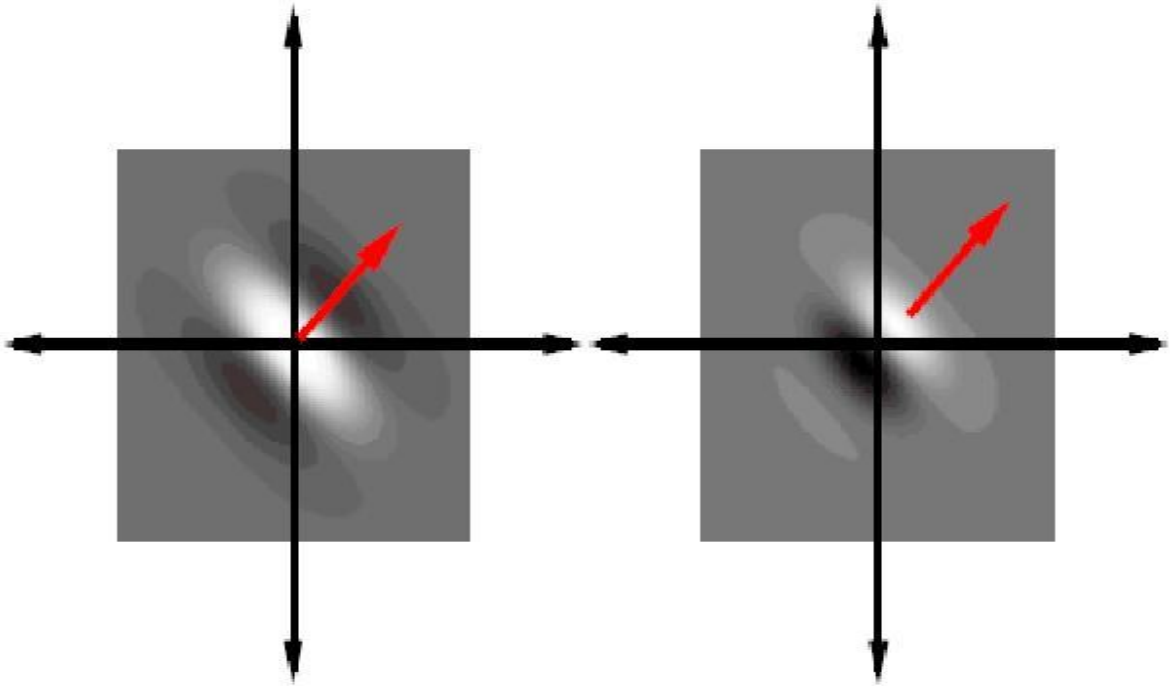
Kompleksna Gaborova funkcija je definirana sa slijedećih 9 parametara:

- $K$  : Skalira magnitudu Gaussove ovojnice
- $(a,b)$  : skaliranje dvaju osi Gaussove ovojnice
- $\theta$  : rotacijski kut Gaussove ovojnice
- $(x_0, y_0)$  : lokacija maksimuma Gaussove ovojnice
- $(u_0, v_0)$  : prostorne frekvencije nositelja u kartezijevim koordinatama
- $P$  : faza sinusoidalnog nositelja

Svaki kompleksni Gabor se sastoji od dvije funkcije u kvadraturi (s razlikom u fazi od  $90^\circ$ ), a locirani su na realnom i imaginarnom dijelu kompleksne funkcije.

Sada je izraz za kompleksnog Gabora u prostornoj domeni:

$$g(x, y) = Ke^{-\pi(a^2(x-x_0)_r^2 + b^2(y-y_0)_r^2)} e^{j(2\pi(u_0x+v_0y)+P)} \quad (9)$$



Slika 3: Realni i imaginarni dio kompleksne Gaborove funkcije u prostornoj domeni.

Slike su 128x128 piksela. Parametri su :  $x_0 = y_0 = 0$ ,  $a = 1/50$ ,  $b = 1/40$ ,  $\Theta = -45^\circ$ ,

$$F_0 = \frac{\sqrt{2}}{80}, w_0 = 45^\circ, P = 0^\circ$$

Odziv Gaborovog filtra u frekvencijskoj domeni dobivamo 2-D

Fourierovom transformacijom:

$$G(u, v) = \frac{K}{ab} e^{j[-2\pi(x_0(u-u_0)+y_0(v-v_0)+P)]} e^{-\pi\left[\frac{(u-u_0)_r^2}{a^2} + \frac{(v-v_0)_r^2}{b^2}\right]} \quad (10)$$

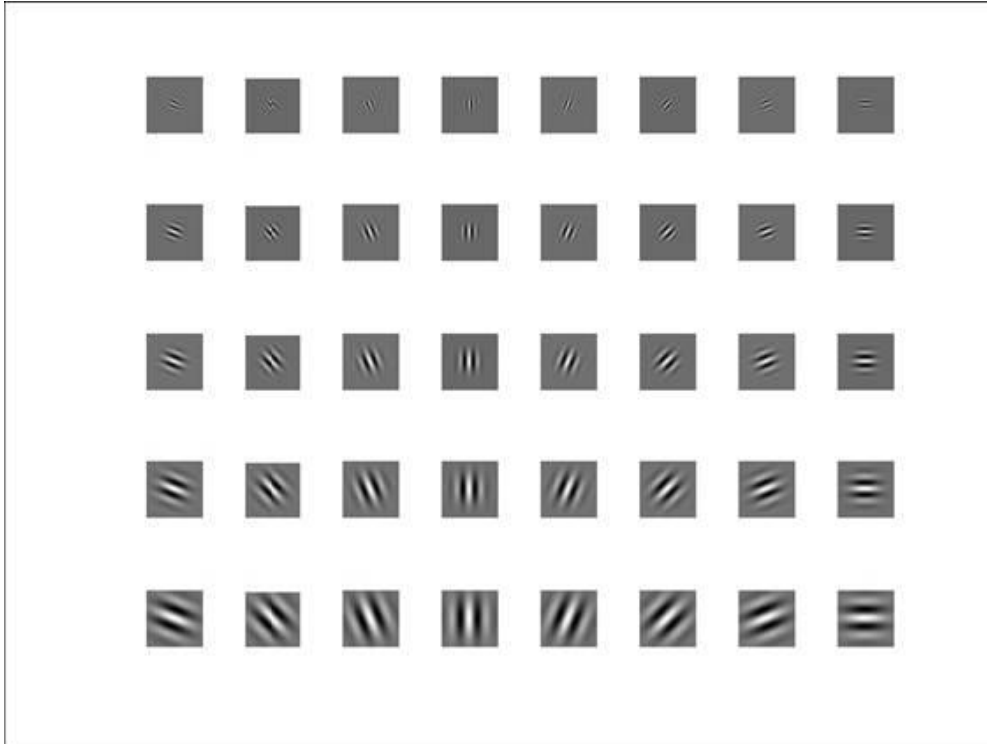
## 2.5. Implementacija Gaborovog filtra

Za implementaciju se koristi modificirani Gaborov filter čiji se izraz može pojednostaviti ako uzmemo da je  $x_0 = y_0 = 0$ ,  $P = 0$ ,  $a = b = \sigma$  (tada rotacija  $\Theta$  nema efekt) te ako uzmemo ograničenje magnitude prostorne frekvencije sinusoidalnog nositelja na  $F_0 = \sqrt{u_0^2 + v_0^2} = \frac{\sigma^2}{\sqrt{2\pi}}$ . Tada se izraz za Gaborov filter u prostornoj domeni pojednostavljuje na:

$$g(x, y) = 2\pi\sigma^2 e^{-\pi\sigma^2(x^2+y^2)} e^{j\sqrt{2\pi}\sigma^2(x\cos w_0 + y\sin w_0)} - e^{-\frac{\sigma^2}{2}} \quad (11)$$

Frekvencijski odziv se dobije fft-om gorenavedene funkcije.

U našoj smo implementaciji koristili pojednostavljenu verziju sa 5 različitih skala i 8 orijentacija. Rezultati su vidljivi na slici 4.



Slika 4. Gaborov filter u prostornoj domeni za 5 skala i 8 orijentacija



## 3. Učenje

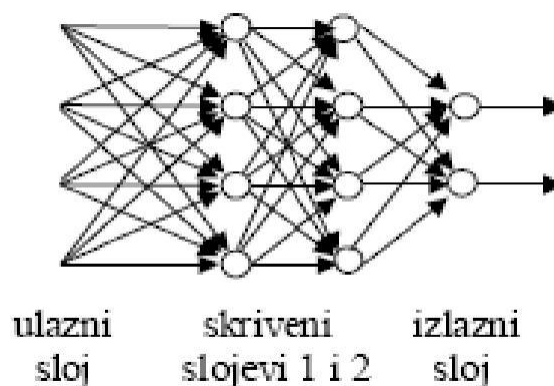
### 3.1. Baza ulaznih uzoraka

Baza ulaznih uzoraka se sastoji od lica i ne-lica veličine 27x18 piksela. S obzirom da je broj uzoraka relativno mali(oko 100) dodatni broj uzoraka se dobije zrcaljenjem slike oko obje osi, rotacijom slike u svim smjerovima za neki mali faktor. Iz jednog lica se dobije deset uzoraka, a iz ne-lica se dobiju četiri uzorka.

Iz slike veličine 27x18 se dobiva vektor dimenzije 2160 koji je dobiven konvolucijom Gaborovog filtra i slike(množenje u frekvencijskoj domeni) te decimacijom u frekvenciji za faktor 3 jer je vektor prevelike dimenzije, a ispravna klasifikacija se može obaviti sa puno manjom dimenzijom(dimenzija uzorka se smanjuje sa 19440 na 2160).

### 3.2. Učenje neuronskom mrežom

Neuronska mreža je alat kojim vršimo klasifikaciju segmenata slike na one koji sadrže lica i one koji ne sadrže lica. Mreža koja se koristi za učenje je bez povratnih veza (engl. **feedforward network**).Principijelna shema jedne takve mreže dana je na slici 5:



Slika 5: Principijelna shema višeslojne feedforward mreže

U našoj implementaciji se koristi mreža sa ulaznim slojem od 2160 neurona, jednim skrivenim slojem od 100 neurona i jednim izlaznim neuronom na koji se stavlja prag koji određuje je li segment slike lice.

Neuronska mreža se uči algoritmom s povratnom propagacijom pogreške (engl. **backpropagation, BP**). Ideja backpropagation algoritma je slijedeća:

- Signal pogreške na izlazu neurona  $j$  u koraku  $n$  je razlika između željenog i dobivenog odziva:

$$e_j(n) = d_j(n) - y_j(n); \quad j \dots \text{izlazni neuron} \quad (12)$$

- Trenutna kvadratna pogreška (u koraku  $n$ ) na izlazu mreže je:

$$E(n) = \sum_{j \in C} e_j^2(n); \quad C \dots \text{skup izlaznih neurona} \quad (13)$$

- Prosječna kvadratna pogreška za sve uzorke jednaka je:

$$E_{sr} = \sum_{n=1}^N E(n) \quad N \dots \text{broj uzoraka u skupu za učenje} \quad (14)$$

- Cilj učenja je odrediti težine tako da se minimizira srednja pogreška  $E_{sr}$

U osnovnoj verziji BP algoritma cilj se postiže modifikacijom LMS (engl. **least mean squares**) algoritma pri čemu se koristi gradijentni spust kao optimizacijski postupak za pronalaženje težina. Na kraju dolazimo do algoritma koji ima 2 prolaza:

#### 1. Prolaz unaprijed

- Težine se ne mijenjaju a izlazi neurona se računaju počevši od ulaznog sloja prema izlaznom sloju
- Signali se računaju prema izrazima za internu aktivnost i izlaz pojedinog neurona:

$$v_j(n) = \sum_{i=0}^p w_{ji}(n) y_i(n) \quad (15)$$

$$y_j(n) = \theta_j(v_j(n)) \quad (16)$$

gdje je  $\theta_j(v_j(n))$  najčešće sigmoidna aktivacijska funkcija:

$$\theta_j(v_j) = \frac{1}{1 + e^{-v_j}} \quad (17)$$

## 2. Povratni prolaz

- Započinje računanje od izlaznog sloja
- Signali pogreške se propagiraju od izlaznog sloja prema ulazu, sloj po sloj, rekursivno računajući gradijent  $\delta$  za svaki neuron
- Korekcija  $\Delta w_{ji}(n)$  za težinu koja ide od neurona **i** prema neuronu **j** se računa delta pravilom:

$$\Delta w_{ji}(n) = \eta \delta_j(n) y_i(n) \quad (18)$$

➤ Ako je neuron **j** izlazni neuron tada:

$$\delta_j(n) = \varphi_j'(v_j(n)) e_j(n) \quad (19)$$

➤ Ako je neuron **j** skriveni neuron tada:

$$\delta_j(n) = \varphi_j'(v_j(n)) \sum_k \delta_k(n) w_{kj}(n) \quad (20)$$

U Matlabu postoje varijacije učenja backpropagation algoritma, a odabrana je metoda konjugiranih gradijenata. Ona se razlikuje od metode gradijentnog spusta po tome što ima konjugirane smjerove (ortogonalni smjerovi koji u n-dimenzionalnom prostoru čine bazu) te se, u iterativnom postupku, gradijentni vektor projicira na najbliži mu vektor koji je konjugiran s obzirom na prethodne odabrane. Na taj način se dobiva rastav

funkcije po konjugiranim vektorima koji daje optimalno rješenje. Može se pokazati da je metoda konjugiranih gradijenata brža od gradijentnog spusta.

Uvjeti zaustavljanja algoritma su 400 iteracija ili srednja kvadratna pogreška(engl. **mean square error, MSE**)  $< 10^{-3}$ .

## **4. Testiranje**

### **4.1. Prva faza**

U prvoj fazi testiranja se slika mora segmentirati kako bi segmenti slike mogli biti obrađeni kao ulazni vektori uzoraka te identificirani kao lice ili ne-lice. Najveći je problem vrlo velik broj segmenata koje treba ispitati.

U implementaciji je to riješeno na način da se ulazna slika konvoluirana sa dva predloška u kojem se nalaze lica ispred svijetle i tamne pozadine te se na slici dobivaju korelacijske regije u obliku lokalnih maksimuma koje su nam potencijalna lica. Tako se rapidno smanjuje broj segmenata koje treba ispitati te se time smanjuje vrijeme pretraživanja slike.

### **4.2. Druga faza**

U drugoj fazi se ispituje svaka regija oko svakog piksela koji ima svojstvo lokalnog maksimuma, odnosno pokazuje veliku koreliranost sa predloškom. Iz tih se regija dobivaju ulazni uzorci koji se dovode na ulaz naučene neuronske mreže. Posebno je važan prag koji određuje je li izlaz neuronske mreže dovoljno velik da ulazni uzorak može biti okarakteriziran kao lice. Ako je izlaz dovoljno velik, oko lokalnog maksimuma se stvara mali krug maksimalne vrijednosti koji nam je ulaz u treću fazu, odnosno na morfološku obradu slike.

### **4.3. Treća faza**

Matrica odziva na mrežu se filtrira funkcijom praga te se na taj način dobiva binarna slika gdje su kružići lica koja trebamo identificirati.

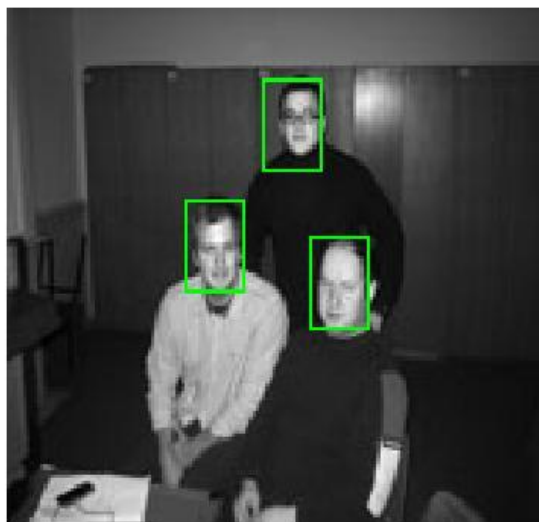
Zatim se na toj slici radi morfološka operacija dilatacije pomoću strukturnog elementa kruga jer je regija lica označena kružnim oblikom. Dilatacija je operacija koja uzima strukturni element kao masku i pomoću maske oko nekog piksela traži lokalni maksimum. Nakon te operacije se ističu oblici slični strukturnom elementu.

Nakon toga se radi labeliranje povezanih komponenti pomoću koje dobivamo regije lica. Nad komponentama se traže centroidi koji služe kao centri za crtanje pravokutnika. Nakon toga se oko centroida crta zeleni pravokutnik koji označava regiju lica te se „lijepi“ na originalnu sliku.

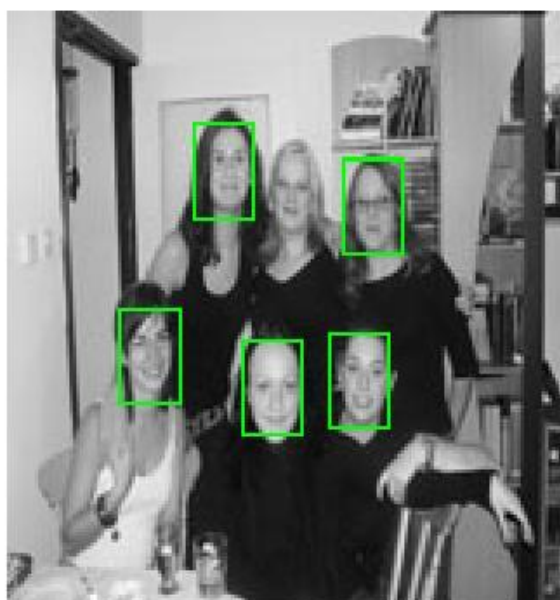
## **5. Rezultati i problemi**

### **5.1. Rezultati**

Rezultati testiranja pokazali su da je za neku sliku koja sadrži manje lica(3-5) vrijeme potrebno za detekciju reda veličine 10 sekundi što je dobro uzimajući u obzir da se testiranje obavlja sekvencijalno, a ne paralelno na potencijalnim regijama (Matlab ne podržava multithreading). Algoritam je pokazao visok stupanj točnosti(preko 90%) za slike koje su prilagođene bazi. Rezultati se mogu vidjeti na slikama:



Slika 6 i 7: slika za testiranje(dolje) i rezultati(gore)



Slika 8 i 9: slika za testiranje(dolje) i rezultati(gore)

## 5.2. Problemi

Osnovni problem leži u činjenici da Matlab alat za praktičnu realizaciju pa je algoritam ipak prilično spor. Što se više pokušava ubrzati algoritam, to je lošiji detection rate. Ako želimo poboljšati algoritam, definitivno je potrebno povećati bazu lica te napraviti piramidu veličina za segmentaciju slike pri čemu bi svaka veličina segmenta slike bila uvećana za neki konstantni parametar. Zbog nedostatka piramide, algoritam uočava lica veličine slične onima za učenje (slab je za velika i mala lica).

Sporost kod testiranja algoritma leži i u činjenici da se kritične regije obrađuju sekvencijalno. Neki oblik receptivnog polja ili multithreading u drugim programskim jezicima riješili bi problem.

Ulazne slike su 27x18 pa su lica izdužena na specifičan način. Zato se detekcija ne obavlja kvalitetno ako rezolucija nije takva da „proizvodi“ lica slična onima za učenje. To je problem u svim algoritmima, ali se oni najčešće rade za konstantne parametre dužine i širine objektiva kamere (rezolucija ne igra važnost kod oblikovanja lica) pa se algoritam uči za takvu sliku.

Elemente baze uzoraka za učenje, lica, još treba dodatno procesirati kako bi se izbjegao efekt da mreža uči oštre konture kao i efekt različitog izvora svjetla.



## **6. Literatura**

1. P. Dayan, L.F. Abbott: Theoretical Neuroscience, The MIT Press, Cambridge Massachusetts, 2001.
2. J. Movellan: Tutorial on Gabor Filters,  
<http://mplab.ucsd.edu/wordpress/tutorials/gabor.pdf>
3. Gabor Filters, <http://gdit.iiit.net/~arul/report/node13.html>
4. S. Lončarić: Predavanja iz kolegija „Neuronske mreže“  
<http://nm.zesoi.fer.hr/>

**Gotovi seminarski, maturski, maturalni i diplomski radovi iz raznih oblasti, lektire , puškice, tutorijali, referati** - specijalizovan tim za usluge visokokvalitetnog pisanja, istraživanja i obradu teksta za kompletan region Balkana.

Posetite nas na sajtovima ispod:

**[WWW.MATURSKIRADOVI.NET](http://WWW.MATURSKIRADOVI.NET)**

**[WWW.SEMINARSKIRAD.ORG](http://WWW.SEMINARSKIRAD.ORG)**

**[WWW.MATURSKI.NET](http://WWW.MATURSKI.NET)**

**[WWW.MATURSKI.ORG](http://WWW.MATURSKI.ORG)**

**[WWW.SEMINARSKIRAD.INFO](http://WWW.SEMINARSKIRAD.INFO)**

Dostupni smo Vam 24h 365 dana u godini.

Za gotove verzije rada obratiti se na mail:

**[maturskiradovi.net@gmail.com](mailto:maturskiradovi.net@gmail.com)**

**061/ 11-00-105**

Seminarski, diplomski, maturski radovi, prevodi na engleski i eseji...