

UNIVERZA V LJUBLJANI
FAKULTETA ZA MATEMATIKO IN FIZIKO
ODDELEK ZA FIZIKO
FIZIKA II. STOPNJA, RAČUNALNIŠKA FIZIKA

Žiga Zaplotnik

UPORABA ANALIZE NEODVISNIH KOMPONENT V
FIZIKALNIH APLIKACIJAH

Magistrsko delo

MENTOR: izred. prof. dr. Simon Širca

Ljubljana, 2014

Izjava o avtorstvu in dovoljenje za objavo elektronske oblike magistrskega dela

Podpisani *Žiga Zaplotnik* izjavljam:

- da sem magistrsko delo z naslovom *Uporaba analize neodvisnih komponent v fizikalnih aplikacijah* izdelal samostojno pod mentorstvom *izred. prof. dr. Simona Širce* in
- da Fakulteti za matematiko in fiziko Univerze v Ljubljani dovoljujem objavo elektronske oblike svojega dela na spletnih straneh.

Ljubljana,

Podpis,

Zahvala

Hvala mentorju, izred. prof. dr. Simonu Širci, za pomoč in svetovanje pri pisanju magistrskega dela. Zahvaljujem se mu za potrpežljivost ob mojem odlašanju z začetkom pisanja, za izjemno odzivnost, sproščenost in vso dobro voljo ob pogovorih o nadaljnjem delu – tako sva skupaj do novih idej še lažje prihajala. Hvala še enkrat. Upam, da ne zadnjič!

Hvala tudi kolegu fiziku, fotografu Oskarju Marku Musiću, za posojilo fotografskega materiala, ki sem ga potreboval pri magistrskem delu.

Hvala za pomoč, urejene zapiske in dolge pogovore sošolkama, meteorologinjama Veroniki Hladnik in Katji Kozjek, ki sta edini z menoj vseh pet let delili študijske klopi in v prvih vrstah bili boj tudi na drugi – včasih oblačni, deževni, a večinoma sončni meteorološki fronti.

Hvala vsem sošolcem in sošolkam – zaradi njih se na fakulteti že dolgo počutim kot doma.

Hvala prijateljem Janu Bohincu, Nejcu Davidoviču, Danielu Grošlju, Svetlani Nikič in bratu Luku za prijetno, smeha polno krajšanje mnogih ur vožnje iz Kranja v Ljubljano na fakulteto ter nazaj. Hvala jim tudi za potrpežljivost ob mojem nenehnem zamujanju.

Hvala puncu Jerci, ker mi je odkrila prej neznani, manj fizikalni in bolj čutni del življenja, s svojo sproščenostjo in igrivostjo pa polepšala tudi najbolj naporene študijske dneve.

Hvala družini: staršem, bratoma in stari mami, ker so me vedno spodbujali – ob negotovem začetku študija in pri nadaljnjih odločitvah ter mi nudili vse in še več, da sem se lahko študiju povsem posvetil.

Povzetek

Analiza neodvisnih komponent (ICA) je numerična metoda, ki iz izmerjene linearne mešanice signalov zgolj na podlagi njihovih statističnih lastnosti rekonstruira izvirne, statistično neodvisne signale. Opisano je teoretično ozadje metode, njena izpeljava in algoritem FastICA, ki metodo učinkovito implementira. Uporaba metode je ilustrirana s teoretičnim fizikalnim primerom merjenja električnega potenciala, ki ga ustvarjajo točkasti delci s časovno spremenljivim električnim nabojem.

Raziskana je tudi možna uporaba metode ICA pri odstranjevanju odbojev svetlobe od dielektričnih površin (stekla) iz zajetih fotografij. Tako zajeta slika je sestavljena iz dveh komponent: iz slike, ki nastane kot posledica odboja svetlobe od predmetov izza stekla in iz slike, ki nastane zaradi odboja žarkov na stekleni površini. S kombinacijo linearne polarizatorja in metode ICA uspešno ločimo sliko, ki jo ustvari delno polarizirana od stekla odbita svetloba, od slike, ki jo ustvari skozi steklo prepuščena nepolarizirana svetloba.

Obravnavano je še optično merjenje frekvence srčnega utripa. S spletno kamero s tremi barvnimi detektorji (R, G, B) merimo variacije od kože odbite svetlobe. Z metodo ICA iz meritev izluščimo iskan pletizmografski signal, ki je posledica krčenja in širjenja arterijskih žil tekom srčnega cikla. Frekvenca pletizmografskega signala ustreza frekvenci srčnega utripa, ki jo z našo metodo uspešno izmerimo.

Ključne besede: analiza neodvisnih komponent, statistična neodvisnost, algoritem FastICA, polarizacija svetlobe, Brewsterjev kot, poševni vpad svetlobe, dielektrična površina, fotopletizmografija, reflektivna spektroskopija, merjenje frekvence srčnega utripa.

PACS števila: 02.30.Zz, 02.50.Sk, 05.45.Tp, 07.07.Df, 42.25.Ja, 42.30.Sy.

Abstract

Independent component analysis (ICA) is a computational method, which is able to estimate independent source signals from a measured linear mixture of source signals. The reconstruction of source signals is based only on the statistical properties of the measured signals. In this master thesis, an idea and derivation of ICA method is provided, as well as FastICA algorithm, which efficiently implements ICA. The usage of ICA method is illustrated with an example of measuring electric potential generated by particles with time-varying electric charge.

The possible application of ICA is also investigated in cancelling reflections from dielectric surfaces (glass) in images. The image taken with the camera through a transparent glass is a superposition of two image components: a real image, which is a result of a light reflection from objects behind the glass and a virtual image, which is a result of light reflection from the glass surface. We show that with the combination of a linear polarizer and ICA method we can successfully separate the image created by unpolarized light that is transmitted through glass and the image created by partly polarized light that is reflected from glass surface.

We also research the remote optical measuring of heart rate. We measure the variations in light reflected from skin with webcam consisting of three (R, G, B) color sensors. We show that with ICA, we are able to extract from those measurements the plethysmographic signal, which is generated by the volumetric changes of arterial blood vessels during the cardiac cycle. The frequency of the plethysmographic signal corresponds to the heart rate frequency.

Keywords: independent component analysis, statistical independence, FastICA algorithm, polarization, Brewster angle, poševni vpad svetlobe, dielectric surface, photoplethysmography, reflective spectroscopy, heart rate measurements.

PACS numbers: 02.30.Zz, 02.50.Sk, 05.45.Tp, 07.07.Df, 42.25.Ja, 42.30.Sy.

Kazalo

1	Uvod	13
1.1	Motivacijski primer	14
2	Analiza neodvisnih komponent	20
2.1	Definicija	20
2.2	Ocena separacijske matrike in negaussovskost	22
2.3	Algoritem FastICA	24
3	Odstranjevanje odbojev od dielektričnih površin iz fotografij	27
3.1	Opis problema	27
3.1.1	Poševen vpad nepolarizirane svetlobe na dielektrično površino	28
3.1.2	Modelski nastavek	31
3.1.3	Uporaba metode analize neodvisnih komponent	32
3.2	Rezultati	33
3.2.1	Ločevanje linearne kombinacije črnobelih slik	33
3.2.2	Odstranjevanje odbojev od dielektričnih površin	34
4	Merjenje fizioloških signalov	39
4.1	Merjenje frekvence srčnega utripa s spletno kamero – opis	39
4.2	Način merjenja pletizmografskega signala	41
4.3	Rezultati meritev in analiza	43
5	Zaključek	47
	Literatura	49
A	Prepoznavanje objektov na sliki	53
A.1	Haarove poteze	54
A.2	Učenje klasifikatorjev	56

A.3 Kaskada klasifikatorjev	58
B Popravek gama	60
C Preslikava gostote svetlobnega toka v relativno svetlost	62

1 Uvod

Fizikalni zakoni in teorije so podlaga za izdelavo matematičnih modelov, s katerimi poenostavljeno opišemo procese v naravi. S temi modeli lahko na primer napovemo rezultate meritev opazovanih količin ali preučujemo vplive različnih dejavnikov na rezultate meritev. Postopku pravimo modeliranje oziroma simuliranje nekega procesa (ang. *forward problem*). Obratni postopek pa imenujemo inverzni problem. Pri slednjem na podlagi meritev opazovanih količin skušamo rekonstruirati model, t.j. poiskati parametre, ki model natančno določajo.

Slikovita analogija modelskemu problemu je v neki situaciji poiskati pravi odgovor na vprašanje, inverzni problem pa bi bil glede na neki znan odgovor poiskati primerno vprašanje. Seveda ima slednji več kot le eno rešitev. Pomembna razlika med (determinističnim) modelskim problemom in inverznim problemom je torej v tem, da ima prvi enolično rešitev pri nekih vhodnih podatkih, zadnji pa ima neskončno mnogo rešitev. To pokažemo na preprostem primeru težnostnega polja Zemlje [1]. Če poznamo porazdelitev gostote $\rho(\mathbf{r})$ v Zemljini notranjosti, lahko točno in enolično izračunamo velikost težnostnega pospeška g kjerkoli na njenem površju. Inverzni problem, ki je pogost tudi v praksi, pa je, kako na podlagi meritev težnostnega polja na površini Zemlje izračunati porazdelitev gostote v njeni notranjosti. Število rešitev za porazdelitev gostote je neskončno – obstajajo namreč različne porazdelitve gostote, ki vse dajo enako težnostno polje na površini. Inverzni problemi običajno niso analitično rešljivi, zato se moramo zateči k numeričnim metodam. Ti problemi so pogosto tudi računsko zelo zahtevni – za dobro oceno modelskih parametrov namreč potrebujemo veliko množico vhodnih podatkov.

V magistrskem delu se osredotočimo na reševanje posebne vrste inverznih problemov, pri katerih ni znan niti fizikalni model niti njegovi parametri. Vsa informacija, ki jo imamo, je torej skupek meritev. Obravnavamo linearne inverzne probleme oblike

$$\mathbf{x} = \mathbf{A}\mathbf{s}, \quad (1.1)$$

kjer je $\mathbf{x}$ vektor znanih meritev, $\mathbf{A}$ neznana matrika, $\mathbf{s}$ pa je vektor neznanih izvirov (vhodnih podatkov modela). Enačba (1.1) torej opredeljuje linearni model, ki predstavlja model preslikave $\mathbf{s}$ v $\mathbf{x}$. Matriki $\mathbf{A}$ včasih rečemo tudi model oz. mešalna matrika (ang. *mixing matrix*). Poleg teh obravnavamo še posplošene nelinearne inverzne probleme oblike

$$\mathbf{x} = f(\mathbf{A}\mathbf{s}), \quad (1.2)$$

kjer je f neznana nelinearna funkcija. V obeh primerih želimo oceniti tako linearen model $\mathbf{A}$ kot tudi izvire $\mathbf{s}$, kljub temu da nimamo nobenih informacij niti o enem niti o drugem. Tako zastavljen problem seveda nima enolične rešitve. Ob dodatnih predpostavkah o statističnih lastnostih izvirov pa rešitev vendarle lahko poiščemo. Že na začetku poudarimo, da ta rešitev – kot se za inverzne probleme spodobi – ni enolična.

Numerična metoda, s katero se v magistrskem delu spopadamo z inverznimi problemi oblike (1.1) in (1.2), je analiza neodvisnih komponent (ang. *independent component analysis*, ICA) [2 – 4]. Ime metode nakazuje, da jo lahko uporabimo ob privzetku, da so izviri med seboj paroma statistično neodvisni. Analizo neodvisnih komponent uvrščamo tudi med metode strojnega učenja (ang. *machine learning*), saj model $\mathbf{A}$ "naučimo" s podatki meritev $\mathbf{x}$. Na področju obdelave signalov (ang. *signal processing*) jo uvrščajo med metode za slepo ločevanje izvirov (ang. *blind source separation*). Podatke o posamezni meritvi x_i lahko namreč razumemo tudi kot časovno odvisen signal $x_i(t)$. Izraz "slepo" izhaja iz dejstva, da nimamo nobene informacije o izvirnih signalih $\mathbf{s}$ niti o modelu $\mathbf{A}$, s katero bi si lahko pomagali pri rekonstrukciji izvirnih signalov.

V magistrskem delu rešujemo torej inverzne probleme oblike (1.1) in (1.2) z uporabo metode ICA. Ti problemi se pojavljajo na mnogih področjih fizike, npr. v akustiki, linearni optiki, spektroskopiji itd. Vseeno pa je za reševanje teh problemov metoda ICA le redko uporabljena, kar je bilo tudi glavna motivacija za to magistrsko delo. Izjema so inverzni problemi v medicinski fiziki in nekaterih področjih geofizike, kjer je metoda ICA že uveljavljena.

V magnetni encefalografiji (MEG), tehniki, s katero spremljamo možgansko aktivnost prek merjenja magnetnega polja, ki ga ustvarjajo šibki električni tokovi v možganih, se metoda uporablja za odpravljanje motenj ali artefaktov v nevromagnetnih signalih, ki jih povzročata mežikanje in premikanje oči [5]. Poleg tega se uporablja še v elektroencefalografiji (EEG) [6], elektrokardiografiji (ECG) [7], funkcionalni magnetni resonanci (fMRI) [8], v elektromiografiji (sEMG) [9] in elektrogastrografiji (EGG) [10].

V geofiziki se metoda ICA uporablja za analizo seizmoloških signalov, npr. pri odkrivanju izvirov tresenja tal in določanju lokacije teh izvirov [11]. V fiziki se pogosto uporablja tudi za odkrivanje, identifikacijo in ločevanje fizikalnih vzrokov za variabilnost nekega dinamičnega sistema – torej, kadar dinamike nekega sistema ne znamo pojasniti, saj je ta lahko posledica številnih fizikalnih procesov [12].

Matematični opis analize neodvisnih komponent podamo v naslednjem poglavju, najprej pa kot motivacijski primer opišimo preprost fizikalni problem, ki ga s to metodo lahko rešimo.

1.1 Motivacijski primer

Zamislimo si sledeč problem iz teorije električnega polja. V ravnini naj bo razporejenih n točkastih delcev s časovno spremenljivim električnim nabojem $e_j = e_j(t)$. Predpostavljamo, da se delci nahajajo na točno določenih pozicijah $\mathbf{r}_j$, ki se s časom ne spreminjajo, torej $\mathbf{r}_j \neq \mathbf{r}_j(t)$.

Ti nabiti delci ustvarjajo električno polje. Skupno električno polje v neki točki $\mathbf{R}_i$ je superpozicija električnih polj $\mathbf{E}_j$, ki bi jih ustvarili posamezni naboji e_j :

$$\mathbf{E}(\mathbf{R}_i, t) = \sum_{j=1}^n \mathbf{E}_j(\mathbf{r}_j, t) = \frac{1}{4\pi\epsilon_0} \sum_{j=1}^n \frac{e_j(t)}{r_{ij}^3} \mathbf{r}_{ij}, \quad (1.3)$$

kjer je $\mathbf{r}_{ij} = \mathbf{R}_i - \mathbf{r}_j$ vektor razdalje od točke v polju do naboja in $r_{ij} = |\mathbf{r}_{ij}|$ njena dolžina. Še raje kot z električnim poljem, ki je vektorska količina, delamo z električnim (coulombskim) potencialom $U = -\oint_C \mathbf{E} \cdot d\mathbf{r}$. Skupni električni potencial v neki točki $\mathbf{R}_i$ zopet zapišemo kot linearno kombinacijo potencialov U_j , ki bi jih povzročili posamezni naboji e_j :

$$U(\mathbf{R}_i, t) = U_i(t) = \sum_{j=1}^n U_j(\mathbf{r}_j, t) = \frac{1}{4\pi\epsilon_0} \sum_{j=1}^n \frac{e_j(t)}{r_{ij}}. \quad (1.4)$$

Električne naboje si torej lahko predstavljamo kot "izvire" električnega potenciala.

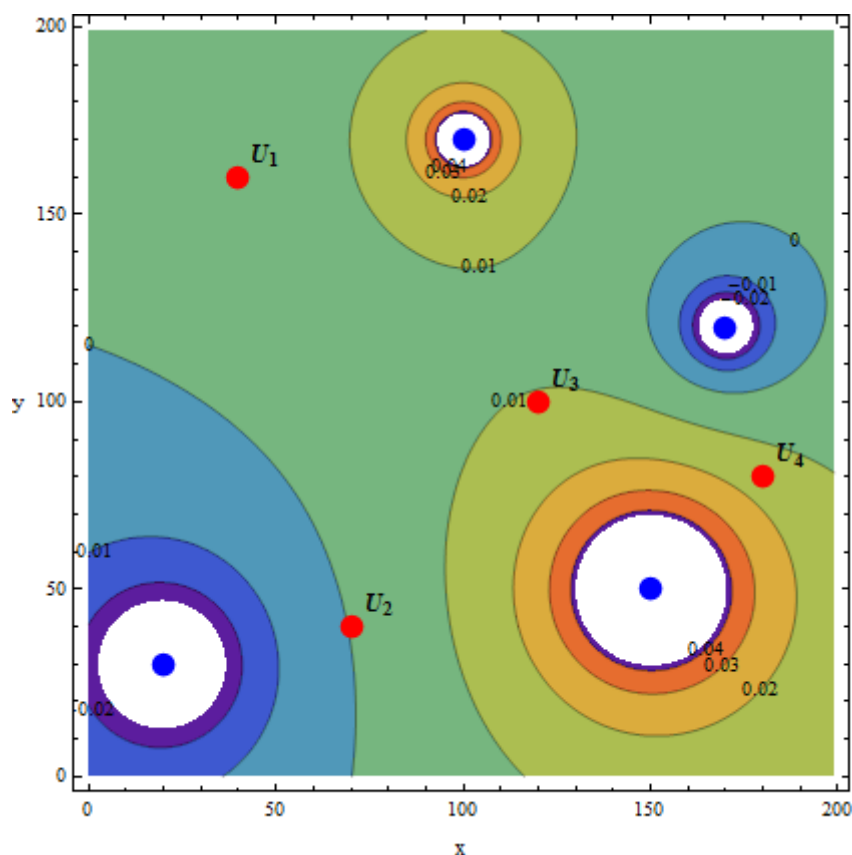
Predstavljajmo si sedaj, da v nekaj točkah, ki ne sovpadajo s položaji nabojev, izmerimo časovno spreminjajoč se električni potencial – ta je kar enak napetosti med to točko in neko referenčno točko, npr. ozemljeno točko. Zanima nas, ali lahko iz teh meritev ugotovimo časovni potek posameznih nabojev, torej "izvirov", ki ustvarjajo električno polje. Izkaže se, da ob določenih predpostavkah, t.j. da so naboji posameznih delcev statistično neodvisni ter da je število meritev večje ali enako številu izvirov, opisani inverzni problem lahko rešimo z metodo analize neodvisnih komponent.

Število nabitih delcev v ravnini in število meritev električnega potenciala je v našem primeru $n = 4$ (slika 1.1). Konstanto $1/4\pi\epsilon_0$ zaradi boljše preglednosti postavimo na 1. Enačbo (1.4) zapišemo v matrični obliki kot

$$\begin{bmatrix} U_1(t) \\ U_2(t) \\ U_3(t) \\ U_4(t) \end{bmatrix} = \begin{bmatrix} r_{11}^{-1} & r_{12}^{-1} & r_{13}^{-1} & r_{14}^{-1} \\ r_{21}^{-1} & r_{22}^{-1} & r_{23}^{-1} & r_{24}^{-1} \\ r_{31}^{-1} & r_{32}^{-1} & r_{33}^{-1} & r_{34}^{-1} \\ r_{41}^{-1} & r_{42}^{-1} & r_{43}^{-1} & r_{44}^{-1} \end{bmatrix} \cdot \begin{bmatrix} e_1(t) \\ e_2(t) \\ e_3(t) \\ e_4(t) \end{bmatrix}. \quad (1.5)$$

Točke $\mathbf{R}_i$, v katerih merimo električni potencial U_i , poznamo, ne pa tudi položajev $\mathbf{r}_j$ nabitih delcev. Imamo torej inverzni problem oblike (1.1), kjer je vektor meritev $\mathbf{x}(t) = [U_i(t)]^T$, vektor izvirov $\mathbf{s}(t) = [e_j(t)]^T$ in $\mathbf{A} = [r_{ij}^{-1}]$ od časa neodvisen model. Matrike $\mathbf{A}$ ne poznamo vnaprej, saj ne poznamo razdalj r_{ij} . Električni potencial U_i merimo ob diskretnih časih $t_N = N\Delta t$, kjer je $N = 0, 1, \dots, 1000$ in $\Delta t = 0.1$. Meritve potencialov $U_{i=1, \dots, 4}(t)$ so prikazane na sliki 1.2. Nabor podatkov za vsako meritev U_i je tako dovolj velik, da lahko izračunamo statistične lastnosti meritev, ki jih potrebujemo za oceno statističnih lastnosti izvirov.

Izvire sedaj poiščemo tako, da napravimo analizo neodvisnih komponent za vektor meritev $\mathbf{x}(t)$. Rezultat tega izračuna sta model oz. mešalna matrika $\mathbf{A}$ in vektor izvirov $\mathbf{s}(t)$, ki pa nista

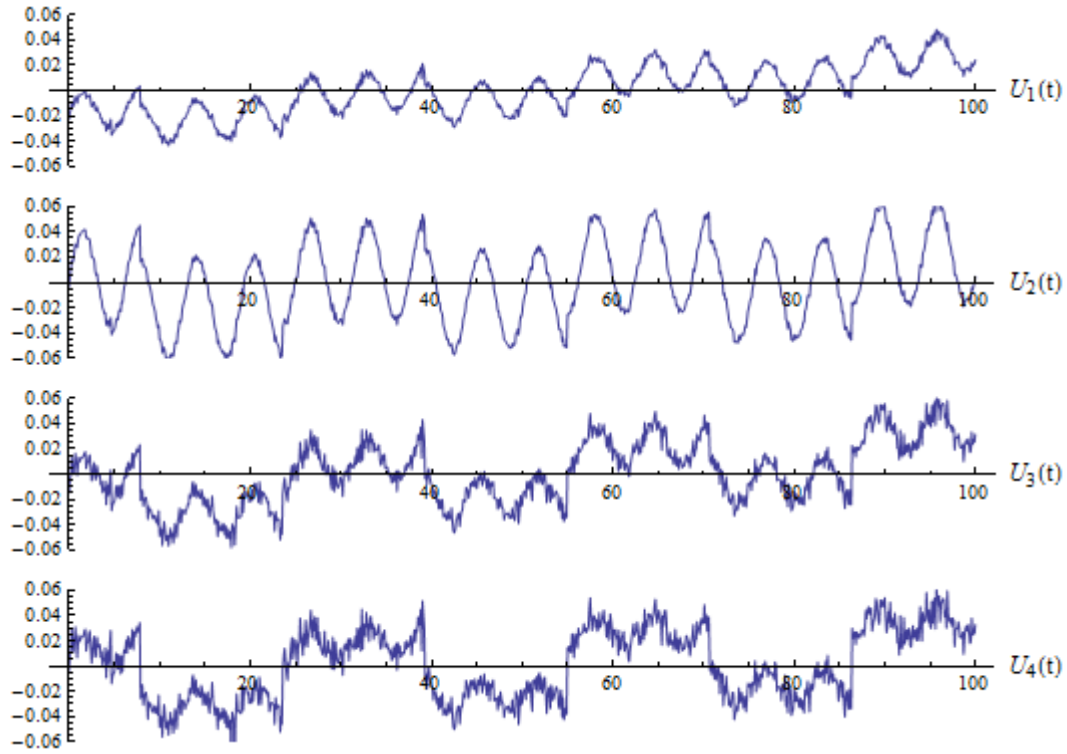


Slika 1.1. Električni potencial $U(x, y)$, ki ga ustvarjajo nabiti delci (modri krogi) v ravnini; ob nekem času t . Naboj na naibitih delcih se s časom spreminja. Pozicij nabitih delcev, izvirov električnega polja/potenciala, ne poznamo vnaprej. Električni potencial merimo v štirih točkah (rdeči krogi), katerih pozicije poznamo. Potencial merimo ob diskretnih časih, na vsaki izmed točk dobimo torej kot meritev časovno spreminjanje potenciala $U(t)$.

enolično določena. Ker namreč lahko mešalno matriko in vektor izvirov reskaliramo,

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \cdot \begin{bmatrix} s_1(t) \\ s_2(t) \\ s_3(t) \\ s_4(t) \end{bmatrix} = \begin{bmatrix} \frac{a_{11}}{\alpha} & \frac{a_{12}}{\beta} & \frac{a_{13}}{\gamma} & \frac{a_{14}}{\delta} \\ \frac{a_{21}}{\alpha} & \frac{a_{22}}{\beta} & \frac{a_{23}}{\gamma} & \frac{a_{24}}{\delta} \\ \frac{a_{31}}{\alpha} & \frac{a_{32}}{\beta} & \frac{a_{33}}{\gamma} & \frac{a_{34}}{\delta} \\ \frac{a_{41}}{\alpha} & \frac{a_{42}}{\beta} & \frac{a_{43}}{\gamma} & \frac{a_{44}}{\delta} \end{bmatrix} \cdot \begin{bmatrix} s_1(t)\alpha \\ s_2(t)\beta \\ s_3(t)\gamma \\ s_4(t)\delta \end{bmatrix}, \quad (1.6)$$

pomeni, da predznak in velikost matričnih elementov oz. izvirov nista enolično določena. V obravnavanem primeru zahtevamo, da so vsi matrični elementi pozitivni, saj razdalje r_{ij}^{-1} ne morejo biti negativne. Ker velja (1.6), lahko vsak negativen stolpec j v matriki $\mathbf{A}$ in pripadajoč izvir s_j pomnožimo z -1 in s tem zadostimo pogoju $a_{ij} > 0$.



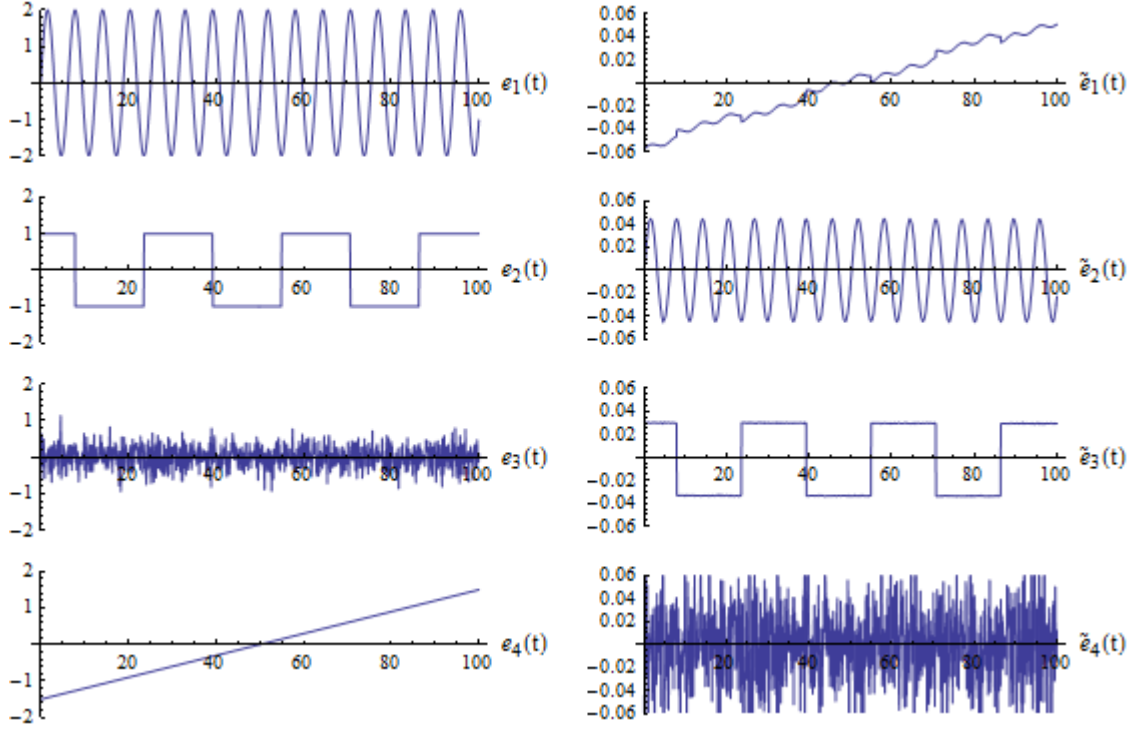
Slika 1.2. Meritve električnega potenciala v štirih točkah v ravnini. Položaji posameznih točk so označeni na sliki 1.1.

Tudi vrstni red ocenjenih izvirov ni enolično določen. Matrika $\mathbf{A}$ je nedoločena do poljubne permutacije natanko, npr.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \cdot \begin{bmatrix} s_1(t) \\ s_2(t) \\ s_3(t) \\ s_4(t) \end{bmatrix} = \begin{bmatrix} a_{14} & a_{11} & a_{12} & a_{13} \\ a_{24} & a_{21} & a_{22} & a_{23} \\ a_{34} & a_{31} & a_{32} & a_{33} \\ a_{44} & a_{41} & a_{42} & a_{43} \end{bmatrix} \cdot \begin{bmatrix} s_4(t) \\ s_1(t) \\ s_2(t) \\ s_3(t) \end{bmatrix}. \quad (1.7)$$

Slika 1.3 prikazuje z metodo ICA ocenjeno časovno spreminjanje električnega naboja $\tilde{e}_j(t)$ na delcih. Rezultate primerjamo z dejanskim časovnim spreminjanjem nabojev $e_j(t)$. Razen linearno naraščajočega naboja so oblike vseh izvirnih signalov dobro ocenjene. Vrstni red ocenjenih izvirov potenciala je dejansko permutiran po (1.7), amplitude izvirov pa so bistveno manjše od dejanskih. S samo metodo ICA torej lahko rekonstruiramo le oblike signalov.

Zanima nas še, ali bi lahko kako ocenili tudi dejanske amplitude izvirnih signalov (faktorje $\alpha, \beta, \gamma, \delta$) in celo njihove položaje. Izkaže se, da s samo metodo ICA tega ne moremo narediti, vseeno pa z metodo dobimo ključno informacijo – mešalno matriko $\mathbf{A}$. Ključno je tudi poznavanje točk $\mathbf{R}_i$ meritev potenciala. Iz (1.5) in (1.6) sledi sistem enačb, iz katerega lahko določimo tako položaj vsakega posameznega nabitega delca kot tudi dejansko velikost električnega naboja na njem. Za primer prvega nabitega delca zapišemo



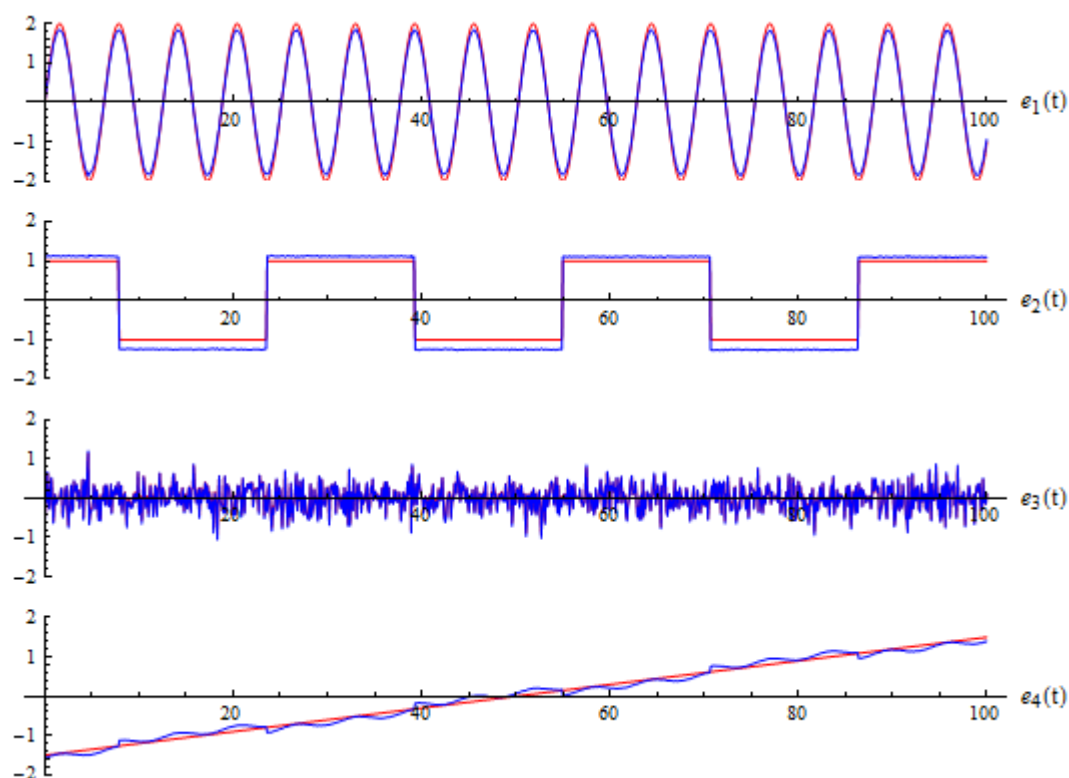
Slika 1.3. [LEVO] Dejansko časovno spreminjanje naboja na delcih; $e_1(t) = 2 \sin(t)$, $e_2(t) = \text{sign}\left(\cos\left(\frac{t}{5}\right)\right)$, $e_3(t) = 0.3 N(0,1)$ in $e_4(t) = 1.5 \left(-1 + 2 \left(\frac{t}{T}\right)\right)$, kjer $T = 100$. [DESNO] Ocenjene vrednosti časovnega spreminjanja naboja $\tilde{e}_j(t)$, ki jih dobimo iz meritev na sliki 1.2 z metodo ICA. Ocenjeni signali so reskalirani (1.6) tako, da so njihove variance enake 1 (razlog za to je razložen v razdelku 2.1). Amplitude ocenjenih signalov so zato precej nižje od amplitud prvotnih signalov.

$$\begin{aligned}
 r_{11} &= |\mathbf{R}_1 - \mathbf{r}_1| = a_{11}^{-1} \alpha, \\
 r_{21} &= |\mathbf{R}_2 - \mathbf{r}_1| = a_{21}^{-1} \alpha, \\
 r_{31} &= |\mathbf{R}_3 - \mathbf{r}_1| = a_{31}^{-1} \alpha, \\
 r_{41} &= |\mathbf{R}_4 - \mathbf{r}_1| = a_{41}^{-1} \alpha.
 \end{aligned} \tag{1.8}$$

To je sistem štirih enačb s tremi neznankami: skalarjem α in koordinatama $\mathbf{r}_1 = (x_1, y_1)$ prvega delca. Koeficiente a_{ij} dobimo iz mešalne matrike, ki smo jo izračunali z metodo ICA. Točke meritev $\mathbf{R}_i = (X_i, Y_i)$ pa tudi poznamo. Sistem rešimo numerično s katero od minimizacijskih metod, npr. Powellovo minimizacijsko metodo [13]. Minimizarati želimo cenilno funkcijo

$$f(x_1, y_1, \alpha) = \sum_{i=1}^n \left(\sqrt{(X_i - x_1)^2 + (Y_i - y_1)^2} - a_{i1}^{-1} \alpha \right)^2. \tag{1.9}$$

Povsem enako napravimo minimizacijo še za ostale nabite delce. Tako smo torej rekonstruirali položaje $\mathbf{r}_j$ nabitih delcev in poiskali faktorje $\alpha, \beta, \gamma, \delta$, ki točno določajo časovno spreminjanje



Slika 1.4. Končna ocena naboja na delcih $\hat{e}_j(t)$ (modro) v primerjavi z dejanskim nabojem $e_j(t)$ (rdeče). Ujemanje ocenjenih in dejanskih vrednosti je zelo dobro.

naboja na delcih $\hat{e}_j(t)$. Rezultati so prikazani na sliki 1.4. Ker za reševanje sistema (1.8) potrebujemo najmanj toliko enačb, kot je neznakov, sledi, da mora biti pri ravninskem (2D) problemu število meritev in izvirov $n \geq 3$, pri prostorskem (3D) problemu pa $n \geq 4$, če želimo popolnoma rekonstruirati časovno spreminjanje električnega naboja na delcih.

Obravnavani primer iz teorije električnega polja je različica slavnega (ang.) *cocktail-party* problema¹. Pokazali smo, da je iz meritev x_i , $i = 1, 2, \dots, n$, ki so linearna kombinacija neznanih izvirov, z metodo ICA mogoča rekonstrukcija enakega števila izvirov s_j , $j = 1, 2, \dots, n$, do nekega faktorja natančno. To pomeni, da lahko ugotovimo oblike izvirov, ne pa tudi njihovih amplitud. Informacijo, ki jo dobimo iz mešalne matrike $\mathbf{A}$, pa lahko nadalje uporabimo, da s pomočjo drugih numeričnih metod ocenimo še dejansko velikost signala. V motivacijskem primeru smo torej pokazali, da ima metoda ICA v kombinaciji z drugimi numeričnimi inverznimi metodami izjemen potencial pri problemih rekonstrukcije neznanih izvirov.

¹ *Cocktail-party* problem je problem ločevanja izvirov zvoka. Primer je množica hkrati govorečih ljudi v isti sobi (npr. na zabavi – odtod prihaja tudi ime problema). Poslušalec pri tem sliši mešanico zvokov iz različnih izvirov, želi pa se osredotočiti le na posamezen izvir, torej na govor določene osebe. Človeški možgani in sistem za zaznavanje sluha so tako izpopolnjeni, da lahko ločijo posamezne izvirne signale. Pri tem pa se postavi vprašanje, ali lahko te izvire ločimo tudi z digitalno obdelavo. Problem je prvi analiziral Colin Cherry [14].

2 Analiza neodvisnih komponent

Analiza neodvisnih komponent (ICA) je statistična numerična metoda, s katero v splošnem rešujemo linearne inverzne probleme oblike (1.1). Klasična analiza neodvisnih komponent obravnava take vrste problemov, uspešna pa je tudi pri problemih oblike (1.2), kadar je f zvezna šibko nelinearna funkcija. Poleg tega obstaja še mnogo drugih različic metode ICA: poleg klasične ICA metode je najbolj uporabljena jedrna analiza neodvisnih komponent (ang. *Kernel Independent Component Analysis*, KICA [15]), ki pa je v magistrskem delu ne obravnavamo.

2.1 Definicija

Klasična analiza neodvisnih komponent je statična linearna brezšumna analiza neodvisnih komponent. Predpostavimo, da vsak vektor koreliranih meritev $\mathbf{x} = (x_1, x_2, \dots, x_m)^T$ nastane z linearnim mešanjem neodvisnih izvirov $\mathbf{s} = (s_1, s_2, \dots, s_n)^T$. Velja torej

$$\mathbf{x} = \mathbf{A}\mathbf{s}, \quad (2.1)$$

kjer je $\mathbf{A} \in \mathbb{R}^{m \times n}$ mešalna matrika, ki je neodvisna od časa ($\mathbf{A} \neq \mathbf{A}(t)$); odtod poimenovanje "statična" ICA). Zgornjo enačbo lahko zapišemo tudi kot

$$\mathbf{x} = \sum_{j=1}^n \mathbf{a}_j s_j, \quad (2.2)$$

kjer je $\mathbf{a}_j = (a_{1j}, a_{2j}, \dots, a_{mj})^T$ j -ti stolpec matrike $\mathbf{A}$. Koeficienti a_{ij} v mešalni matriki določajo, kako močno je zastopan nek izvir s_j v meritvi x_i .

V modelu (2.1) privzamemo, da je mešanje izvirov z mešalno matriko $\mathbf{A}$ takojšnje (ang. *instantaneous*). To pomeni, da ni dovoljen zamik med časi, pri katerih posamezni izviri dosežejo isti senzor/merilec. Natančneje, če predpostavimo, da je frekvenca vzorčenja merilca ν , potem to pomeni, da se mora informacija iz izvira do vseh senzorjev razširiti vsaj v času $t_0 = 1/\nu$. V primeru z električnim poljem (razdelek 1.1) je bilo takojšnje mešanje zagotovljeno,

saj se električno polje vzpostavi po vsej domeni s svetlobno hitrostjo. Predpostavimo tudi, da je število izvirov manjše ali enako številu meritev, $n \leq m$. Napake meritev, ki bi jih lahko dodali na desni strani enačbe (2.1), zanemarimo oz. privzamemo, da je merski šum vsebovan že v izviri $\mathbf{s}$ (odtod ime "brezšumna", ang. *noiseless*, ICA) [16]. Prav tako predpostavimo, da so izviri statistično neodvisni. To je tudi v realnih problemih zelo verjetna predpostavka. Neodvisnost izvirov matematično formalno izrazimo z izjavo, da je njihova skupna verjetnostna gostota produkt posameznih verjetnostnih gostot, $p(s_1, s_2, \dots, s_n) = p(s_1)p(s_2) \cdot \dots \cdot p(s_n)$. Zato velja tudi za pričakovane vrednosti $\langle f(s_i)g(s_j) \rangle = \langle f(s_i) \rangle \langle g(s_j) \rangle$ pri poljubnih funkcijah f in g . Statistična neodvisnost je višja zahteva kot nekoreliranost – par statistično neodvisnih spremenljivk $(s_i, s_{j \neq i})$ je nekoreliran in ima kovarianco $\text{cov}(s_i, s_j) = \langle s_i s_j \rangle - \langle s_i \rangle \langle s_j \rangle = 0$, obratno pa ni nujno res. Verjetnostna porazdelitev posameznega izvira s_j (neodvisne komponente) je sicer lahko kakršnakoli, razen gaussovska. To je bistveni pogoj, da je rešitev (2.1) z metodo ICA mogoča. V nasprotnem primeru lahko mešalno matriko ocenimo le do ortogonalne transformacije natančno. Gaussovska je lahko največ ena komponenta [4].

Nadalje privzamemo, da je povprečje vektorja izvirov enako nič, $\langle \mathbf{s} \rangle = \mathbf{0}$. Da bo to držalo, moramo tudi povprečje vektorja meritev popraviti na nič. Zato napravimo transformacijo $\mathbf{x} \leftarrow \mathbf{x} - \langle \mathbf{x} \rangle$. Zaradi statistične neodvisnosti izvirov je kovariančna matrika izvirov diagonalna,

$$\Sigma_{ss} = \text{cov}(\mathbf{s}, \mathbf{s}) = \mathbf{s} \mathbf{s}^T = \mathbf{I}. \quad (2.3)$$

Izviri s_j v modelu (2.1) so latentne spremenljivke. To pomeni, da jih ne moremo neposredno izmeriti. Prav tako ne poznamo mešalne matrike $\mathbf{A}$. Naloga metode ICA je torej samo iz podatkov meritev $\mathbf{x}$ izluščiti prvotne izvire signalov in določiti način mešanja izvirov v merjene signale. Pri tem niti matrika $\mathbf{A}$ niti izviri $\mathbf{s}$ ne morejo biti enolično določeni. Iz modela (2.1) je razvidno, da:

1. Variance neodvisnih komponent ne morejo biti enolično določene.

Ker sta tako $\mathbf{A}$ kot tudi $\mathbf{s}$ neznana, lahko namreč katerikoli faktor v enem izmed izvirov s_j skrajšamo z deljenjem ustreznega stolpca $\mathbf{a}_j$ v matriki $\mathbf{A}$ z istim faktorjem (1.6). Zato je najbolj naravno v modelu predpostaviti, da imajo neodvisne komponente enotsko varianco, $\langle s_j^2 \rangle = 1$, (enačba (2.3)). Temu pogoju bo v rešitvi prilagojena tudi matrika $\mathbf{A}$.

2. Vrstni red neodvisnih komponent ne more biti enolično določen.

Ker sta tako $\mathbf{A}$ in $\mathbf{s}$ neznana, lahko namreč prosto zamenjamo vrstni red v vsoti (2.2). V model (2.1) na desni strani vstavimo permutacijsko matriko $\mathbf{P}$ in njen inverz. Dobimo $\mathbf{x} = \mathbf{A} \mathbf{P}^{-1} \mathbf{P} \mathbf{s}$. Elementi $\mathbf{P} \mathbf{s}$ so prvotni neodvisni izviri s_j , le njihov vrstni red je premešan. Matrika $\mathbf{A} \mathbf{P}^{-1}$ pa je nova mešalna matrika.

2.2 Ocena separacijske matrike in negaussovskost

Za matriko $\mathbf{A}$ s polnim rangom obstaja glede na model (2.1) separacijska matrika (ang. *separation matrix* ali *unmixing matrix*) $\mathbf{W}$, s katero lahko izvire $\mathbf{s}$ rekonstruiramo iz meritev $\mathbf{x}$:

$$\mathbf{s} = \mathbf{W}\mathbf{x}, \quad \mathbf{W} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T = (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n)^T. \quad (2.4)$$

Tu je $\mathbf{W}$ posplošen inverz matrike $\mathbf{A}$, $\mathbf{w}_j$ pa so vrstice v matriki $\mathbf{W}$. Za izračun neodvisnih komponent s_j je torej treba oceniti separacijsko matriko $\mathbf{W}$. Pri tem moramo upoštevati čim bolj splošne predpostavke o modelu. Najbolj pogosto uporabljena je predpostavka o negaussovskosti neodvisnih komponent [3]. Maksimizacija negaussovskosti pomeni enako kot minimizacija statistične odvisnosti [4, 17]. Najti poskušamo torej takšne vektorje $\mathbf{w}_j$, da bo verjetnostna porazdelitev enorazsežne projekcije izmerjenih signalov,

$$y_j = \mathbf{w}_j^T \mathbf{x}, \quad j = 1, 2, \dots, n, \quad (2.5)$$

čim manj podobna gaussovski (čim bolj negaussovska). Če je vektor $\mathbf{w}_j$ enak neki vrstici posplošenega inverza $\mathbf{W}$, je projekcija y_j enaka eni od neodvisnih komponent, $y_j = s_j$. Iskanje takšnih vektorjev je bistveni del analize neodvisnih komponent.

Mera za (ne)gaussovskost neke zvezne naključne spremenljivke y je Shannonova entropija

$$H(y) = - \int p(y) \log p(y) dy, \quad (2.6)$$

kjer je $p(y)$ verjetnostna gostota spremenljivke y . Od vseh naključnih spremenljivk z enako (npr. enotsko) varianco pa ima ravno gaussovska naključna spremenljivka največjo entropijo; to pomeni, da je najbolj naključna in najmanj napovedljiva [18]. Slednje je bistven postulat teorije informacij. Še bolj uporabna mera za negaussovskost je v praksi negentropija

$$J(y) = H(y_{\text{Gauss}}) - H(y), \quad y_{\text{Gauss}} \sim N(0,1). \quad (2.7)$$

Ta je namreč vedno nenegativna in enaka nič le v primeru, da je y porazdeljena po Gaussovi porazdelitvi. Verjetnostne gostote $p(y)$ ne poznamo vnaprej, zato za negentropijo uporabljamo približke. Najpogosteje uporabljamo približek [4]

$$J(y) = [\langle G(y) \rangle - \langle G(y_{\text{Gauss}}) \rangle]^2, \quad (2.8)$$

kjer je $G(y)$ nekadratična funkcija. Za uporabne so se izkazale sledeče izbire funkcije G [4]:

$$G_1(y) = \frac{1}{a} \log \cosh ay, \quad G_2(y) = -\exp\left(-\frac{y^2}{2}\right), \quad G_3(y) = \frac{1}{4}y^4, \quad (2.9)$$

kjer je $1 \leq a \leq 2$. Funkcija G_1 je priporočljiva za splošno rabo, funkcija G_2 je priporočljiva za signale z izrazitimi ubežnimi vrednostmi oz. s super-gaussovskimi porazdelitvami (dolgi repi v verjetnostnih gostotah, ki potekajo nad repi normalne porazdelitve), funkcija G_3 pa je kar sploščenost in je priporočljiva za subgaussovske signale, ki ne vsebujejo ubežnih vrednosti.

Oceno y_j za neodvisno komponento s_j dobimo torej tako, da poiščemo tak vektor $\mathbf{w}_j$, da bo imela projekcija $y_j = \mathbf{w}_j^T \mathbf{x}$ maksimalno negentropijo $J(y_j)$. Tedaj bo njena verjetnostna gostota najmanj podobna gaussovski. Pri tem moramo upoštevati še pogoj $\langle y_j^2 \rangle = 1$. Neodvisne komponente dobimo, ko zaporedoma poiščemo vse $\mathbf{w}_j$, $j = 1, 2, \dots, n$, pri katerih imajo negentropije $J(y_j)$ maksimume. Vektorje (vrstice) $\mathbf{w}_j$ nato zložimo v separacijsko matriko $\mathbf{W}$ in po enačbi (2.4) izračunamo neodvisne komponente (izvire) $\mathbf{s}$.

Za reševanje problemov oblike (2.1) z metodo ICA obstaja več algoritmov, najbolj pogosto uporabljeni so Infomax [20], JADE [21] in FastICA [4, 16, 17]. Algoritmi za metodo ICA se najbolj razlikujejo po načinu ocenjevanja separacijske matrike $\mathbf{W}$. Algoritem FastICA, ki ga za izračune v magistrskem delu tudi implementiramo, uporablja za oceno separacijske matrike $\mathbf{W}$ iteracijsko metodo (ang. *fixed-point algorithm*). To opišemo v nadaljevanju.

Vrnimo se k ocenjevanju vektorjev $\mathbf{w}_j$. Zanje imamo nelinearen optimizacijski problem

$$J(\mathbf{w}_j) = [\langle G(\mathbf{w}_j^T \mathbf{x}) \rangle - \langle G(y_{\text{Gauss}}) \rangle]^2 = \max. , \quad \langle (\mathbf{w}_j^T \mathbf{x})^2 \rangle = 1, \quad (2.10)$$

kjer operator $\langle \cdot \rangle$ označuje pričakovano vrednost. Po Karush-Kuhn-Tuckerjevih pogojih [19] je maksimum negentropije dosežen v točkah $\mathbf{w}_j$, kjer velja

$$F(\mathbf{w}_j) = \langle \mathbf{x} G'(\mathbf{w}_j^T \mathbf{x}) \rangle - \beta \mathbf{w}_j = 0, \quad (2.11)$$

Pri tem je β konstanta, ki jo v optimalnih točkah preprosto iz vrednotimo kot $\beta = \langle \mathbf{w}_j^T \mathbf{x} G'(\mathbf{w}_j^T \mathbf{x}) \rangle$, G' pa je odvod funkcije G , ki jo izberemo iz nabora funkcij (2.9). Za iskanje maksimuma negentropije pri pogojih (2.11) uporabimo Newton-Raphsonovo metodo. Pri tem je Jacobijeva matrika

$$\begin{aligned} J_F(\mathbf{w}_j) &= \frac{\partial F}{\partial \mathbf{w}_j} = \langle \mathbf{x} \mathbf{x}^T G''(\mathbf{w}_j^T \mathbf{x}) \rangle - \beta \mathbf{I} \\ &\approx \langle \mathbf{x} \mathbf{x}^T \rangle \langle G''(\mathbf{w}_j^T \mathbf{x}) \rangle - \beta \mathbf{I} \\ &= \langle G''(\mathbf{w}_j^T \mathbf{x}) \rangle \mathbf{I} - \beta \mathbf{I}. \end{aligned} \quad (2.12)$$

Jacobijevo matriko $J_F(\mathbf{w}_j)$ smo torej poenostavili tako, da je postala diagonalna in tako preprosto obrnljiva. Iteracijski korak z Newton-Raphsonovo metodo tako lahko zapišemo kot

$$\begin{aligned} \mathbf{w}_j^+ &= \mathbf{w}_j - J_F^{-1}(\mathbf{w}_j) F(\mathbf{w}_j) \\ &= \mathbf{w}_j - \frac{\langle \mathbf{x} G'(\mathbf{w}_j^T \mathbf{x}) \rangle - \beta \mathbf{w}_j}{\langle G''(\mathbf{w}_j^T \mathbf{x}) \rangle - \beta}. \end{aligned} \quad (2.13)$$

Zgornji postopek lahko še poenostavimo, če obe strani (2.13) pomnožimo z $\beta - \langle G''(\mathbf{w}_j^T \mathbf{x}) \rangle$:

$$\begin{aligned} \mathbf{w}_j^+ (\beta - \langle G''(\mathbf{w}_j^T \mathbf{x}) \rangle) &= \mathbf{w}_j (\beta - \langle G''(\mathbf{w}_j^T \mathbf{x}) \rangle) + \langle \mathbf{x} G'(\mathbf{w}_j^T \mathbf{x}) \rangle - \beta \mathbf{w}_j \\ &= \langle \mathbf{x} G'(\mathbf{w}_j^T \mathbf{x}) \rangle - \langle G''(\mathbf{w}_j^T \mathbf{x}) \rangle \mathbf{w}_j. \end{aligned} \quad (2.14)$$

Skalarni faktor $\beta - \langle G''(\mathbf{w}_j^T \mathbf{x}) \rangle$ na levi strani zgornje enačbe lahko kar izpustimo, saj nov izračunan vektor $\mathbf{w}_j^+$ tako ali tako po vsaki iteraciji normiramo. Iteracijski korak tako končno zapišemo kot

$$\mathbf{w}_j^+ = \langle \mathbf{x} G'(\mathbf{w}_j^T \mathbf{x}) \rangle - \langle G''(\mathbf{w}_j^T \mathbf{x}) \rangle \mathbf{w}_j, \quad (2.15a)$$

$$\mathbf{w}_j^+ \leftarrow \frac{\mathbf{w}_j^+}{\|\mathbf{w}_j^+\|}. \quad (2.15b)$$

Pri Newtonovi metodi lahko včasih naletimo na težave s konvergenco. V takih primerih uporabimo stabilizirano različico iteracijskega koraka (2.13):

$$\mathbf{w}_j^+ = \mathbf{w}_j - \mu \frac{\langle \mathbf{x} G'(\mathbf{w}_j^T \mathbf{x}) \rangle - \beta \mathbf{w}_j}{\langle G''(\mathbf{w}_j^T \mathbf{x}) \rangle - \beta}, \quad (2.16a)$$

$$\mathbf{w}_j^+ \leftarrow \frac{\mathbf{w}_j^+}{\|\mathbf{w}_j^+\|}, \quad (2.16b)$$

kjer je $\beta = \langle \mathbf{w}_j^T \mathbf{x} G'(\mathbf{w}_j^T \mathbf{x}) \rangle$ in μ velikost koraka. Če je $\mu \ll 1$, bo Newtonova metoda konvergirala z veliko večjo gotovostjo. Dobra strategija je začeti z $\mu = 1$, nato pa velikost koraka postopoma zmanjševati.

2.3 Algoritem FastICA

Signale $\mathbf{x}_k = (x_1, x_2, \dots, x_m)_k^T$ izmerimo v N zaporednih instancah, $k = 0, 1, \dots, N-1$. Običajno (ni pa nujno) so ta "instanca" časi. Izmerjene signale uredimo v matriko

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_{N-1} \end{bmatrix} = \begin{bmatrix} x_{0,1} & x_{0,2} & \cdots & x_{0,m} \\ x_{1,1} & x_{1,2} & \cdots & x_{1,m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N-1,1} & x_{N-1,2} & \cdots & x_{N-1,m} \end{bmatrix} \in \mathbb{R}^{N \times m}, \quad (2.17)$$

v kateri vsaka od N vrstic predstavlja en m -variatni podatek (npr. izmerjeno električno polje na m merilcih ob času $t = N\Delta t$). Podatke nato pobelimo (ang. *whitening*). Tako dobimo nov vektor $\mathbf{x}_k$, ki je bel – to pomeni, da so njegove komponente nekorelirane in njegova varianca enaka 1. Postopek beljenja je sledeč. Najprej tvorimo povprečni vektor dimenzije m

$$\bar{\mathbf{x}} = \frac{1}{n} \sum_{k=0}^{N-1} \mathbf{x}_k. \quad (2.18)$$

Izračunamo še kovariančno matriko podatkov

$$\mathbf{\Sigma}_{xx} = \frac{1}{N-1} \sum_{k=0}^{N-1} (\mathbf{x}_k - \bar{\mathbf{x}})(\mathbf{x}_k - \bar{\mathbf{x}})^T, \quad (2.19)$$

ki jo diagonaliziramo, $\mathbf{\Sigma}_{xx} = \mathbf{U}\mathbf{D}\mathbf{U}^T$. Iz diagonalizacije dobimo ortogonalno matriko $\mathbf{U}$ in diagonalno matriko $\mathbf{D}$. Podatke obdelamo s transformacijo

$$\mathbf{x}_k \leftarrow \mathbf{D}^{-1/2} \mathbf{U}^T (\mathbf{x}_k - \bar{\mathbf{x}}), \quad k = 0, 1, \dots, N-1. \quad (2.20)$$

Kovariančna matrika je sedaj enaka $\mathbf{\Sigma}_{xx} = \mathbf{I}$, kar je trivialno dokazati. Nato sledimo algoritmu:

1. Izberemo število neodvisnih komponent r , ki bi jih radi določili; $r \leq m$.
2. Naključno določimo vektorje $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_r$ ($\mathbf{w}_j \in \mathbb{R}^m$), tako da ima vsak normo $\|\mathbf{w}_j\|_2 = 1$. Nato jih uredimo v matriko

$$\mathbf{W} = \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \\ \vdots \\ \mathbf{w}_r \end{bmatrix} \in \mathbb{R}^{r \times m}. \quad (2.21)$$

3. Napravimo simetrično ortogonalizacijo matrike $\mathbf{W}$:

$$\mathbf{W} \leftarrow (\mathbf{W}\mathbf{W}^T)^{-1/2} \mathbf{W}. \quad (2.22)$$

V tem koraku poskrbimo, da so vsi do sedaj poiskani $\mathbf{w}_j$ med seboj ortogonalni. S tem dekoreliramo enorazsežne projekcije izmerjenih signalov (2.5). Ta korak je nujen, sicer različni vektorji $\mathbf{w}_j$ konvergirajo k istemu vektorju $\mathbf{w}$, pri katerem je dosežen maksimum negentropije (2.10).

Inverzni koren $(\mathbf{W}\mathbf{W}^T)^{-1/2}$ pri tem izračunamo iz razcepa na lastne vrednosti (ang. *eigenvalue decomposition*) $\mathbf{W}\mathbf{W}^T = \mathbf{F}\mathbf{D}\mathbf{F}^T$ kot $(\mathbf{W}\mathbf{W}^T)^{-1/2} = \mathbf{F}\mathbf{D}^{-1/2}\mathbf{F}^T$. Še preprostejša alternativa je iterativni postopek [17]:

- a. Naj bo $\mathbf{W} = \mathbf{W} / \sqrt{\|\mathbf{W}\mathbf{W}^T\|_2}$

- b. $\mathbf{W} = \frac{3}{2} \mathbf{W} - \frac{1}{2} \mathbf{W}\mathbf{W}^T \mathbf{W}$.

Ponavljamo b., dokler ne dosežemo konvergence

4. Za vsak $j = 1, 2, \dots, r$ izračunamo nove vektorje po enačbi (2.15a):

$$\mathbf{w}_j \leftarrow \frac{1}{n} \left\{ \sum_{k=0}^{N-1} \mathbf{x}_k G'(\mathbf{w}_j^T \mathbf{x}_k) - \mathbf{w}_j \sum_{k=0}^{N-1} G''(\mathbf{w}_j^T \mathbf{x}_k) \right\} \quad (2.23)$$

Ta korak je ključni del algoritma. Poskrbi, da se v zaporednih iteracijah premikamo k negibni točki, ki ustreza maksimumu negentropije (2.10). Dobljene vektorje še normiramo po (2.15b), $\mathbf{w}_j \leftarrow \mathbf{w}_j / \|\mathbf{w}_j\|_2$.

5. Točki 3 in 4 ponavljamo, dokler ne dosežemo konvergence vseh vektorjev $\mathbf{w}_j$. Konvergenco vektorja dosežemo, ko se smer vektorja $\mathbf{w}_j$ v zaporednih iteracijah $(\tau - 1)$ in (τ) ne spreminja več, torej kadar je absolutna vrednost skalarnega produkta $|\mathbf{w}_j^{(\tau)T} \mathbf{w}_j^{(\tau-1)}|$ blizu 1. Postopek torej ustavimo, ko velja

$$\left| |\mathbf{w}_j^{(\tau)T} \mathbf{w}_j^{(\tau-1)}| - 1 \right| < \varepsilon, \quad j = 1, 2, \dots, r, \quad (2.24)$$

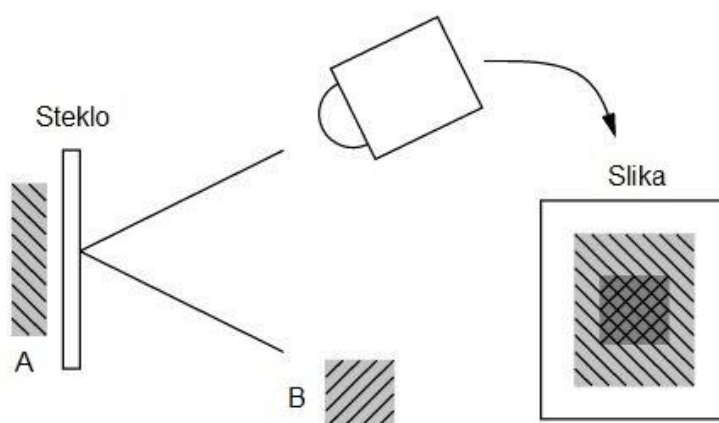
kjer je ε neki vnaprej izbran majhen pozitiven parameter.

6. Neodvisne komponente so potem vrstice matrike $\mathbf{S} = \mathbf{W}\mathbf{X}^T \in \mathbb{R}^{r \times N}$, kjer je $\mathbf{X}$ matrika po transformaciji (2.20).

3 Odstranjevanje odbojev od dielektričnih površin iz fotografij

3.1 Opis problema

Ko gledamo ali fotografiramo skozi okno, se pogosto zgodi, da je slika, ki jo zaznamo ali zajamemo, sestavljena iz dveh komponent: iz slike, ki nastane zaradi odboja svetlobe od predmetov izza okna, in iz slike, ki nastane zaradi odboja žarkov na stekleni površini okna (slika 3.1). V tem primeru predmetov za oknom ne moremo videti povsem čisto, saj je zaznani sliki primešan še odboj. Ta nastane na steklu, ki je dielektrična površina. Odbita svetloba je delno polarizirana. To pomeni, da lahko intenziteto odbite svetlobe (gostoto svetlobnega toka), ki jo zajamemo z merilnikom svetlobe (fotometrom – oko, fotoaparati), uravnavamo s spreminjanjem orientacije linearne polarizatorja, ki ga postavimo pred merilnik. Vseeno se z uporabo linearne polarizatorja ne moremo povsem znebiti odbojev, razen v primeru, ko je vpadni kot



Slika 3.1. Slika, ki jo zaznamo, ko gledamo skozi stekleno okno, je pogosto sestavljena iz dveh komponent: iz slike, ki nastane zaradi odboja svetlobe od predmetov izza okna, (A) in iz slike, ki nastane zaradi odboja žarkov na stekleni površini okna (B).

odbite svetlobe enak Brewsterjevemu kotu in je vsa odbita svetloba popolnoma linearno polarizirana. Z uporabo polarizatorja in analize neodvisnih komponent pa se odbojev lahko učinkovito znebimo tudi pri ostalih vpadnih kotih. V nadaljevanju problem pojasnimo podrobneje ter podamo celovito fizikalno ozadje. Problem je bil že obravnavan v [22, 23], vendar pa njegova interpretacija ni temeljila na fizikalnem opisu.

3.1.1 Poševen vpad nepolarizirane svetlobe na dielektrično površino

Svetloba, ki se odbije od predmetov izza stekla (območje A), je običajno difuzna in nepolarizirana. Pri prehodu skozi steklo se je majhen del odbije nazaj v okolico, večji del pa jo steklo prepusti. Tudi po prehodu skozi steklo je ta svetloba še vedno nepolarizirana in kot taka vpade v merilnik svetlobe. Pri obravnavi našega problema pa je pomembnejša svetloba, ki vpada na steklo iz območja B in se potem na steklu zrcalno odbije v merilnik. V tem primeru je vpadna svetloba ravno tako običajno nepolarizirana in difuzna, odbiti del pa je delno polariziran.

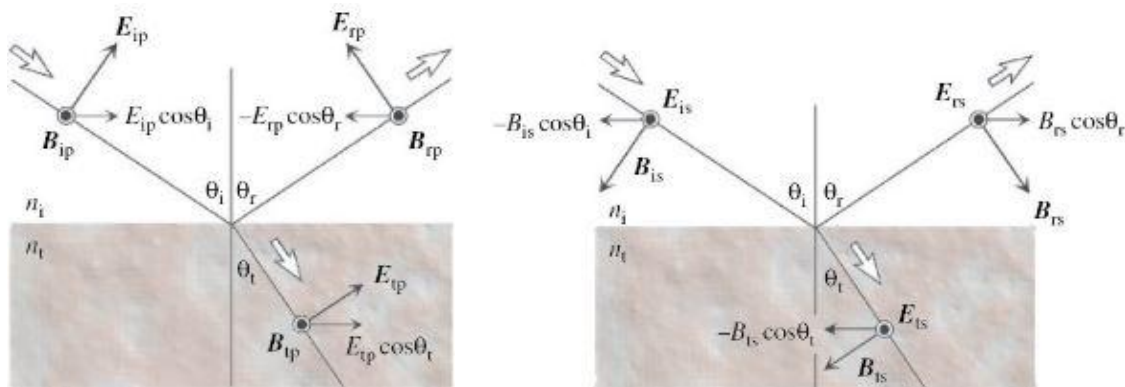
Opišimo dogajanje podrobneje. Svetlobo (elektromagnetno valovanje), ki se širi v smeri osi z , opišemo kot

$$\mathbf{E} = \mathbf{E}_0 \exp[i(\omega t - kz + \delta)], \quad (3.1a)$$

$$\mathbf{B} = \mathbf{B}_0 \exp[i(\omega t - kz + \delta)], \quad (3.1b)$$

kjer je $\mathbf{E}_0 = c\mathbf{B}_0$. Ker sta električna poljska jakost $\mathbf{E}$ in gostota magnetnega polja $\mathbf{B}$ povezani količini, v nadaljnjem opisu delamo zgolj z električno poljsko jakostjo. Razcepimo jo na pravokotni komponenti, ki sta hkrati pravokotni na smer širjenja valovanja: $\mathbf{E} = E_x \hat{\mathbf{x}} + E_y \hat{\mathbf{y}}$. Kadar je smer $\mathbf{E}$ konstantna, imamo opravka z linearno polarizirano svetlobo. Kadar pa se smer $\mathbf{E}$ naključno spreminja, gre za nepolarizirano svetlobo. Za popolnoma nepolarizirano svetlobo velja $\langle E_x \rangle = \langle E_y \rangle$, kar pomeni, da se vektor $\mathbf{E}$ spreminja v vseh smereh enakomerno oz. imamo enakomerno mešanico v vseh smereh linearno polarizirane svetlobe.

V prejšnjem odstavku smo poudarili razliko med linearno polarizirano in nepolarizirano



Slika 3.2. Vpad [LEVO] p- in [DESNO] s-polarizirane svetlobe iz zraka ($n \approx 1$) na steklo ($n \approx 1.5$), torej na optično gostejšo dielektrično površino. Indeks i (ang. *incident*) ponazarja vpadno, r (ang. *reflected*) odbito in t (ang. *transmitted*) prepuščeno svetlobo. Povzeto po [24].

svetlobo. Obravnavajmo sedaj poševen vpad nepolarizirane svetlobe (iz območja B; slika 3.1) iz zraka ($n_i \approx 1$) na steklo ($n_t \approx 1.5$), torej na optično gostejšo dielektrično površino. Glede na vpadno ravnino ima nepolarizirana svetloba dve komponenti, ki se spreminjata: p- (transverzalno magnetno, TM) in s- (transverzalno električno, TE) komponento (slika 3.2). Če ima svetloba le eno od obeh komponent, je linearno polarizirana. V tem primeru je lahko glede na vpadno ravnino svetloba p- ali s-polarizirana. Primera se razlikujeta v smereh osciliranja električnega in magnetnega polja. Pri p-polarizirani svetlobi vpadno in odbito električno polje oscilirata v isti ravnini – v vpadni ravnini. Pri s-polarizirani svetlobi pa to velja za magnetno polje. Robni pogoji za elektromagnetno valovanje na meji med dvema sredstvom pravijo, da sta komponenti električnega in magnetnega polja, ki sta vzporedni z mejo, enaki pred in po prehodu. Iz tega sledijo Fresnelove enačbe za prepuščeno in odbito svetlobo:

$$r_p \equiv \frac{E_{rp}}{E_{ip}} = \frac{n_t \cos \theta_i - n_i \cos \theta_t}{n_t \cos \theta_i + n_i \cos \theta_t}, \quad (3.2a)$$

$$t_p \equiv \frac{E_{tp}}{E_{ip}} = \frac{2n_i \cos \theta_i}{n_t \cos \theta_i + n_i \cos \theta_t}, \quad (3.2b)$$

$$r_s \equiv \frac{E_{rs}}{E_{is}} = \frac{n_i \cos \theta_i - n_t \cos \theta_t}{n_i \cos \theta_i + n_t \cos \theta_t}, \quad (3.2c)$$

$$t_s \equiv \frac{E_{ts}}{E_{is}} = \frac{2n_i \cos \theta_i}{n_i \cos \theta_i + n_t \cos \theta_t}. \quad (3.2d)$$

Gostota energijskega toka pri EM valovanju (oz. gostota svetlobnega toka, ang. *light intensity*) z amplitudo električne poljske jakosti E je enaka $j = \frac{1}{2} \varepsilon_0 c_0 E^2$. Odtod sledijo izrazi za odbojnost (reflektivnost, R):

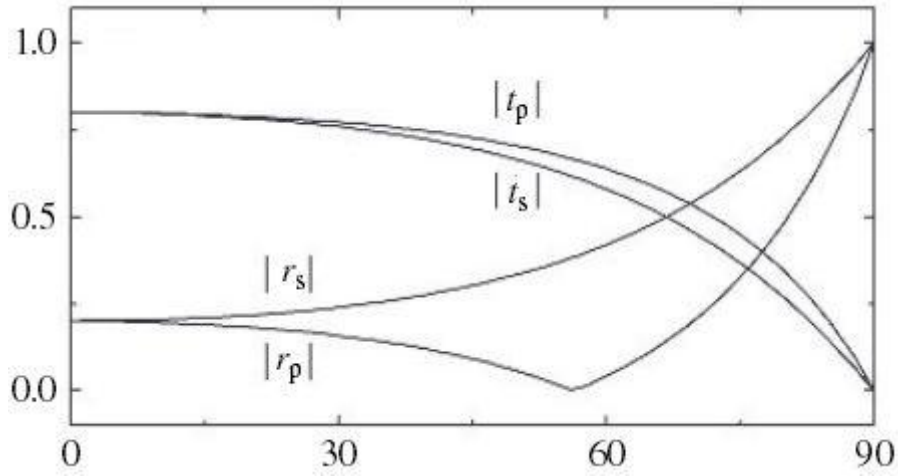
$$R_p \equiv \frac{j_{rp}}{j_{ip}} = \left| \frac{E_{rp}}{E_{ip}} \right|^2 = |r_p|^2, \quad R_s \equiv \frac{j_{rs}}{j_{is}} = \left| \frac{E_{rs}}{E_{is}} \right|^2 = |r_s|^2, \quad (3.3)$$

in prepustnost (transmisivnost, T) p- in s-polarizirane svetlobe:

$$T_p \equiv \frac{j_{tp} \cos \theta_t}{j_{ip} \cos \theta_i} = \left(\frac{n_t \cos \theta_t}{n_i \cos \theta_i} \right) \left| \frac{E_{tp}}{E_{ip}} \right|^2 = \left(\frac{n_t \cos \theta_t}{n_i \cos \theta_i} \right) |t_p|^2, \quad (3.4a)$$

$$T_s \equiv \frac{j_{ts} \cos \theta_t}{j_{is} \cos \theta_i} = \left(\frac{n_t \cos \theta_t}{n_i \cos \theta_i} \right) \left| \frac{E_{ts}}{E_{is}} \right|^2 = \left(\frac{n_t \cos \theta_t}{n_i \cos \theta_i} \right) |t_s|^2. \quad (3.4b)$$

Pri poševnem vpadu nepolarizirane svetlobe s spremenljivima komponentama p in s na dielektrično površino je torej odbita svetloba deloma polarizirana, pri tem je njena komponenta s precej večja od komponente p (slika 3.3), $|r_s| > |r_p|$. Pri vpadnem kotu θ_i , ki je enak Brewsterjevemu kotu,



Slika 3.3. Absolutne vrednosti koeficientov (3.2) pri različnih vpadnih kotih svetlobe θ_i ; za prehod svetlobe čez mejo zrak-steklo. Pri Brewsterjevem kotu $\theta_i = \theta_B \approx 56^\circ$ je amplituda električne poljske jakosti p-polarizirane svetlobe $E_{rp} = 0$ in zato $|r_p| = 0$. Za nepolarizirano (naravno) svetlobo velja $t_p \approx t_s$ za vse θ_i [22].

$$\theta_B = \arctan\left(\frac{n_t}{n_i}\right), \quad (3.5)$$

pa je amplituda električne poljske jakosti $E_{rp} = 0$ (za razlago glej [22]), zato je koeficient $r_p = 0$ in posledično odbojnost $R_p = 0$, kar pomeni, da odbita svetloba nima komponente p. Odbita svetloba je torej v tem primeru popolnoma linearno polarizirana, s-polarizirana. Za mejo med zrakom in steklom je Brewsterjev kot enak $\theta_B \approx 56^\circ$.

Kadar je vpadni kot svetlobe enak Brewsterjevemu kotu, $\theta_i = \theta_B$, lahko torej z idealnim linearnim polarizatorjem, ki ga postavimo pred merilnik svetlobe, pri določeni orientaciji polarizatorja povsem izločimo s-polarizirano svetlobo, ki se odbije od dielektrične površine. Orientacija polarizatorja je odvisna od njegove prepustne smeri. Prepustna smer polarizatorja $\mathbf{r}$ pa je določena s smerjo električne poljske jakosti $\mathbf{E}$ v linearno polariziranem valovanju, ki ga polarizator prepusti neoslabljenega [25]. Prepustna smer je smer, v kateri velja

$$\mathbf{E} = E\mathbf{r}, \quad (3.6)$$

kjer je E amplituda popolnoma prepuščene električne poljske jakosti. V vmesni legi $\mathbf{r}'$ pa prepusti polarizator del vpadnega valovanja s komponento jakosti električnega polja v prepustni smeri, torej

$$E' = E\mathbf{r}\mathbf{r}' = E \cos \varphi, \quad (3.7)$$

kjer je φ kot med prepustno smerjo $\mathbf{r}$ in vmesno lego $\mathbf{r}'$. Velja še $j \propto E^2$. Odtod potem sledi Malusov zakon [26] za gostoto prepuščenega svetlobnega toka v odvisnosti od orientacije polarizatorja

$$j' = j \cos^2 \varphi, \quad (3.8)$$

kjer je j gostota na polarizator vpadlega svetlobnega toka.

Obravnavamo še drug, v praksi seveda mnogo bolj pogost problem, ko vpadni kot $\theta_i \neq \theta_B$. Takrat je odbita svetloba le deloma polarizirana in ima tako p- kot s-komponento, $|r_s| > |r_p|$. Odbite svetlobe zato z linearnim polarizatorjem ne moremo povsem izločiti. V tem primeru si pri odstranjevanju odbojev lahko pomagamo z metodo analize neodvisnih komponent.

3.1.2 Modelski nastavek

Najprej zato zapišemo izraz za svetlobni tok $j^{(P)}$, ki ga prejme merilnik svetlobe v neki točki. Enak je

$$j^{(P)} = \left[\frac{T_p + T_s}{2} \right]_{\theta_i^{(A)}} j^{(A)} + \left[\frac{R_p + R_s}{2} \right]_{\theta_i^{(B)}} j^{(B)}. \quad (3.9)$$

Tu je $j^{(A)}$ gostota svetlobnega toka nepolarizirane svetlobe iz izvira A, ki ga steklo prepusti pri vpadnem kotu $\theta_i^{(A)}$ s skupno prepustnostjo $T_n = (T_p + T_s)/2$. Pri prehodu skozi steklo ostane svetloba iz izvira A nepolarizirana. Gostota svetlobnega toka $j^{(B)}$ prav tako nepolarizirane svetlobe iz izvira B pa se pri vpadnem kotu $\theta_i^{(B)}$ od stekla odbije s skupno odbojnostjo $R = (R_p + R_s)/2$ kot deloma polarizirana ($R_s > R_p$) nazaj v merilnik. V modelu (3.9) smo zanemarili absorpcijo svetlobe v steklu.

Ko pred merilnik svetlobe postavimo še linearni polarizator pri poljubni orientaciji, model (3.9) dopolnimo v skladu z enačbo (3.8):

$$j_{POL}^{(P)} = \underbrace{\left[\langle \cos^2 \varphi \rangle \frac{T_p + T_s}{2} \right]_{\theta_i^{(A)}}}_{\tilde{a}} j^{(A)} + \underbrace{\left[\frac{R_p \cos^2 \varphi + R_s \cos^2 \left(\frac{\pi}{2} + \varphi \right)}{2} \right]_{\theta_i^{(B)}}}_{\tilde{b}} j^{(B)}, \quad (3.10)$$

kjer je φ kot med smerjo p-komponente prepuščene svetlobe iz izvira A (oz. analogno odbite svetlobe iz izvira B) in prepustno smerjo polarizatorja ter $\langle \cos^2 \varphi \rangle = 1/2$. V popolnoma nepolarizirani svetlobi so namreč linearne polarizacije v vseh smereh (pri vseh kotih) enakomerno porazdeljene. Iz modela (3.10) je jasno, da bo del $\tilde{a}$ v vsoti izjemno malo odvisen od orientacije polarizatorja (majhna odvisnost bo zgolj na račun tega, da svetloba ni popolnoma nepolarizirana), del $\tilde{b}$ pa bo močno odvisen. Prav tako smo zanemarili dejstvo, da je svetlobni tok, ki prihaja v točke merilnika svetlobe (te se odražajo na sliki v pikslih), za različne točke posledica različnih vpadnih kotov $\theta_i^{(A)}$ in $\theta_i^{(B)}$. Vseeno lahko, predvsem pri fotografiranju pri večjih oddaljenostih od stekla, predpostavimo, da je $\theta_i^{(A)} = \text{konst.}$ in $\theta_i^{(B)} = \text{konst.}$ Poenostavljeno zato lahko (3.10) zapišemo kar kot linearen model za gostoto svetlobnega toka, ki velja v vsaki točki merilnika svetlobe:

$$j_{POL}^{(P)} = \tilde{a}j^{(A)} + \tilde{b}j^{(B)}. \quad (3.11)$$

V razdelku 3.2 se bo izkazalo, da je približek smiseln in da z njim dobimo dobre rezultate. Na fotografiji pa se z merilnikom (fotoaparatom) v neki točki zajeta svetloba odraža kot piksel. Posnamemo sliko velikosti $C = X \times Y$ pikslov. V naših poskusih vsak piksel predstavlja numerična vrednost $y \in [0,1]$ za intenziteto svetlobe oz. relativno svetlost v črno-belem barvnem prostoru (ang. *grayscale*). Iz numerične vrednosti y smo vpliv nelinearnega popravka gamma že izločili. Ta je nujen za pravilen prikaz slike na ekranih (dodatek B). V dodatku C pa pokažemo, da je preslikava med gostoto svetlobnega toka in intenziteto svetlobe v črno-belem barvnem prostoru y linearna, zaradi česar lahko (3.11) za intenziteto svetlobe zapišemo kar kot

$$y_{POL}^{(P)} = ay^{(A)} + by^{(B)}. \quad (3.12)$$

V nadaljevanju pokažemo, kako z metodo analize neodvisnih komponent odstranimo iz fotografij odboje na dielektričnih površinah. Pri tem je nujna predpostavka o statistični neodvisnosti izvirov $y^{(A)}$ in $y^{(B)}$, ki pa je seveda smiselna.

3.1.3 Uporaba metode analize neodvisnih komponent

Enako fotografijo posnamemo skozi steklo pri m različnih orientacijah linearnega polarizatorja. Vsako sliko za potrebo obdelave z metodo ICA predstavimo tako, da njene vrstice eno za drugo zložimo v vektor (vrstico) $\mathbf{y} \in \mathbb{R}^{1 \times C}$, kjer je C število vseh pikslov. Vrednost posameznega piksla je $y \in [0,1]$. Glede na definicijo metode ICA (razdelek 2.1) tu posamezno meritev x_i zamenjamo z vrstico $\mathbf{y}_i$, ki predstavlja posnetek pri neki orientaciji linearnega polarizatorja. Zapišemo torej po vzoru (2.1) model za (3.12)

$$\underbrace{\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_m \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ \vdots & \vdots \\ a_m & b_m \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} \mathbf{y}^{(A)} \\ \mathbf{y}^{(B)} \end{bmatrix}}_{\mathbf{s}}. \quad (3.13)$$

Naš namen je z metodo ICA rešiti zgornji linearni inverzni problem, ki je oblike (1.1). Ločiti torej želimo oba izvira – sliko $\mathbf{y}^{(B)}$, ki nastane kot posledica odbojev žarkov od steklene površine, od slike $\mathbf{y}^{(A)}$, ki je posledica od predmetov odbite svetlobe izza stekla. V našem primeru vedno izberemo $m = 2$, posnamemo torej sliko pri dveh različnih orientacijah linearnega polarizatorja.

Stolpci z metodo ICA rekonstruirane matrike $\mathbf{A}$ so določeni do skalarne faktorja natančno (razdelek 2.1). Vse negativne stolpce $\mathbf{a}_j$ in pripadajoče izvire s_j zato najprej pomnožimo s faktorjem (-1) :

$$\mathbf{a}_j \leftarrow \mathbf{a}_j(-1), \quad s_j \leftarrow s_j(-1). \quad (3.14)$$

Nato z linearno preslikavo

$$\mathbf{a}_j \leftarrow \frac{\mathbf{a}_j}{c}, \quad s_j \leftarrow cs_j + d, \quad (3.15)$$

izvire še "napihnemo", tako da bodo vrednosti relativnih svetlosti za posamezen izvir zajemale celoten interval $s_j \in [0,1]$.

3.2 Rezultati

3.2.1 Ločevanje linearne kombinacije črnobelih slik

Najprej prikažemo, kako deluje algoritem ICA pri ločevanju linearne kombinacije črnobelih slik. Tri izvorne slike $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3]^T$ velikosti 128×128 pikslov (slika 3.4) linearno zmešamo z matriko

$$\mathbf{A} = \begin{bmatrix} 0.30 & 0.20 & 0.50 \\ 0.60 & 0.15 & 0.25 \\ 0.25 & 0.60 & 0.15 \end{bmatrix} \quad (3.14)$$

in tako dobimo nove slike $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3]^T$ (slika 3.5). Z metodo ICA jih potem zopet ločimo. Ocenjene izvorne slike so prikazane na sliki 3.6.



Slika 3.4. Izvirne slike velikosti 128×128 pikslov, ki jih po vrsti linearno zmešamo z mešalno matriko $\mathbf{A}$ (enačba (3.14)).



Slika 3.5. Z mešalno matriko $\mathbf{A}$ (enačba (3.14)) zmešane izvorne slike.



Slika 3.6. Z metodo ICA ocenjene izvirne slike. Vrstni red ocenjenih slik je seveda nedoločen. Za vsako izmed slik izračunamo še koren povprečja kvadratov razlik med relativni svetlostmi ocenjenih in pripadajočih izvirnih slik: $\sqrt{\langle \tilde{s} - s \rangle}$. Za zgornje tri slike dobimo po vrsti vrednosti 0.153, 0.146 in 0.095.

3.2.2 Odstranjevanje odbojev od dielektričnih površin

S fotoaparatom, pred katerega objektiv postavimo linearni polarizator, posnamemo dve fotografiji skozi stekleno površino, vsakič pri različni orientaciji polarizatorja. Skladno z modelom (3.12) velja, da je enkrat na fotografiji močnejše zastopana slika, ki nastane pri odboju svetlobe od steklene površine, drugič pa šibkeje, odvisno od orientacije polarizatorja. Ker vpadni kot odbite svetlobe v splošnem ni enak Brewsterjevemu kotu, se odbojev le z uporabo polarizatorja ne moremo znebiti. Zato uporabimo metodo ICA.

Najprej prikažemo uporabo na primeru fotografiranja skozi vetrobransko steklo avtomobila. Posneti črno-beli fotografiji y_1 in y_2 sta prikazani na sliki 3.7. Slika, ki nastane pri odboju svetlobe od steklene površine, je na spodnji fotografiji zastopana močnejše kot na zgornji. Privzamemo, da velja model (3.12) in uporabimo metodo ICA na način, opisan v podrazdelku 3.1.3. Z metodo ICA dobimo oceno za mešalno matriko:

$$\mathbf{A} = \begin{bmatrix} 0.936 & 0.291 \\ 0.931 & 0.445 \end{bmatrix}. \quad (3.15)$$

Rezultat je skladen z našo opazko (glej razlago pod enačbo (3.10)), da bo gostota svetlobnega toka nepolarizirane prepuščene svetlobe iz izvira A skoraj neodvisna od orientacije polarizatorja – vrednosti v prvem stolpcu matrike $\mathbf{A}$ sta namreč skoraj identični. Gostota svetlobnega toka svetlobe iz izvira B, ki se od stekla kot deloma polarizirana odbije v fotoaparatu, pa je precej odvisna od orientacije polarizatorja – odtod tudi različni vrednosti v drugem stolpcu mešalne matrike $\mathbf{A}$. Rezultate, torej z metodo ICA ločena izvira – sliko $y^{(A)}$, ki je posledica skozi steklo prepuščene nepolarizirane svetlobe, in sliko $y^{(B)}$, ki je posledica odboja svetlobe od steklene površine (pa tudi od drugih delov v notranjosti avtomobila) – prikažemo na sliki 3.8.

Zgornji problem je torej rekonstrukcija problema, obravnavanega v [22, 23]. V magistrskem delu pa pokažemo še, da lahko odboje od dielektričnih površin ločimo tudi na barvnih fotografijah. Posnamemo torej še dve barvni fotografiji, y_1 in y_2 , z istim motivom. V primeru barvnih fotografij vsak piksel na sliki predstavlja tri vrednosti, ki pomenijo zastopnost osnovnih barv: R (rdeča), G (zelena) in B (modra). Kombinacija osnovnih barv namreč enolično določa katerokoli barvo znotraj RGB barvnega prostora [27]. Obe fotografiji razstavimo na tri osnovne barvne komponente. Tako dobimo vektorje y_{1R} , y_{1G} , y_{1B} in y_{2R} , y_{2G} , y_{2B} za relativno svetlost posamezne osnovne barve (slika 3.9). V dodatku C pokažemo še, da je preslikava



Slika 3.7. Črnobeli fotografiji [ZGORAJ] y_1 in [SPODAJ] y_2 , velikosti 640×480 pikslov, posneti skozi vetrobransko steklo avtomobila. Pred objektiv fotoaparata smo postavili linearni polarizator in spreminjali njegovo orientacijo. Slika armaturne plošče in predmetov na njej, ki nastane pri odboju svetlobe od steklene površine, je [SPODAJ] zato zastopana močnejše kot [ZGORAJ]. Sprememba orientacije polarizatorja pri enem in drugem posnetku je bila $\Delta\varphi \approx 30^\circ$.



Slika 3.8. Z metodo ICA ločeni izvorni sliki, [ZGORAJ] je slika skozi steklo prepuščne svetlobe brez odbojev, $y^{(A)}$, [SPODAJ] pa slika $y^{(B)}$, ki jo ustvari svetloba, ki se odbije na stekleni površini. Na slednji so vidni armaturna plošča, časopisni papir na njej ter nosilec vzvratnega ogledala. Na odbiti sliki so vidni tudi obrisi oddaljenih oblakov. Svetloba se namreč pri sipanju na oblačnih kapljicah odbije kot deloma polarizirana. Z metodo ICA torej v principu ločujemo nepolarizirano svetlobo od deloma polarizirane.

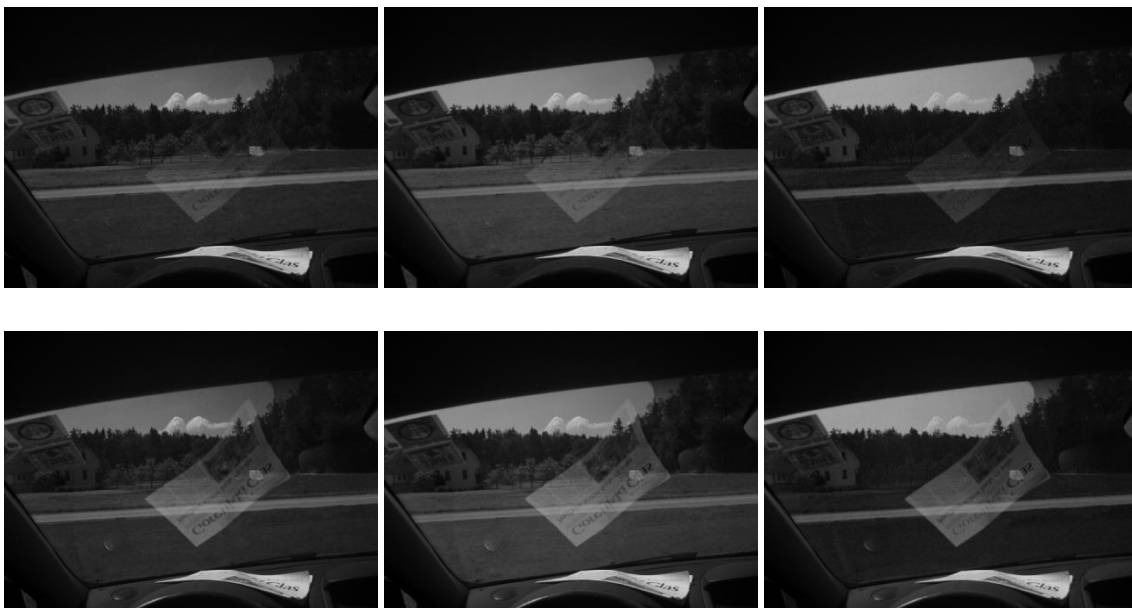
gostote svetlobnega toka v relativno svetlost katerekoli osnovne barve linearna, kar pomeni, da tudi za relativno svetlost osnovnih barv velja model (3.12). Z metodo ICA nato ocenimo mešalno matriko za vsako osnovno barvo posebej:

$$\mathbf{A}_R = \begin{bmatrix} 0.388 & 0.113 \\ 0.399 & 0.033 \end{bmatrix}, \quad (3.16a)$$

$$\mathbf{A}_G = \begin{bmatrix} 0.420 & 0.155 \\ 0.441 & 0.061 \end{bmatrix}, \quad (3.16b)$$

$$\mathbf{A}_B = \begin{bmatrix} 0.385 & 0.119 \\ 0.413 & 0.031 \end{bmatrix}. \quad (3.16c)$$

Mešalne matrike so si med seboj precej podobne, iz česar lahko sklepamo, da velja za cel spekter vidne svetlobe podoben model. Rezultat so z metodo ICA ločeni izviri v vseh treh osnovnih barvah: $\mathbf{y}_R^{(A)}$, $\mathbf{y}_R^{(B)}$ in $\mathbf{y}_G^{(A)}$, $\mathbf{y}_G^{(B)}$ ter $\mathbf{y}_B^{(A)}$, $\mathbf{y}_B^{(B)}$. Ločena barvna izvira dobimo tako, da nazaj združimo barvne komponente posameznega izvira, torej $\mathbf{y}^{(A)} = [\mathbf{y}_R^{(A)}, \mathbf{y}_G^{(A)}, \mathbf{y}_B^{(A)}]$ in $\mathbf{y}^{(B)} = [\mathbf{y}_R^{(B)}, \mathbf{y}_G^{(B)}, \mathbf{y}_B^{(B)}]$. Ločeni barvni slike sta prikazani na sliki 3.10. Izkaže se, da je ločevanje nepolarizirane in deloma polarizirane svetlobe z linearnim polarizatorjem in metodo ICA izjemno učinkovito.



Slika 3.9. Barvni fotografiji $\mathbf{y}_1$ in $\mathbf{y}_2$, velikosti 640×480 pikslov, posnamemo skozi vetrobransko steklo avtomobila – zopet pri različnih orientacijah linearnega polarizatorja. Fotografiji razstavimo na tri osnovne barve: R, G in B. Na sliki [ZGORAJ] so prikazane relativne svetlosti osnovnih barv (po vrsti) R, G in B slike $\mathbf{y}_1$, torej $\mathbf{y}_{1R}$, $\mathbf{y}_{1G}$ in $\mathbf{y}_{1B}$, na sliki [SPODAJ] pa so prikazane relativne svetlosti osnovnih barv slike $\mathbf{y}_2$, torej $\mathbf{y}_{2R}$, $\mathbf{y}_{2G}$ in $\mathbf{y}_{2B}$.



Slika 3.10. Z metodo ICA ločena barvna izvira. [ZGORAJ] je prikazana slika $\mathbf{y}^{(A)} = [\mathbf{y}_R^{(A)}, \mathbf{y}_G^{(A)}, \mathbf{y}_B^{(A)}]$, ki jo ustvari svetloba, ki se odbije od predmetov izza okna, [SPODAJ] pa slika $\mathbf{y}^{(B)} = [\mathbf{y}_R^{(B)}, \mathbf{y}_G^{(B)}, \mathbf{y}_B^{(B)}]$, ki nastane zaradi odboja žarkov na stekleni površini okna ter na svetleči površini armaturne plošče.

4 Merjenje fizioloških signalov

4.1 Merjenje frekvence srčnega utripa s spletno kamero – opis

Razvoj medicine je tesno povezan z napredkom v fiziki, še posebej v optiki. Optično merjenje, kot je npr. možganska tomografija z rentgenskimi valovi, nam omogoča vpogled v notranjost telesa brez stika merilnikov s telesom. Optično merjenje bioloških tkiv je mogoče, saj se tkiva med seboj razlikujejo v optičnih lastnostih: prepustnosti, odbojnosti, sipalnosti in absorptivnosti. Biološka tkiva so optično nehomogena snov, skupno pa jim je, da je njihov lomni količnik večji od lomnega količnika zraka [28]. Pri optičnem merjenju frekvence srčnega utripa uporabljamo fotopletizmografijo (ang. *photoplethysmography*, PPG). To je tehnika, s katero merimo fiziološke signale, ki so posledica sprememb volumna opazovanih delov telesa [29]. Problem v nadaljevanju opišemo podrobneje.

V vsakem srčnem ciklu se srčni ventrikli skrčijo, tlak v njih naraste in kri kot posledica gradienta tlaka po žilah steče proti obrobju telesa, torej proti koži. Kljub temu, da udarni val (območje povečanega tlaka v žilah kot posledica hitre deformacije ventriklov) v žilnem sistemu precej oslabi, preden pride do kože, je vseeno dovolj močan, da povzroči razširitev arterij in kapilarnih arteriol v podkožnem tkivu. Spreminjanje volumna žil pomeni, da se spreminja tudi dolžina poti svetlobe skozi to tkivo. Beer-Lambertov zakon,

$$j(l) = j_0 \exp(-\mu l), \quad (4.1)$$

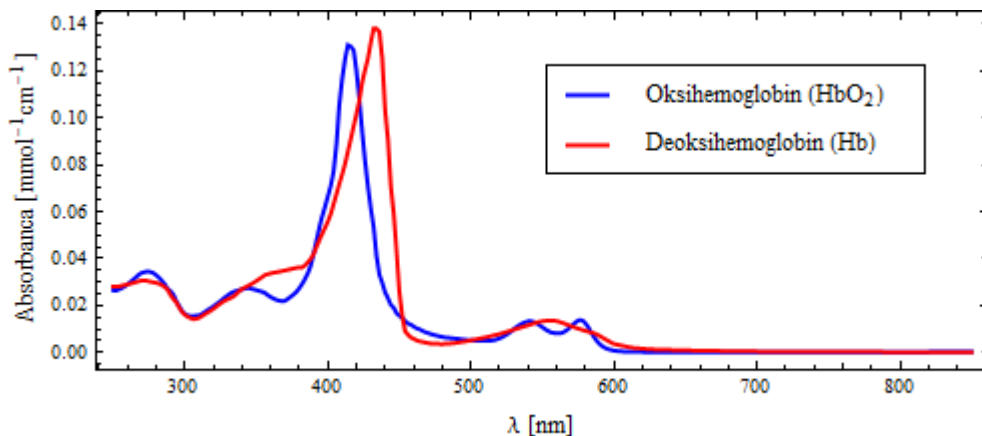
opisuje oslabitev gostote vpadnega svetlobnega toka j_0 vzdolž poti dolžine l v snovi z absorpcijskim koeficientom μ . Odtod sledi za absorptivnost snovi izraz

$$A = -\ln\left(\frac{j(l)}{j_0}\right). \quad (4.2)$$

Spreminjanje volumna žil v podkožnem tkivu tekom srčnega cikla torej pomeni, da se spreminja dolžina poti l svetlobe skozi žile, zato se hkrati spreminja tudi absorptivnost podkožnega tkiva. Merilnik svetlobe pa to zazna kot spremembe v svetlobi, ki se od kože odbije nazaj vanj. Izmerjenemu signalu, ki je posledica sprememb v odbiti svetlobi zaradi spreminjanja volumna žil, pravimo (foto)pletizmografski signal, načinu merjenja pa pogosto rečemo tudi reflektivna

(foto)pletizmografija [30] oz. reflektivna spektroskopija [28]. Frekvenca variacij v odbiti svetlobi, ki so posledica srčnega cikla, torej frekvenca pletizmografskega signala, seveda sovpada s frekvenco srčnega utripa. Dolgo je veljalo, da je merjenje pletizmografskega signala možno le, kadar je vpadna svetloba rdeča (z valovno dolžino $\lambda \approx 600$ nm) oz. infrardeča ($\lambda > 700$ nm). Ta namreč prode najgloblje v tkivo, 8 – 10 mm globoko [28]. Nedavni poskusi [31] pa so pokazali, da je fotopletizmografija možna tudi pri dnevni svetlobi oz. pri svetlobi, ki jo sevajo običajna umetna svetila.

V našem primeru merimo spremembe v svetlobi, ki se z obraza odbije v preprost merilnik svetlobe – spletno kamero (ang. *webcam*). Običajna spletna kamera ima tri svetlobne detektorje z različno spektralno občutljivostjo $dj_\lambda/d\lambda$. Ti trije detektorji pokrivajo območje vidnega in tudi infrardečega spektra. Glede na to, za katero barvo svetlobe so najbolj občutljivi oz. v katerem delu vidnega spektra je njihova relativna občutljivost največja, pa jih poimenujemo kar rdeči (R), zeleni (G) in modri (B) barvni detektor. Vsak detektor zazna mešanico pletizmografskega signala in drugih izvirov fluktuacij svetlobe, ki so lahko posledica premikov opazovanega dela telesa ali sprememb v samih virih osvetlitve, npr. v zunanji svetlosti [32]. Izvirom svetlobe, ki niso posledica pletizmografskega signala, z eno besedo rečemo artefakti. Vsak detektor zazna omenjeno mešanico z nekoliko drugačnimi utežmi. Razlog je v tem, da so že uteži samega pletizmografskega signala na vsakem senzorju različne, saj je absorptivnost hemoglobina¹, molekule, ki v krvi po arterijskih žilah prenaša kisik, močno odvisna od valovne dolžine svetlobe (slika 4.1) [33].



Slika 4.1. Molarna absorbanca (absorbanca raztopine s koncentracijo 1 mol/L) hemoglobina (Hb) pri različnih valovnih dolžinah svetlobe; za oksihemoglobin (molekula Hb s kisikom O₂) in deoksihemoglobin (Hb). Graf temelji na podatkih [34]. Iz slike je tudi jasno razvidno, zakaj je oksigenirana kri bolj svetlo rdeča kot deoksigenirana.

¹ Hemoglobin (Hb) je molekula, protein, ki v krvi po žilah prenaša kisik od dihalnih organov k ostalim delom telesa. Tam odda molekulo kisika. Glede na to ali molekula hemoglobina prenaša kisik ali ne, ločimo oksihemoglobin (z molekulo kisika) in deoksidiran hemoglobin oz. deoksihemoglobin (brez vezane molekule kisika).

Opisali smo torej fizikalno ozadje merjenja pletizmografskega signala. V naslednjem podrazdelku pa opišemo tehnično izvedbo meritev in povezavo z metodo analize neodvisnih komponent.

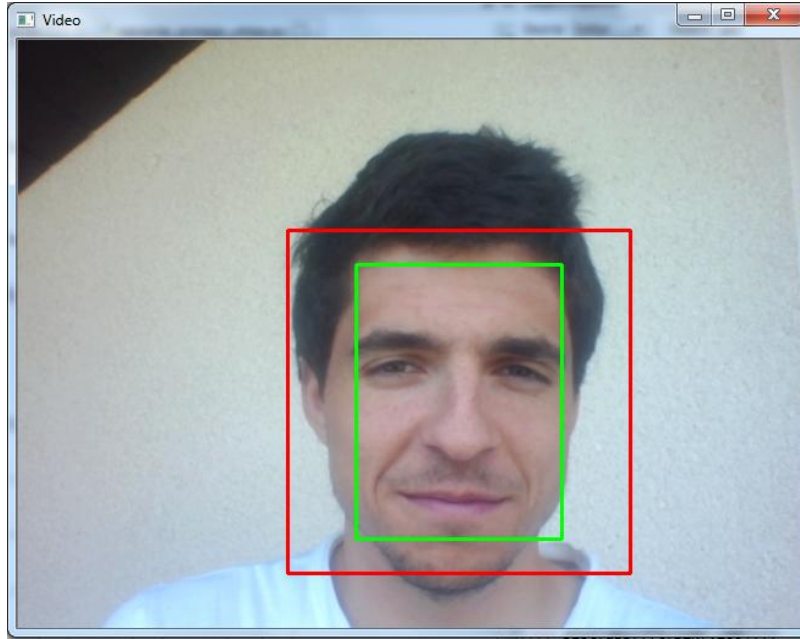
4.2 Način merjenja pletizmografskega signala

Najprej opišemo samo postavitev eksperimenta. Ker merimo pletizmografski signal, torej variacije odbite svetlobe z obraza opazovane osebe, mora biti celoten obraz ves čas znotraj vidnega polja kamere. Poleg tega sta omejitvi še, da premiki obraza niso prenašli (gibanje glave naj bo čimbolj naravno) ter da je obraz ves čas merjenja osvetljen. Vir osvetlitve je lahko tako zunanja dnevna svetloba kot tudi svetloba, ki jo oddajajo običajna umetna svetila. Pri naših eksperimentih obraz nikoli ni bil dlje kot dva metra od spletne kamere.

Posamezno meritev je predstavljal 30-sekundni posnetek obraza. Pri tem je bila frekvenca zajemanja slik (ang. *frame rate*) spremenljiva; vendar okrog 15 slik na sekundo. Frekvenca zajemanja je spremenljiva, saj je odvisna od števila računskih korakov med zajemom dveh zaporednih slik, ti pa se prav tako spreminjajo (število računskih korakov za prepoznavo obraza namreč ni konstantno). Vseeno to ne predstavlja problema, saj sproti beležimo tudi čas zajema slike, da lahko kasneje napravimo interpolacijo signala tako, da je interval Δt med dvema instancama meritve konstanten. Ločljivost zajetih slik je bila 640×480 pikslov.

Območje obraza na posnetku določimo z algoritmom za prepoznavanje objektov na sliki (dodatek A). Na vsaki zajeti sliki prepoznamo obraz (slika 4.2) in si zapomnimo njegov položaj, t.j. središče območja, omejenega s pravokotnikom. To je nujno potrebno, saj algoritem za prepoznavanje obrazov ni popolnoma učinkovit. Včasih se zgodi, da algoritem na isti sliki prepozna več obrazov, čeprav ostali zaznani objekti v resnici niso obrazi. V tem primeru predpostavimo, da predstavlja pravi obraz zgolj območje na sliki, katerega položaj je najbližje položaju obraza na prejšnji zajeti sliki. Včasih se zgodi, da algoritem na sliki ne prepozna nobenega obraza, kljub temu da se vsaj en celoten obraz nahaja na sliki. V tem primeru uporabimo kar položaj obraza iz prejšnje zajete slike. S tem postopkom zmanjšamo možnost, da bi do napak v meritvah pletizmografskega signala prihajalo zaradi napak algoritma za prepoznavanje obrazov.

Pletizmografski signal, torej variacije z obraza opazovane osebe odbite svetlobe, zaznamo s povsem preprostim optičnim merilnikom – spletno kamero s tremi barvnimi detektorji; za rdečo (R), zeleno (G) in modro (B) svetlobo. Svetloba, ki jo zaznajo trije barvni detektorji v neki točki, se na sliki odraža v treh barvnih vrednostih za vsak piksel. Vsak piksel torej predstavlja tri 8-bitne celoštevilске barvne vrednosti RGB, vrednost vsake se nahaja na intervalu $[0, 255]$. Vrnimo se k meritvam pletizmografskega signala. Najprej določimo območje obraza na zajeti sliki in ga omejimo s pravokotnikom z levim spodnjim ogliščem na položaju (a_1, b_1) in desnim zgornjim ogliščem na položaju (a_2, b_2) na sliki. Ta pravokotnik je označen na sliki 4.2 z zeleno. Nato za vsak barvni kanal posebej krajevno povprečimo barvne vrednosti tistih pikslov, ki se nahajajo znotraj pravokotnika. Tako dobimo iz slike, ki jo zajamemo ob nekem času t , tri povprečene vrednosti



Slika 4.2. Prepoznavanje obraza na sliki. Rdeč okvir predstavlja območje obraza, kakor ga določi algoritem za prepoznavanje obrazov (dodatek A), zelen okvir pa predstavlja območje, kakršnega uporabimo pri naših meritvah; njegova širina je $2/3$ rdečega okvira, višina pa $4/5$ rdečega. Iz slike, zajete ob času t , dobimo amplitude signalov $x_R(t)$, $x_G(t)$, $x_B(t)$ tako, da za piksele znotraj zelenega okvira povprečimo posebej barvne vrednosti R, G, B.

$$x_i(t) = \frac{1}{S} \sum_{\substack{a_1 \leq k \leq a_2, \\ b_1 \leq l \leq b_2}} y_i^{k,l}, \quad i = R, G, B, \quad (4.3)$$

kjer je S število vseh pikslov znotraj pravokotnika, ki omejuje območje obraza, $y_i^{k,l}$ pa je i -ta barvna vrednost piksla, ki se nahaja na položaju (k, l) na sliki. V prejšnjem podrazdelku smo zapisali, da vsak detektor zazna mešanico pletizmografskega signala, ki ga želimo izluščiti, in artefaktov. Privzamemo, da je mešanje izvirnih signalov linearno in zapišemo

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t), \quad (4.4)$$

kjer je $\mathbf{x}(t) = [x_R(t), x_G(t), x_B(t)]^T$ vektor izmerjenih signalov na barvnih detektorjih, $\mathbf{A} \in \mathbb{R}^3 \times \mathbb{R}^3$ neznana mešalna matrika, $\mathbf{s}(t) = [s_1(t), s_2(t), s_3(t)]^T$ pa vektor neznanih izvirnih signalov. Zopet imamo torej problem oblike (2.1), ki ga (kot pokažemo v naslednjem podrazdelku) uspešno rešimo z metodo analize neodvisnih komponent. Z metodo ICA torej ocenimo izvirne signale, med katerimi se skriva tudi iskani pletizmografski signal. Iz rekonstruiranega pletizmografskega signala pa nato lahko določimo frekvenco srčnega utripa opazovane osebe.

Merjenje srčnega utripa s spletno kamero je torej nekoliko podobno merjenju s pomočjo kamere in LED (ang. *light emitting diode*) izvorov svetlobe s točno določenimi emitiranimi valovnimi dolžinami – običajno so te v rdečem delu vidnega spektra in v infrardečem območju

spektra svetlobe [35, 36]. V tem primeru nek tanek del telesa, npr. prst ali pa uho, presvetlimo z diodo in merimo variacije v svetlobi, ki se odbije nazaj v kamero. Ta pletizmografski signal pa je dovolj močan (brez artefaktov), da nadaljnja obdelava z metodo ICA ni potrebna.

4.3 Rezultati meritev in analiza

Najprej prikažemo rezultate za primer 30-sekundne meritve srčnega utripa ene osebe s spletno kamero. Pri meritvi je bilo gibanje obraza opazovane osebe naravno, torej takšno kot je običajno pri delu z računalnikom. Na sliki 4.3 so prikazane meritve, narejene s tremi barvnimi detektorji: $x_R(t)$, $x_G(t)$, $x_B(t)$. Nato izračunamo še normirane merjene signale

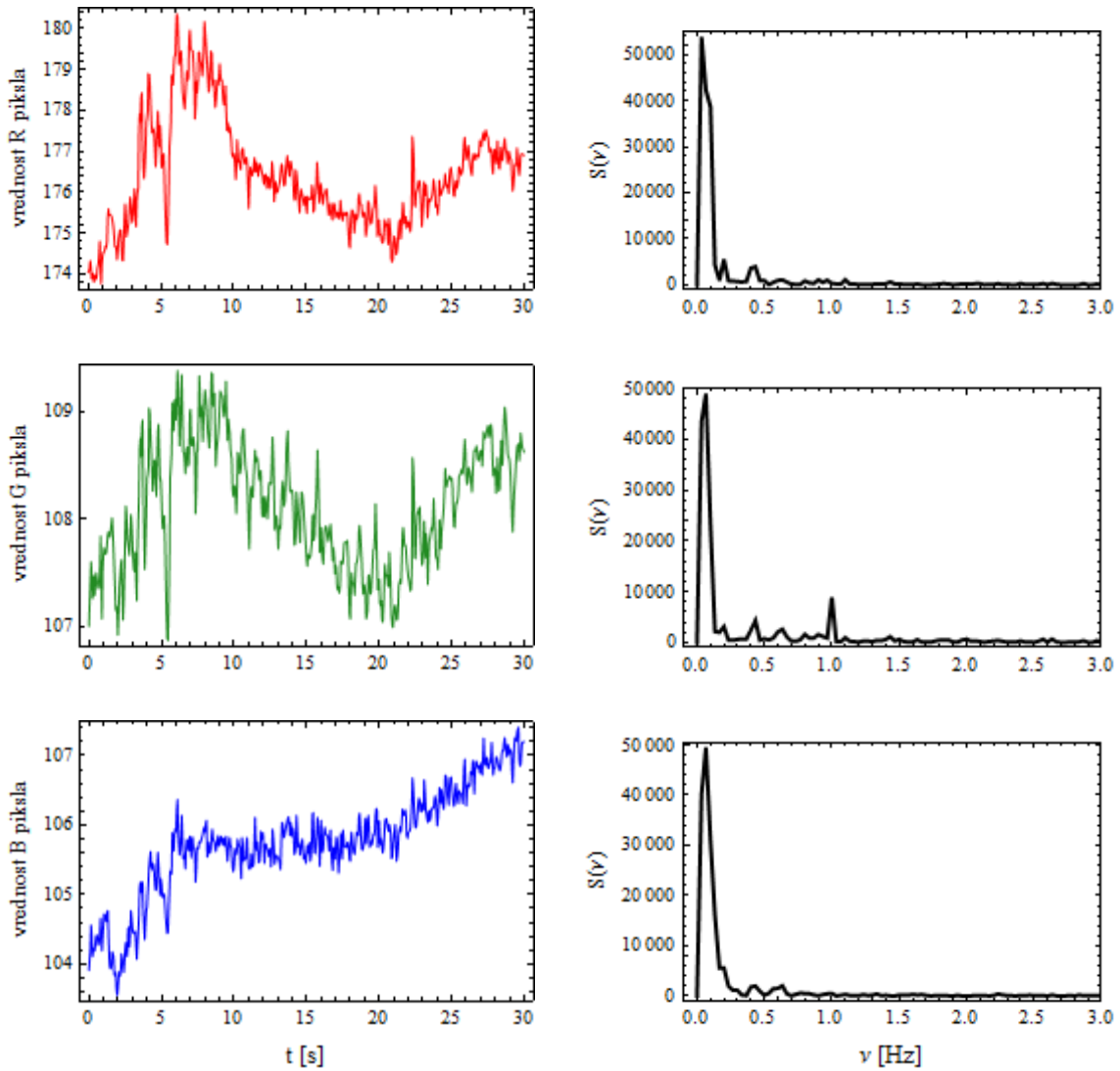
$$x_i^N(t) = \frac{x_i(t) - \mu_i}{\sigma_i}, \quad i = R, G, B. \quad (4.5)$$

Zanje potem lahko izračunamo spektralne gostote moči, $S(\nu) = |F(\nu)|^2$, pri čemer smo Fourierove transformirane $F(\nu)$ pri diskretnih frekvencah ν izračunali s hitro Fourierovo transformacijo (FFT). Po formuli (4.5) normiranim merjenim signalom pripadajoči grafi spektralne porazdelitve moči so prikazani na sliki 4.3. Opazimo, da se povprečne vrednosti pikslov na posameznih barvnih kanalih izjemno malo spreminjajo. V merjenem signalu $x_G(t)$ lahko opazimo 'naložen' periodičen signal z majhno amplitudo (manjšo od 0.5 barvne vrednosti piksla). Ta se odraža tudi na pripadajočem grafu spektralne gostote moči. Na grafu je namreč, poleg glavnega vrha pri frekvenci 0.1 Hz, dokaj izrazit vrh še pri frekvenci 1 Hz, kar pri srčnem ciklu ustreza lažje predstavljamim 60 utripom na minuto (ang. *beats per minute*, bpm). Iz tega sklepamo, da je signalu $x_G(t)$ s precejšnjo utežjo primešan pletizmografski signal, ki je posledica srčnega cikla.

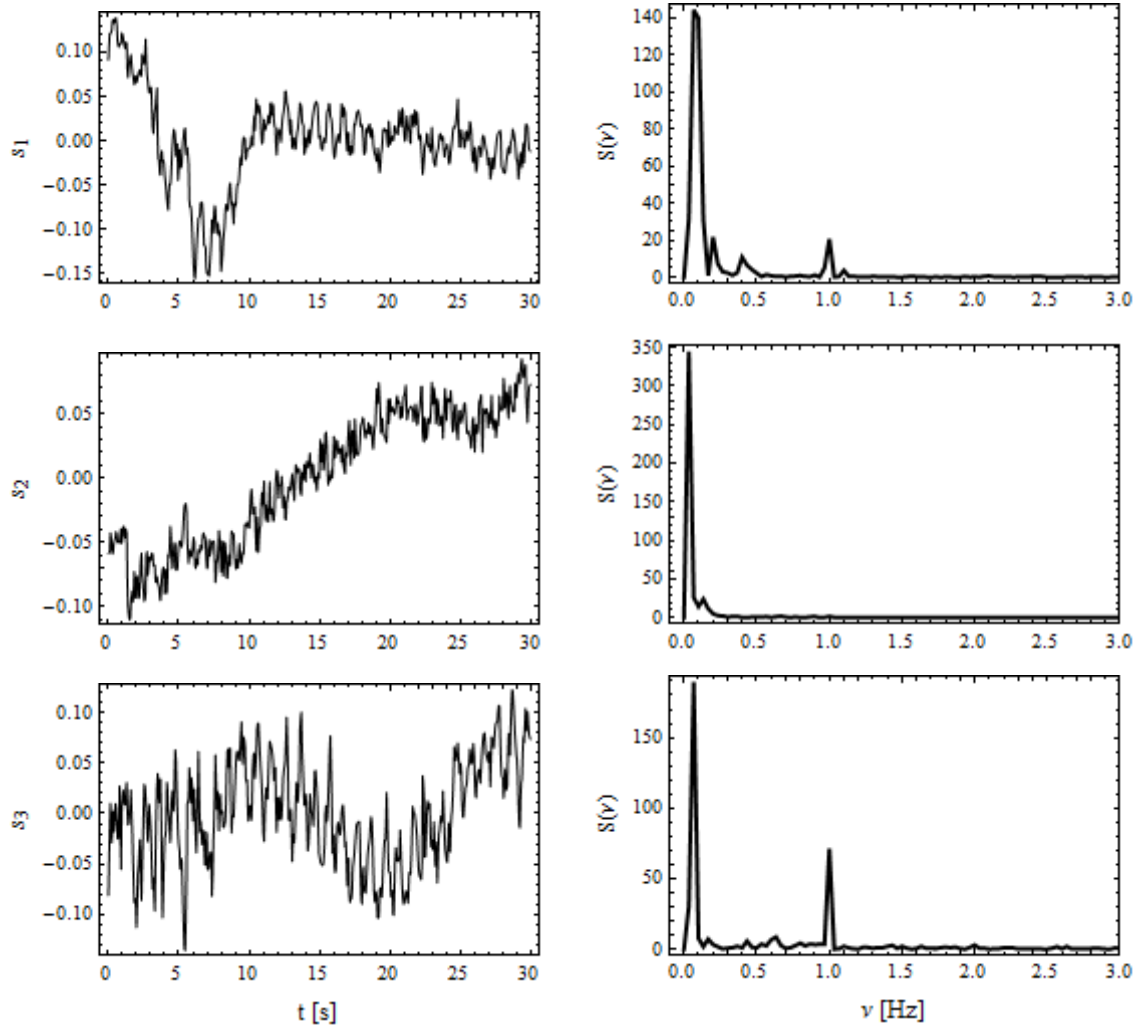
Iz izmerjenih signalov nato z metodo ICA ocenimo izvirne signale, ki so skupaj s njihovimi grafi spektralnih gostot moči prikazani na sliki 4.4. Prva dva rekonstruirana izvirna signala, torej $s_1(t)$ in $s_2(t)$, ne vsebujeta nobene pomembne periodičnosti, kar potrjujeta tudi močnostna spektra. Izvirni signal $s_3(t)$ pa je periodičen, njegov močnostni spekter pa razkriva izrazit vrh v moči signala pri frekvenci 1 Hz. To je torej frekvenca na frekvenčnem intervalu $\nu \in [0.5, 5]$ Hz oz. $\nu \in [30, 300]$ bpm, pri kateri je moč signala največja. Zanj v nadaljevanju uporabljamo oznako ν_M . O periodičnosti nekega signala $s(t)$ se še bolje prepričamo, če izračunamo njegovo avtokorelacijsko funkcijo [16], zapišemo jo kot $R_{ss}(\tau) = \langle s(t)s(t+\tau) \rangle_t$. Avtokorelacijska funkcija periodičnega signala je namreč tudi sama striktno periodična. Na sliki 4.5 sta prikazani avtokorelacijski funkciji signalov $x_G(t)$ in $s_3(t)$, torej obeh, pri katerih smo opazili periodičnosti. Izkaže se, da je avtokorelacijska funkcija $R_{s_3s_3}$ ocenjene neodvisne komponente $s_3(t)$ periodična s periodo $t_0 = 1/\nu_M = 1$ s. Periodičnost z manjšo amplitudo pa je zaznati tudi v avtokorelacijski funkciji $R_{x_Gx_G}$. Pravilno smo torej domnevali, da je meritvi $x_G(t)$ z nekoliko večjo utežjo kot pri ostalih meritvah ($x_R(t)$ in $x_B(t)$) primešan pletizmografski signal.

Neodvisna komponenta $s_3(t)$ je torej iskani pletizmografski signal, katerega frekvenco lahko povežemo s frekvenco ponavljanja srčnega cikla. Iz opravljenih meritev lahko zaključimo, da je bila frekvenca bitja srca opazovane osebe med 30-sekundno meritvijo 60 utripov na minuto. S

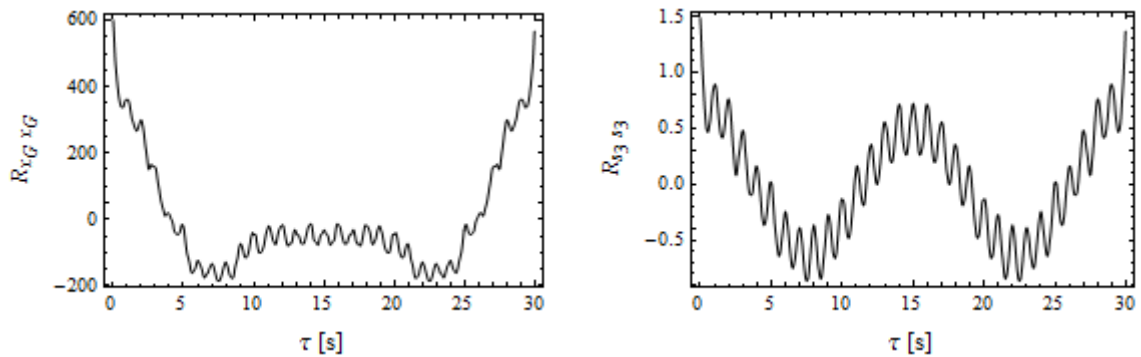
poskusom smo hkrati tudi pokazali, da je s preprostim merilnikom svetlobe, kot je spletna kamera, mogoče merjenje srčnega utripa, če pri tem uporabimo še robustno numerično metodo analize neodvisnih komponent. Izkaže se, da je privzete (4.4) o linearnem mešanju izvornih signalov dovolj dober, prav tako pa metoda odlično loči iskan pletizmografski signal od artefaktov.



Slika 4.3. [LEVO] Meritve signalov $x_R(t)$, $x_G(t)$ in $x_B(t)$ s treh barvnih detektorjev in [DESNO] po formuli (4.5) normiranim meritvam $x_R^N(t)$, $x_G^N(t)$ in $x_B^N(t)$ pripadajoči grafi spektralne gostote moči. Vrednosti pikselov na posameznih barvnih kanalih se izjemno malo spreminjajo. Razberemo lahko tudi nekoliko bolj izrazit vrh v močnostnem spektru signala $x_G(t)$. Vrh se nahaja pri frekvenci 1 Hz, kar ustreza 60 utripom na minuto (bpm). Iz tega sklepamo, da bi bil lahko pletizmografski signal močnejše zastopan v signalu $x_G(t)$.



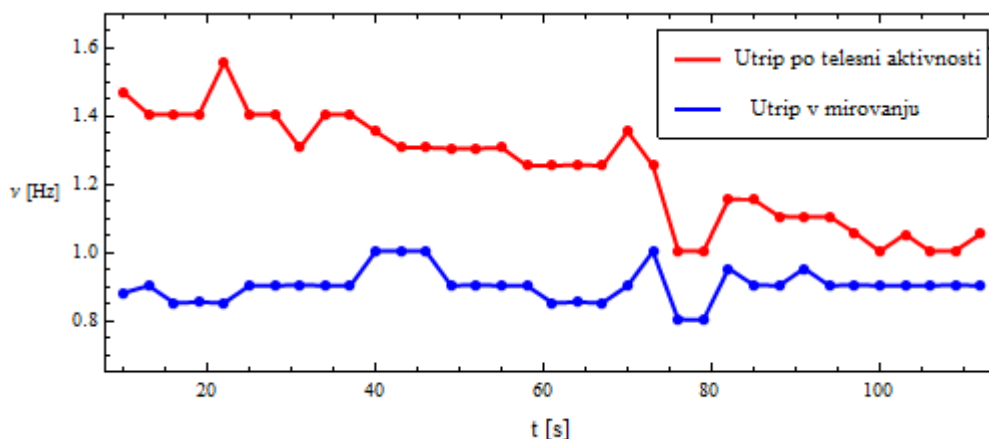
Slika 4.4. [LEVO] Neodvisne komponente oz. izvorni signali $s_1(t)$, $s_2(t)$ in $s_3(t)$, ocenjeni z metodo ICA in [DESNO] pripadajoči grafi spektralne gostote moči. Dobro je viden izrazit vrh pri $\nu = 1 \text{ Hz} = 60 \text{ bpm}$ v močnostnem spektru neodvisne komponente $s_3(t)$. Izvirni signal $s_3(t)$ torej predstavlja pletizmografski signal, njegova frekvenca pa ustreza frekvenci bitja srca.



Slika 4.5. Avtokorelacijska funkcija signala [LEVO] $x_G(t)$ in [DESNO] izrazito periodičnega $s_3(t)$.

Nazadnje napravimo še daljšo, dvominutno meritev frekvence srčnega utripa. Frekvenco srčnega utripa ob času t določimo tako, da znotraj dinamičnega časovnega okna, ki zajame časovni interval $[t - \Delta, t + \Delta t]$, analiziramo signale z barvnih detektorjev in jih obdelamo z metodo ICA. Tako rekonstruiramo izvirne signale. Zanje izračunamo še spektralne gostote moči. Nato objektivno določimo, kateri izmed ocenjenih signalov predstavlja pletizmografski signal. To storimo tako, da ugotovimo, za kateri ocenjeni signal je vrh v močnostnem spektru na frekvenčnem intervalu $\nu \in [0.5, 5]$ Hz največji v primerjavi z vrhom v celotnem močnostnem spektru signala. Signal, za katerega velja zgornji pogoj, je torej pletizmografski signal, frekvenca, pri kateri je dosežen vrh v močnostnem spektru na omejenem frekvenčnem intervalu, pa je kar frekvenca srčnega utripa.

Dinamično okno je v naših meritvah obsegalo obdobje 20 sekund, torej velja $\Delta t = 10$ s. Okno premaknemo na vsakem koraku za 3 sekunde v času naprej. To pomeni, da imamo en 'izmerk' frekvence srčnega utripa vsake 3 sekunde. Na sliki 4.6 je prikazano s spletno kamero izmerjeno spreminjanje frekvence pletizmografskega signala, ki ustreza spreminjanju frekvence srčnega utripa. Napravimo dve meritvi frekvence srčnega utripa, enkrat v času, ko je opazovana oseba v popolnem mirovanju in drugič v času mirovanja po lažji telesni aktivnosti. V prvem primeru je utrip pričakovano konstanten, v drugem primeru pa se frekvenca utripa s časom znižuje in se približuje frekvenci utripa v mirovanju.



Slika 4.6. S spletno kamero samodejno izmerjeno spreminjanje frekvence pletizmografskega signala, ki ustreza spreminjanju frekvence srčnega utripa. Meritev napravimo v času, ko je opazovana oseba v popolnem mirovanju – v tem primeru je utrip konstanten, variacije v frekvenci pa so večinoma posledica napak v samodejni obdelavi signalov. Drugo meritev pa napravimo v času mirovanja po lažji telesni aktivnosti. V tem primeru se frekvenca srčnega utripa pričakovano s časom zmanjšuje.

5 Zaključek

V magistrski nalogi smo raziskovali možnosti uporabe numerične metode analize neodvisnih komponent (ICA) v fizikalnih aplikacijah. Metodo smo implementirali za reševanje različnih fizikalnih problemov in problemov obdelave signalov s fiziko v ozadju. Bistvo metode ICA je, da naredi razcep izmerjene linearne mešanice signalov na izvirne statistično neodvisne komponente.

Obravnavali smo preprost teoretični problem točkastih delcev s časovno spremenljivim nabojem, ki so bili na neznanih položajih v ravnini. Iz meritev električnega potenciala na znanih položajih smo s pomočjo metode ICA in numerične minimizacije natančno rekonstruirali izvirne signale, t.j. časovne odvisnosti nabojev na delcih. Poleg tega smo rekonstruirali tudi položaje teh delcev v ravnini. Sklepamo, da bi povsem analogno lahko s kombinacijo teh dveh metod reševali na primer tudi tridimenzionalne (geo)fizikalne probleme s področja magnetnega polja, pri katerih ne moremo neposredno opraviti meritev izvirnih signalov, lahko pa izmerimo različne linearne kombinacije teh izvirnih signalov. Metodo ICA torej lahko uporabimo pri meritvah vseh količin, za katere velja načelo linearne oz. šibko nelinearne superpozicije.

Ukvarjali smo se še z odstranjevanjem odbojev od dielektričnih površin iz fotografij. Pogosto se namreč zgodi, da je slika, ki jo zajamemo pri fotografiranju skozi steklo, sestavljena iz dveh komponent: iz slike, ki nastane zaradi odboja svetlobe od predmetov izza stekla, in iz slike, ki nastane zaradi odboja od steklene površine. Fizikalno smo utemeljili, zakaj v tem primeru za gostoto svetlobnega toka, ki ga izmerimo z merilnikom svetlobe, dovolj dobro drži linearni model – to pomeni, da je izmerjena gostota svetlobnega toka linearna superpozicija gostote svetlobnega toka iz izvira izza stekla in gostote svetlobnega toka iz izvira pred steklom. Ker je odbita svetloba delno polarizirana, so koeficienti v tej linearni kombinaciji odvisni od orientacije linearnega polarizatorja, ki ga namestimo pred objektiv merilnika svetlobe. Kadar je kot pri odboju žarka od dielektrične površine enak Brewsterjevemu kotu, je odbita svetloba celo popolnoma polarizirana, zato jo z linearnim polarizatorjem lahko povsem izločimo. Z našim poskusom, pri katerem smo uporabili linearni polarizator v kombinaciji z metodo ICA, pa smo pokazali, da se lahko deloma polarizirane odbite svetlobe znebimo pri skoraj poljubnih odbojnih kotih. Sliko, ki nastane zaradi odboja svetlobe od steklene površine, torej lahko na fotografijah

uspešno izločimo. Izjema so le primeri, ko se žarek svetlobe odbije od dielektrične površine pri kotih, ki so blizu 0° ali 90° glede na vpadno pravokotnico.

Nazadnje smo se ukvarjali še z optičnim merjenjem fizioloških signalov. Definirali in implementirali smo metodo, s katero je možno samodejno merjenje frekvence srčnega utripa z obraza opazovane osebe s preprostim merilnikom svetlobe – spletno kamero. Za potrebe te metode smo implementirali še algoritem za prepoznavanje obrazov in ga priredili našim potrebam. Merili smo torej spremembe v svetlobi, ki se z obraza opazovane osebe odbije v merilnik s tremi senzorji – po enim za rdečo, zeleno in modro svetlobo. Z metodo ICA smo iz treh izmerjenih signalov rekonstruirali izvirne signale, med katerimi se je nahajal tudi iskan pletizmografski signal, ki je posledica srčnega cikla. Z analizo frekvenčnih spektrov izvirnih signalov smo objektivno določili, kateri signal je pletizmografski. Nato smo poiskali na intervalu možnih frekvenc človeškega utripa frekvenco, pri kateri je pletizmografski signal najmočnejši – ta frekvenca pa ustreza frekvenci srčnega utripa. Tako smo pokazali, da je z metodo ICA možno poiskati tudi najbolj skrite signale.

Literatura

- [1] A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation*. (Society for Industrial and Applied Mathematics, Philadelphia, 2005).
- [2] J. Herault, C. Jutten in B. Ans, "Circuits neuronaux a synapses modifiables: décodage de messages composites par apprentissage non supervisé", C.-R. de l'Academie des Sciences **299**(III-13), 525 (1984).
- [3] P. Comon, "Independent component analysis – a new concept?", Signal Processing **36**, 287 (1994).
- [4] A. Hyvärinen in E. Oja, "Independent component analysis: algorithms and applications", Neural Networks **13**, 411-430 (2000).
- [5] R. Vigario, V. Jousamäki, M. Hämmäläinen, R. Hari in E. Oja, "Independent component analysis for identification of artifacts in magnetoencephalographic recordings", Advances in Neural Information Processing Systems **10**, 229 (1998).
- [6] T.P. Jung, S. Makeig, C. Humphries, T.W. Lee, M.J. McKeown, V. Iragui in T.J. Sejnowski, "Removing electroencephalographic artifacts by blind source separation", Psychophysiology **37**, 163 (2000).
- [7] T. He, G. Clifford in L. Tarassenko, "Application of independent component analysis in removing artefacts from the electrocardiogram", Neural Computing & Applications **15**, 105 (2006).
- [8] M.J. McKeown, S. Makeig, G.G. Brown, T.P. Jung, S.S. Kindermann, A.J. Bell in T.J. Sejnowski, "Analysis of fMRI data by blind separation into independent spatial components", Human Brain Mapping **6**, 160 (1998).
- [9] B. Azzerboni, G. Finocchio, M. Ipsale, F. La Foresta in F.C. Morabito, "A new approach to detection of muscle activation by independent component analysis and wavelet transform", Lecture Notes in Computer Science **2486**, 109 (2002).

- [10] H. Liang, "Adaptive independent component analysis of multichannel electrogastrograms", *Medical Engineering & Physics* **23**, 91 (2001).
- [11] A. Ciaramella, E. De Lauro, S. De Martino, B. Di Lieto in R. Tagliaferri, "Characterization of Strombolian events by using independent component analysis", *Nonlinear Processes in Geophysics* **11**, 453 (2004).
- [12] F. Aires, A. Chedin in J.-P. Nadal, "Independent component analysis of multivariate time series: Application to the tropical SST variability", *J. Geophys. Res.* **105**(D13), 17437 (2000).
- [13] M.J.D. Powell, "An efficient method for finding the minimum of a function of several variables without calculating derivatives", *Computer Journal* **7**, 155 (1964).
- [14] C.E. Cherry, "Some experiments on the recognition of speech, with one and with two ears", *The Journal of the Acoustical Society of America* **25**, 975 (1953).
- [15] F.R. Bach in M.I. Jordan, "Kernel independent component analysis", *The Journal of Machine Learning Research* **3**, 48 (2003).
- [16] S. Širca in A. Horvat, *Računske metode za fizike* (DMFA – založništvo, Ljubljana, 2010).
- [17] A. Hyvärinen, "Fast and robust fixed-point algorithms for independent component analysis", *IEEE Trans. Neural Networks* **10**, 626 (1999).
- [18] A. Papoulis, *Probability, random variables and stochastic processes*, 3rd edition (McGraw-Hill, New York, 1991).
- [19] J. Nocedal in S.J. Wright, *Numerical Optimization* (Springer, New York, 2006).
- [20] A.J. Bell in T.J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution", *Neural Computation* **7**, 1004 (1995).
- [21] J. Cardoso, "Source separation using higher order moments", *International Conference on Acoustics, Speech and Signal Processing*, 2109 (1989).
- [22] M. Yamazaki, Y.-W. Chen in G. Xu, "Separating Reflections from Images Using Kernel Independent Component Analysis", *IEEE International Conference on Pattern Recognition*, 194 (2006).
- [23] H. Farid in E. Adelson, "Separating Reflections and Lightning Using Independent Components Analysis", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 262 (1999).
- [24] H. Fujiwara, *Spectroscopic Ellipsometry; Principles and Applications* (John Willey & Sons, Ltd., 2007).
- [25] J. Strnad, *Fizika, 2. del; Električna, optika* (DMFA – založništvo, Ljubljana, 2005).
- [26] M. Born in E. Wolf, *Principles of Optics* (Cambridge University Press, Cambridge, 1999).

-
- [27] R.W.G. Hunt, *The Reproduction of Colour* (Wiley-IS&T Series in Imaging Science and Technology, Chichester, VB, 2004).
- [28] V.V. Tuchin, "Light scattering study of tissues", *Physics – Uspekhi* **40**, 495 (1997).
- [29] K. Shelley in S. Shelley, *Pulse Oximeter Waveform: Photoelectric Plethysmography in Clinical Monitoring* (W.B. Saunders Company, 2001).
- [30] K. Shelley, "Photoplethysmography: beyond the calculation of arterial oxygen saturation and heart rate", *Anesthesia & Analgesia* **105**, S31-6 (2007).
- [31] W. Verkruyse, L.O. Svaasand in J.S. Nelson, "Remote plethysmographic imaging using ambient light", *Optics Express* **16**, 21434 (2008).
- [32] M.-Z. Poh, D.J. McDuff in R.W. Picard, "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation", *Optics Express* **18**, 10762 (2011).
- [33] W.G. Zijlstra, A. Buursma in W.P. Meeuwse-van der Roest, "Absorption Spectra of Human Fetal and Adult Oxyhemoglobin, De-Oxyhemoglobin, Carboxyhemoglobin, and Methemoglobin", *Clinical Chemistry* **37**, 1633 (1991).
- [34] Human Hemoglobin Optical Characteristics. Dostopno na:
http://www.spectros.com/fileadmin/user_upload/library/Data_Hemoglobin_Standards.pdf
(25.6.2014).
- [35] K. Humphreys, T. Ward, C. Markham, "Noncontact simultaneous dual wavelength photoplethysmography: a further step toward noncontact pulse oximetry", *The Review of Scientific Instruments* **78**, 044304 (2007).
- [36] F.P. Wieringa, F. Mastik in A.F. van der Steen, "Contactless multiple wavelength photoplethysmographic imaging: a first step toward "SpO₂ camera" technology", *Annals of Biomedical Engineering* **33**, 1034 (2005).
- [37] P. Viola in M.J. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features", *IEEE International Conference on Computer Vision and Pattern Recognition* (2001).
- [38] R. Lienhart in J. Maydt, "An Extended Set of Haar-like Features for Rapid Object Detection", *IEEE International Conference on Image Processing* (2002).
- [39] Y. Freund in R.E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting", *Computational Learning Theory: Eurocolt '95*, 23-36 (Springer-Verlag, 1995).
- [40] Y. Freund in R.E. Schapire, "A Short Introduction to Boosting", *Journal of Japanese Society for Artificial Intelligence* **14**, 771 (1999).
- [41] G. Bradski, "The opencv library", *Doctor Dobbs Journal* **25**, 120-126 (2000).
- [42] C.A. Poynton, *Digital Video and HDTV: Algorithms and Interfaces* (Morgan Kaufmann, 2003).

- [43] CIE, *Commission internationale de l'Eclairage proceedings, 1931* (Cambridge University Press, Cambridge, 1932)
- [44] T. Smith in J. Guild, "The C.I.E. colorimetric standards and their use", *Transactions of the Optical Society* **33**, 73 (1932).

A Prepoznavanje objektov na sliki

Pri implementaciji metode za zaznavanje frekvence srčnega utripa uporabimo algoritem, ki na vsaki sliki v vhodnem video-signalu natančno določi lege delov telesa, ki jih želimo spremljati. V našem primeru za spremljanje tega fiziološkega procesa sledimo celotnemu obrazu. Za prepoznavanje obraza na slikah uporabimo algoritem za prepoznavanje objektov, ki sta ga razvila Viola in Jones [37], pozneje pa sta ga izboljšala Lienhart in Maydt [38]. Algoritem se angleško imenuje *Haar Feature-based Cascade Classifier for Object Detection*. To je bil prvi algoritem, s katerim je bilo možno v realnem času prepoznavati tako zapletene grafične objekte, kot so npr. človeški obrazi. Hkrati pa je pri prepoznavanju delov človeškega telesa tudi najbolj zanesljiv. Ker uspešnost prepoznavanja posameznih objektov na sliki pomembno vpliva na uspešnost metode, opisane v magistrskem delu, algoritem v nadaljevanju podrobneje predstavimo.

Algoritem spada v skupino učnih algoritmov (algoritmov za strojno učenje, ang. *machine learning algorithms*). Ti se znajo učiti na osnovi množice vhodnih podatkov. Vsak vhodni podatek nosi neke lastnosti. Produkt tega učenja je model znanja (klasifikator). Ta zna tudi za nove, še nepoznane vhodne podatke na podlagi njihovih lastnosti odločiti o vrsti (razredu) podatka.

Algoritem za prepoznavanje objektov na sliki je zato potrebno najprej naučiti. V našem primeru prepoznavanja obrazov je algoritem naučen na nekaj tisoč pozitivnih primerih (pravih slikah obrazov) in na podobnem številu negativnih primerov, ki niso slike obrazov. Vse slike za učenje so skalirane na isto standardno velikost 24×24 ločljivostnih točk (pikslov) in transformirane v črnobeli barvni prostor (ang. *grayscale*) (podrobnosti o transformaciji v dodatku C). Vsaka vrednost piksla je tako mera za relativno svetlost oz. intenziteto svetlobe. Ko je klasifikator naučen, ga lahko apliciramo na neko še nepoznano testno sliko, ki jo prej seveda preslikamo v črnobeli barvni prostor. Testno sliko pregledujemo z iskalnim oknom. Klasifikator vrne vrednost 1 (True), če je dovolj verjetno, da iskalno okno vsebuje objekt (obraz). To pomeni, da je klasifikator na sliki prepoznal iskan objekt. V nasprotnem primeru, torej če klasifikator v iskalnem oknu ne prepozna objekta, vrne vrednost 0 (False). Iskanje objekta poteka tako, da po sliki premikamo iskalno okno in pri vsaki poziciji iskalnega okna klasifikator skuša prepoznati objekte. Prav tako lahko klasifikator s "približevanjem" ali "oddaljevanjem" apliciramo na manjše ali večje območje slike. Kakšne velikosti bodo objekti v primerjavi z

velikostjo slike namreč vnaprej ne vemo, zato moramo sliko pregledati z različno velikimi iskalnimi okni. Pri tem algoritmu kot vhodni podatek podamo minimalno in maksimalno velikost iskalnega okna ter faktor, za katerega zmanjšamo površino iskalnega okna na vsakem koraku pregledovanja slike.

Končni klasifikator v nekem iskalnem oknu sestavlja v resnici kaskada preprostejših klasifikatorjev. To pomeni, da klasifikatorje enega za drugim (po stopnjah kaskade) apliciramo na območje v iskalnem oknu. Postopek ustavimo, ko je na neki stopnji klasifikator zavržen ali pa smo pregledali že vse. Ti preprostejši klasifikatorji, ki jih uporabimo na vsaki stopnji kaskade so tudi sami kompleksni in sestavljeni iz še preprostejših – osnovnih klasifikatorjev. Ti na podlagi potez (ang. *features*) območja slike znotraj iskalnega okna ocenijo, ali območje vsebuje objekt ali ne. Zgoraj smo torej opisno predstavili korake algoritma za prepoznavanje objektov. V nadaljevanju posamezne korake algoritma še podrobneje predstavimo.

A.1 Haarove poteze

Poteze območja slike želimo kvantitativno numerično ovrednotiti. Osnovne poteze močno spominjajo na Haarove valčne funkcije¹. Zato tem potezam pravimo tudi Haarove poteze. Uporabljamo tri vrste pravokotnih potez (slika A.1). Vrednost poteze, ki je sestavljena iz dveh pravokotnikov (ang. *two-rectangle feature*), dobimo tako, da izračunamo razliko med vsotama vrednosti pikslov v obeh pravokotnikih. Pravokotnika sta enako velika in tako horizontalno kot vertikalno poravnana. Vrednost poteze, ki je sestavljena iz treh pravokotnikov (ang. *three-rectangle feature*), izračunamo tako, da vsoti zunanjih pravokotnikov odštejemo od vsote osrednjega pravokotnika. Vrednost poteze, ki je sestavljena iz štirih pravokotnikov (ang. *four-rectangle feature*), pa izračunamo tako, da med seboj odštejemo vsoti diagonalnih pravokotnikov. Če označimo pravokotnike kot je prikazano na sliki A.1, lahko zapišemo splošno formulo za vrednost V neke poteze takole

$$V = \sum_{(x,y) \in \text{siv}} i(x,y) - \sum_{(x,y) \in \text{bel}} i(x,y), \quad (\text{A.1})$$

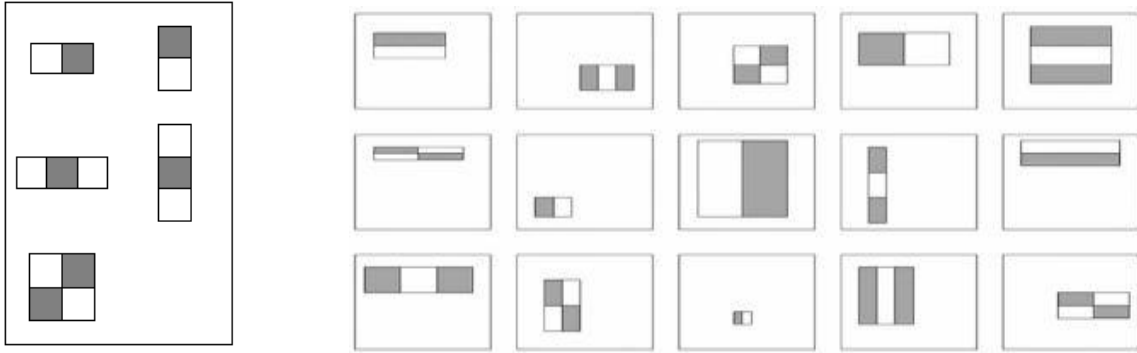
kjer je $i(x,y)$ vrednost piksla na poziciji (x,y) .

Če je velikost iskalnega okna 24×24 pikslov, to pomeni, da je vseh potez skupaj več kot 160000. Vse te poteze seveda ne predstavljajo baze – ta sestoji le iz 576 potez (kolikor je pikslov). Na preprostem primeru iskalnega okna velikosti 4×4 pikslov pokažimo, kako smo prišli do tega števila vseh potez. Vseh možnih gradnikov, s katerimi lahko sestavimo neko

¹ V matematiki so Haarovi valčki vrsta funkcij pravokotnih oblik, ki skupaj tvorijo ortonormalno bazo. Poljubno funkcijo $f(x)$ potem lahko zapišemo z vrsto $f(x) = c_0 + \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} c_{jk} \psi_{jk}(x)$, kjer velja

$$\psi_{jk}(x) = \psi(2^j x - k) \text{ za } j \geq 0 \text{ in } 0 \leq k \leq 2^j - 1, \text{ pri čemer je osnovni val } \psi(x) = \begin{cases} 1; & 0 \leq x < \frac{1}{2}, \\ -1; & \frac{1}{2} < x \leq 1, \\ 0; & \text{sicer.} \end{cases}$$

Haarova vrsta je poimenovana po madžarskem matematiku Alfredu Haaru, ki jo je prvi opisal leta 1909.



Slika A.1. [LEVO] Osnovni gradniki pravokotnih potez. Osnovni gradniki so velikosti 1×2 , 2×1 , 1×3 , 3×1 in 2×2 . [DESNO] Primeri pravokotnih potez v iskalnem oknu. Vrednost poteze izračunamo tako, da vsoto pikslov v belih pravokotnikih odštejemo od vsote pikslov v sivih pravokotnikih.

potezo, je pet (slika A.1). To so gradniki velikosti 1×2 , 2×1 , 1×3 , 3×1 in 2×2 . Eno potezo lahko zgradimo samo z gradniki iste velikosti. Na primer z gradnikom 1×2 lahko v 4×4 iskalnem oknu zgradimo 12 različnih potez velikosti 1×2 , lahko zgradimo 4 različne poteze velikosti 1×4 , lahko 9 različnih potez velikosti 2×2 itn. Vseh možnih potez je torej že za majhno 4×4 iskalno okno kar 136. Ker je število potez pri večjih iskalnih oknih ogromno, je računsko zahtevno določiti vrednosti V vseh potez. Veliko večino jih moramo zato ovreči že prej in uporabiti le tiste, ki so pomembne pri procesu opisovanja objekta. Ta izjemno pomemben proces brez katerega sprotna detekcija objektov ne bi bila mogoča, opišemo v naslednjih razdelkih.

Opišimo sedaj postopek, po katerem vrednosti V pravokotnih potez izjemno hitro izračunamo. Seštevanje vrednosti pikslov po enačbi (A.1) bi bilo namreč zelo zamudno. Zato uvedemo integralno sliko (ang. *integral image*). Integralna slika na lokaciji (x, y) je vsota vrednosti vseh pikslov levo in navzgor od točke (x, y) , torej

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y'), \quad (\text{A.2})$$

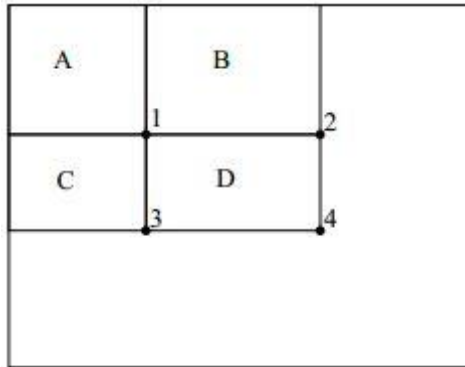
kjer je $ii(x, y)$ integralna slika in $i(x, y)$ originalna slika. Če uporabimo sledeč par zaporedij

$$s(x, y) = s(x, y - 1) + i(x, y), \quad (\text{A.3})$$

$$ii(x, y) = ii(x - 1, y) + s(x, y), \quad (\text{A.4})$$

kjer je $s(x, y)$ kumulativna vsota vrednosti stolpca pod (x, y) in velja $s(x, -1) = 0$ ter $ii(-1, y) = 0$, potem lahko integralno sliko izračunamo v enem samem "sprehodu" čez originalno sliko.

Bistvena prednost integralne slike je torej, da lahko vsoto vrednosti v katerem koli pravokotniku v potezi izračunamo s samo štirimi klici integralne slike (slika A.2). Vrednost poteze, ki je sestavljena iz dveh pravokotnikov, izvednotimo s šestimi klici. Nadalje potrebujemo osem klicev za potezo iz treh pravokotnikov in devet klicev za potezo iz štirih pravokotnikov.



Slika A.2. Vsoto pikslov znotraj pravokotnika D lahko izračunamo s samo štirimi klici na integralne slike. Vrednost integralne slike na lokaciji 1 je vsota vseh pikslov v pravokotniku A. Vrednost na lokaciji 2 je A + B, vrednost na lokaciji 3 je A + C, vrednost na lokaciji 4 pa je A + B + C + D. Vsoto pikslov v pravokotniku D preprosto izračunamo kot $4 + 1 - 2 + 3$ [37].

A.2 Učenje klasifikatorjev

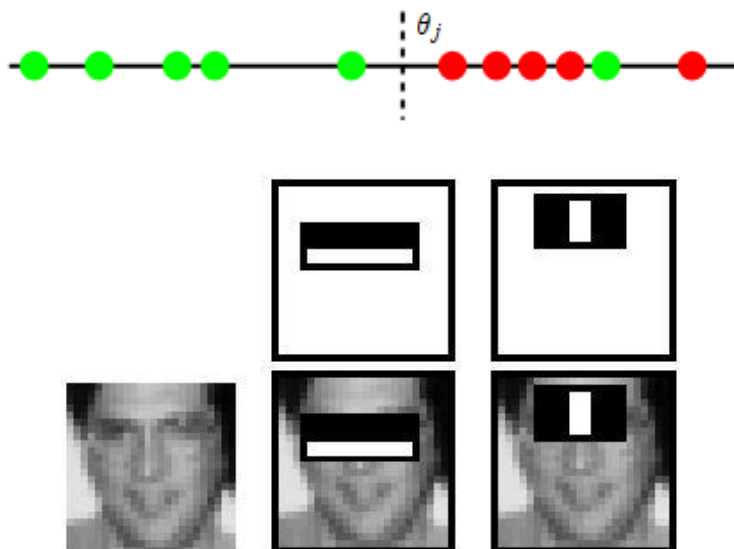
Ko imamo nabor potez in nabor vhodnih pozitivnih (vsebujejo iskan objekt) in negativnih (ne vsebujejo iskanega objekta) primerov, lahko algoritem naučimo. V magistrskem delu je bil za učenje izbran algoritem AdaBoost [39]. Ime zanj izhaja iz angleškega izraza *boosting*. Ta označuje metodo, ki lahko na podlagi množice nenatančnih in komaj uporabnih podatkov napravi zelo natančno napoved. V našem primeru prepoznavanja objektov je informacija, ki jo dobimo od posamezne poteze v iskalnem oknu, dokaj neuporabna, že nekaj potez pa lahko s pomočjo AdaBoost algoritma povežemo v natančen in učinkovit klasifikator. Pri tem je glavni problem poiskati tiste poteze, ki nosijo največ informacije.

To storimo tako, da med vsemi potezami poiščemo tiste, ki najboljše ločijo pozitivne in negativne primere. Za vsako potezo poiščemo tako klasifikacijsko funkcijo (slika A.3), da bo število napačno uvrščenih (klasificiranih) primerov minimalno. Zapišemo torej šibki klasifikator h_j (direktni prevod iz ang. *weak classifier*)

$$h_j(x) = \begin{cases} 1; & p_j f_j(x) < p_j \theta_j, \\ 0; & \text{sicer,} \end{cases} \quad j = 1, \dots, n, \quad (\text{A.5})$$

kjer je f_j poteza, x iskalno okno velikosti 24×24 , θ_j mejna vrednost, na podlagi katere uvrstimo posamezen primer v eno izmed obeh skupin, p_j določa neenačaj, n pa je število vseh potez. Vsak šibki klasifikator h_j potem uporabimo na naboru vhodnih pozitivnih in negativnih primerov. Za opis iskanega objekta so nam torej zanimive le tiste poteze f_j , pri katerih je pripadajoč šibki klasifikator h_j napačno uvrstil najmanj primerov. Te šibke klasifikatorje potem povežemo v močni klasifikator (ang. *strong classifier*). Ta je kar linearna kombinacija šibkih klasifikatorjev

$$H(x) = w_1 h_1(x) + w_2 h_2(x) + w_3 h_3(x) + \dots \quad (\text{A.6})$$



Slika A.3. [ZGORAJ] Primer klasifikacijske funkcije, ki z mejno vrednostjo θ_j razdeli množico zelenih (pozitivnih) in rdečih (negativnih) primerov tako, da je število napačno uvrščenih minimalno. [SPODAJ] Prva in druga najpomembnejša poteza za primer prepoznavanja obrazov, izbrana objektivno z algoritmom AdaBoost. Prva poteza meri razliko v intenziteti svetlobe med območjem oči in območjem zgornjega dela lic. Ta poteza je logična izbira – območje oči je namreč pogosto temnejše kot območje lic. Druga poteza pa meri razliko v intenziteti svetlobe med območjem oči in delom nosa – ta del je običajno svetlejši od dela, kjer se nahajajo oči in obrvi.

Zaradi računske učinkovitosti želimo seveda močni klasifikator sestaviti iz čim manj šibkih klasifikatorjev, saj to pomeni manj vrednotenja potez. Kako poteka izbiranje ustreznih potez, matematično opišemo v psevdokodi [39, 40]:

- Primeri slik $(x_1, y_1), \dots, (x_n, y_n)$, kjer je $y_i = \{0, 1\}$ za negativen in pozitiven primer, x_i pa je slika neke velikosti (v našem primeru 24×24 pikslov).
- Enakomerno utežimo vse primere: pozitivne z $w_{1,i} = \frac{1}{2a}$ in negativne z $w_{1,i} = \frac{1}{2b}$ za $y_i = \{0, 1\}$, kjer je a število pozitivnih in b število negativnih primerov.
- Za vkorak $t = 1, \dots, T$:
 1. Normaliziramo uteži: $w_{t,i} \leftarrow \frac{w_{t,i}}{\sum_{k=1}^n w_{t,k}}$.
 2. Za vsako potezo j naučimo šibki klasifikator h_j . Uteženo napako izračunamo kot $\varepsilon_j = \sum_i w_{t,i} |h_j(x_i) - y_i|$.
 3. Izberemo klasifikator z najmanjšo napako – označimo ga s h_t , napako pa z ε_t .
 4. Na novo izračunamo uteži: $w_{t+1,i} = w_{t,i} \beta_t^{1-e_i}$, kjer je $e_i = 0$, če je x_i uvrščen pravilno, in $e_i = 1$ sicer. $\beta_t = \frac{\varepsilon_t}{1-\varepsilon_t}$. Uteži pravilno uvrščenih primerov i se torej zmanjšajo, uteži napačno uvrščenih pa ostanejo enake (oz. se relativno glede na pravilno uvrščene povečajo). To pomeni, da v naslednjih stopnjah t dajemo več poudarka razvrstitvi napačno uvrščenih primerov.

- Končni močni klasifikator je tako: $H(x) = \begin{cases} 1; & \sum_{t=1}^T \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t, \\ 0; & \text{sicer.} \end{cases}$

Pri tem je $\alpha_t = \log \frac{1}{\beta_t}$.

Na vsakem koraku t torej izberemo eno potezo izmed več kot 160000 možnih potez (v primeru iskalnega okna velikosti 24×24 pikslov). Na sliki A.3 prikažemo še primer prve in druge najpomembnejše poteze pri detekciji obrazov, kot ju dobimo z algoritmom AdaBoost. Če izberemo število vseh potez v močnem klasifikatorju $T = 200$, potem je uspešnost prepoznavanja objektov 95 %. Verjetnost za napačno prepoznavanje, torej da prepoznamo del slike kot objekt, čeprav to ni (ang. *false positive rate*), pa je samo 0.7 %.

Klasifikacija iskalnega okna ("okno vsebuje objekt" ali "okno ne vsebuje objekta") na podlagi močnih klasifikatorjev je hitra, vendar še vedno ne dovolj za prepoznavanje objektov v realnem času. V nadaljevanju zato opišemo še, kako te močne klasifikatorje povežemo v kaskado – uporabljamo večstopenjski sistem za odkrivanje objektov na sliki. Tak sistem močno izboljša detekcijo in hkrati skrajša računski čas.

A.3 Kaskada klasifikatorjev

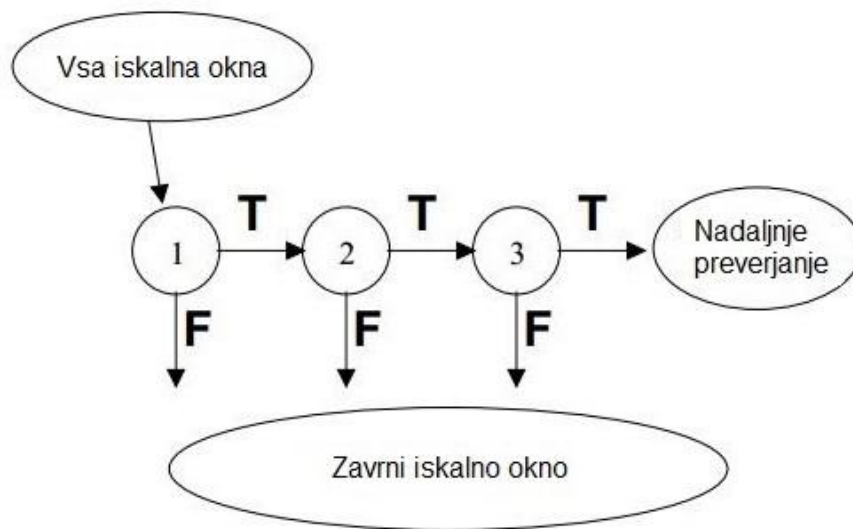
Postopek, s katerim močno zmanjšamo računski čas in hkrati izboljšamo uspešnost prepoznavanja objektov na sliki, je povezovanje močnih klasifikatorjev v kaskado, t.j. večstopenjsko odločitveno drevo (slika A.4). Glavna ideja tega je, da na začetnih stopnjah kaskade skonstruiramo take preproste (iz majhnega števila potez) močne klasifikatorje, ki že takoj zavrnejo ogromno iskalnih oken (iskani objekt naj se tam nebi nahajal). Primer takega preprostega klasifikatorja je klasifikator iz samo dveh potez (slika A.3). Ta napačno izloči približno 0.2% pozitivnih primerov (čeprav ti vsebujejo iskani objekt) in napačno sprejme približno 40% negativnih primerov (čeprav ti ne vsebujejo iskanega objekta). Seveda so napačno sprejeti negativni primeri ovrženi v naslednjih stopnjah kaskade. Šele nato na preostalih iskalnih oknih (kjer je možno, da se objekt nahaja) uporabimo zapletenejše (iz večjega števila potez in zato tudi računsko zahtevnejše) močne klasifikatorje.

Močni klasifikator na vsaki stopnji i kaskade naučimo tako, da bo dosegel željeno uspešnost prepoznavanja D_i in željeno verjetnost napačnega prepoznavanja F_i objektov. To storimo tako, da vsakemu močnemu klasifikatorju dodajamo poteze, dokler ne doseže željenih vrednosti.

Končni klasifikator je seveda kombinacija močnih klasifikatorjev na vseh stopnjah kaskade. Njegovo uspešnost D prepoznavanja objektov in verjetnost F za napačno prepoznavo (prepoznan del slike v resnici ni iskan objekt) potem lahko zapišemo kot

$$D = \prod_i^K D_i, \quad F = \prod_i^K F_i, \quad (\text{A.6})$$

kjer je K število vseh stopenj kaskade. Tako smo opisali celoten algoritem za prepoznavanje obrazov. Posamezni deli algoritma so implementirani v knjižnici OpenCV [41], v magistrskem delu pa jih povežemo v delujočo celoto in priredimo algoritem našim potrebam.



Slika A.4. Shematični prikaz kaskade močnih klasifikatorjev. Številke pomenijo stopnje kaskade – vsaka ustreza nekemu močnemu klasifikatorju. Na prvih stopnjah zavrnemo ogromno iskalnih oken (v njih naj se iskan objekt ne bi nahajal). Klasifikatorje na teh prvih stopnjah kaskade angleško označujemo z *attentional operators*. Izraz pomeni, da po teh stopnjah ostanejo le še tista iskalna okna, ki pritegnejo pozornost algoritma – na njih potem izvedemo nadaljnje stopnje kaskade z zapletenejšimi močnimi klasifikatorji. [37]



Slika A.5. [LEVO] Nabor pozitivnih primerov, na podlagi katerih naučimo algoritem za primer prepoznavanja obrazov. [DESNO] Prepoznavanje obrazov na naključni sliki. [37]

B Popravek gama

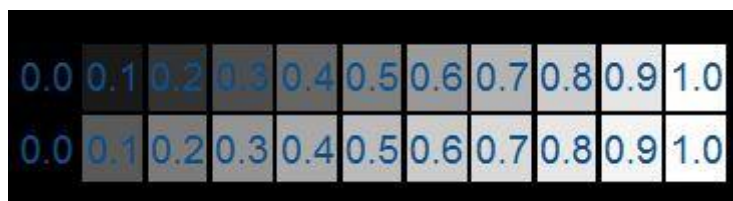
Sliko predstavimo računalniško kot matriko pikslov. V črno-belem barvnem prostoru je vrednost vsakega piksla $\tilde{y} \in [0,1]$. Vrednosti teh pikslov ustrezajo s popravkom gama nelinearno spremenjenim relativnim svetlostim y . Namen popravka je zagotoviti, da se numerične vrednosti $\tilde{y}$ spreminjajo enakomerno s spreminjanjem svetlosti, kot jo zazna človeško oko (slika B.1). V obravnavanih poskusih v magistrskem delu želimo delati zgolj z relativnimi svetlostmi – te so namreč linearno povezane z gostoto svetlobnega toka (dodatek C). Tudi metodo ICA uporabljamo zgolj z relativnimi svetlostmi. Vrednosti $\tilde{y}$ zato želimo transformirati v relativne svetlosti y . Za potrebe prikaza slik na ekranu pa nato želimo transformirati relativno svetlost y nazaj v $\tilde{y}$. Zapišemo obe transformaciji [42]. Transformacija $\tilde{y} \rightarrow y$ je enaka:

$$y = \begin{cases} \tilde{y}/b; & \tilde{y} \leq \tilde{y}_M, \\ \left(\frac{\tilde{y} + a}{1 + a}\right)^{2.4}; & \tilde{y} > \tilde{y}_M, \end{cases} \quad (\text{B.1})$$

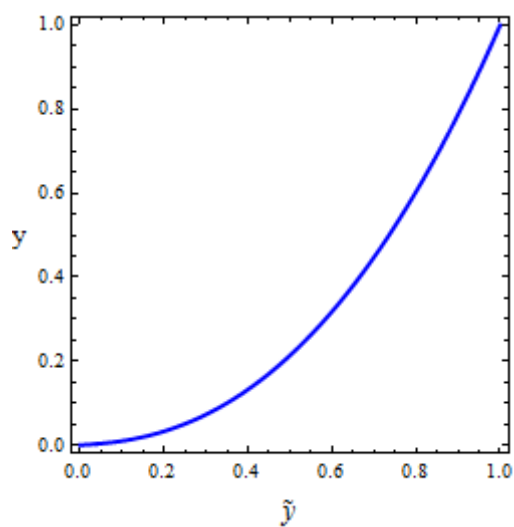
transformacija $y \rightarrow \tilde{y}$ pa je:

$$\tilde{y} = \begin{cases} by; & y \leq y_M, \\ (1 + a) y^{1/2.4} - a; & y > y_M, \end{cases} \quad (\text{B.2})$$

kjer so $a = 0.055$, $b = 12.92$ in mejni vrednosti $\tilde{y}_M = 0.04045$ ter $y_M = 0.0031308$. Zveza med y in $\tilde{y}$ je prikazana na sliki B.2.



Slika B.1. [ZGORAJ] spreminjanje svetlosti z vrednostjo $\tilde{y}$ in [SPODAJ] spreminjanje svetlosti z vrednostjo relativne svetlosti y . Na sliki se jasno vidi, da se svetlost, kot jo subjektivno zazna človeško oko, linearno spreminja z vrednostjo.



Slika B.2. Nelinearna relacija med vrednostjo $\tilde{y}$ in relativno svetlostjo y , ki jo uporabljamo v naših izračunih.

C Preslikava gostote svetlobnega toka v relativno svetlost

Gostota svetlobnega toka j o naravi svetlobe ne pove dovolj. Treba je navesti še njeno porazdelitev po frekvencah ν ali po valovnih dolžinah λ . Spektri svetlobe, s kakršno imamo opravka v naših poskusih, so običajno zvezni. Skupna gostota svetlobnega toka je torej

$$j = \int \frac{dj_\lambda}{d\lambda} d\lambda = \int \frac{dj_\nu}{d\nu} d\nu. \quad (\text{C.1})$$

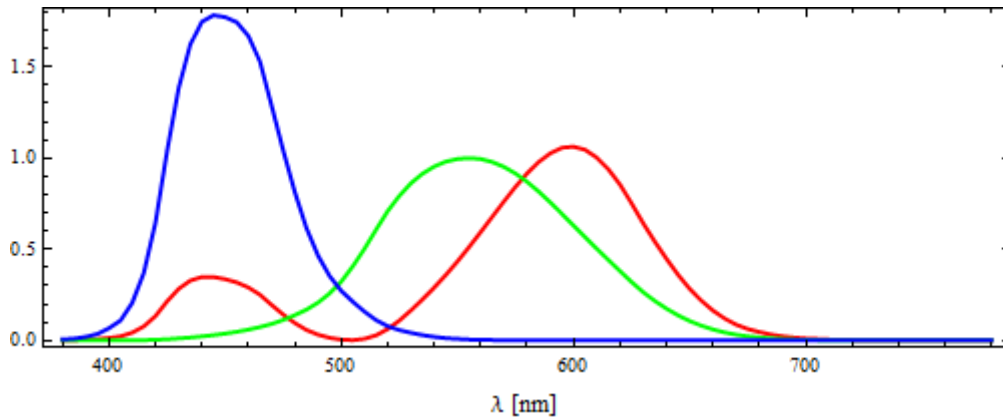
Standardni barvni prostori preslikajo izmerjeno spektralno porazdelitev gostote svetlobnega toka $dj_\lambda/d\lambda$ v tri stimulative vrednosti, katerih linearna kombinacija opiše vse možne barve, ki jih lahko zazna povprečno človeško oko. Tri barvne vrednosti ustrezajo trem različnim vrstam fotopičnih sprejemnikov svetlobe (čepkov) v človeškem očesu, ki imajo različno relativno občutljivost na enobarvno svetlobo z določeno valovno dolžino. Z mednarodnim dogovorom je bil leta 1931 na podlagi poskusov na množici oseb sprejet barvni odziv standardnega opazovalca na svetlobo z določeno spektralno porazdelitvijo gostote svetlobnega toka [43]. Določene so bile t.i. odzivne funkcije $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ in $\bar{z}(\lambda)$, s katerimi numerično opišemo barvni odziv standardnega opazovalca (slika C.1). Lahko si jih predstavljamo kot odzivne krivulje (ang. *sensitivity curves*) treh linearnih detektorjev svetlobe, ki preslikajo spektralno porazdelitev gostote svetlobnega toka v tri stimulative vrednosti X , Y in Z :

$$X = \int_{\lambda_1}^{\lambda_2} \frac{dj_\lambda}{d\lambda} \bar{x}(\lambda) d\lambda, \quad (\text{C.2a})$$

$$Y = \int_{\lambda_1}^{\lambda_2} \frac{dj_\lambda}{d\lambda} \bar{y}(\lambda) d\lambda, \quad (\text{C.2b})$$

$$Z = \int_{\lambda_1}^{\lambda_2} \frac{dj_\lambda}{d\lambda} \bar{z}(\lambda) d\lambda, \quad (\text{C.2c})$$

kjer je λ [nm] valovna dolžina monokromatske svetlobe in meji integrala $\lambda_1 = 380$ nm ter



Slika C.1. Odzivne funkcije (rdeča) $\bar{x}(\lambda)$, (zelena) $\bar{y}(\lambda)$ in (modra) $\bar{z}(\lambda)$, s katerimi numerično opišemo barvni odziv standardnega opazovalca. Definijsko območje odzivnih funkcij je $\lambda \in [380, 780]$ nm, funkcije pa so diskretizirane na $\Delta\lambda = 5$ nm. Odzivna funkcija $\bar{y}(\lambda)$ ustreza kar relativni občutljivosti očesa pri fotopičnem gledanju [25].

$\lambda_2 = 780$ nm, ki ustrežata skrajnim valovnim dolžinam svetlobe, ki jih človeško oko še zazna. Barvne funkcije so diskretizirane na intervalih $\Delta\lambda = 5$ nm. Vrednost Y ustreza svetlosti L (ang. *luminance*)

$$Y = L = \frac{dj_f}{d\Omega}, \quad [L] = \text{cd/m}^2, \quad (\text{C.3})$$

kjer je j_f gostota svetlobnega toka v fiziološkem merilu

$$j_f = (683 \text{ lm/W}) \int_0^\infty \frac{dj_\lambda}{d\lambda} \bar{y}(\lambda) d\lambda, \quad [j_f] = \text{lm/m}^2 \quad (\text{C.4})$$

Vrednost Z ustreza odzivu na modro svetlobo, vrednost X pa je linearna kombinacija odzivnih funkcij vseh treh vrst čepkov. Opisali smo torej CIE XYZ barvni prostor [44].

Relativno svetlost oz. intenziteto svetlobe v črnobelem barvnem prostoru nato izračunamo kot $y = Y/(X + Y + Z)$. Tako smo pokazali, da je preslikava gostote svetlobnega toka j v relativno svetlost y v črnobelem barvnem prostoru linearna, zato ob predpostavki, da velja model (3.11), sledi tudi veljavnost modela (3.12). Ker je tudi preslikava iz CIE XYZ barvnega prostora v CIE RGB barvni prostor linearna,

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 0.41847 & -0.15866 & -0.082835 \\ -0.091169 & 0.25243 & 0.015708 \\ 0.00092090 & -0.0025498 & 0.17860 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}, \quad (\text{C.5})$$

to pomeni, da je prav tako linearna preslikava iz gostote svetlobnega toka j v relativne svetlosti y_R, y_G in y_B osnovnih barv.