
Petar Janićijević

Određivanje smera pogleda korišćenjem jedne fiksne kamere

Predstavljen je sistem za određivanje smera pogleda korisnika korišćenjem samo jedne RGB kamere kakva se može naći u svakodnevnoj upotrebi na kompjuterima, laptopovima i telefonima. Jedan od najvažnijih koraka u radu sistema je određivanje centra zenice na slici, za šta su predstavljene dve metode: centra mase i aproksimacije kruga. Tačnost ovih metoda je ispitana na BioID bazi sa po dva praga tačnosti i dobijeni rezultati iznose 95.5% i 58.4% za prvu, odnosno 92.2% i 64.2% za drugu metodu. Greška konačnog rezultata sistema tj. smera pogleda dobijena je korišćenjem Eye of the Typer baze i iznosi 4.46° kada se koristi metod centra mase, odnosno 5.6° za metod aproksimacije kruga.

Uvod

Sistemi za određivanje smera pogleda korisnika već duže vreme postoje i našli su široku primenu. Određivanje i praćenje smera pogleda se primenjuje u raznim oblastima kao što su interakcija ljudi i kompjutera (kucanje na tastaturi pogledom, upravljanje mišem), psihološka istraživanja o tome kako ljudi posmatraju svet, marketing, neurologija i dijagnoza kliničkih oboljenja. Zavisno od toga kakva je primena, ovi sistemi se mogu izraditi na razne načine. Postoje invazivne, za koje korisnik mora imati elektrode tj. elektrookulografija, i neinvazivne metode, koje pasivno određuju smer pogleda (Chennamma, 2013). Neinvazivni sistemi najčešće koriste jednu ili više kamere. Kamere koje se koriste mogu biti infracrvene kamere ili RGB kamere. Ukoliko se koristi infracrvena kamera potrebno je imati i poseban izvor infracrvene svetlosti.

Ovaj rad se bavi izradom i ispitivanjem neinvazivnog sistema za određivanje smera pogleda

koji koristi jednu RGB kameru, kakva se može naći u svakodnevnoj upotrebi na laptopovima, kompjuterima i telefonima. Ovaj metod omogućava da sistem bude pristupačan za bilo koga ko poseduje uređaj sa kamerom, jer se može podeliti u obliku softvera. Sistem je realizovan sklopom nekoliko odvojenih koraka čiji se rad može pojedinačno ispitivati. Pre korišćenja sistema potrebno je obaviti kalibraciju koja se sastoji iz gledanja u devet tačaka na različitim delovima ekrana. Cilj ovog rada je da se utvrdi preciznost i tačnost sistema, kao i upoređivanje sa uspešnošću ranijih, sličnih sistema.

Metod

Rad sistema je podeljen u nekoliko koraka (slika 1). Prvi deo sistema je određivanje vektora oka. Prvobitno se odrede ključne tačke na licu, od kojih su najznačajnije one oko očiju, a zatim se primenjuje jedna od dve metode za nalaženje centra zenice tj. metod centra mase ili metod aproksimacije kruga. Nakon toga se definiše vektor od unutrašnje ivice oka do centra zenice što predstavlja vektor oka. Sledeći deo sistema je mapiranje vektora oka na tačku na ekranu. Ovaj deo se postiže funkcijom mapiranja koja se dobija u koraku kalibracije pre korišćenja sistema. U toku kalibracije se beleže koordinate prikazanih tačaka kao i vektora oka pri pogledu u te tačke, pa se definiše veza između njih funkcijom mapiranja. Dok se koristi sistem, funkcija mapiranja može na osnovu bilo kog vektora oka dati tačku na ekranu u koju korisnik gleda.

Detekcija lica i ključnih tačaka

Primenom biblioteke *Dlib* (King, 2009) se određuje deo slike u kom se nalazi lice, ukoliko postoji, kao i pozicija 68 ključnih tačaka na licu. Ovo se postiže pozivanjem funkcija koje koriste trenirane neuronske mreže kako bi se lako i efikasno odredile željene informacije. Za lice se

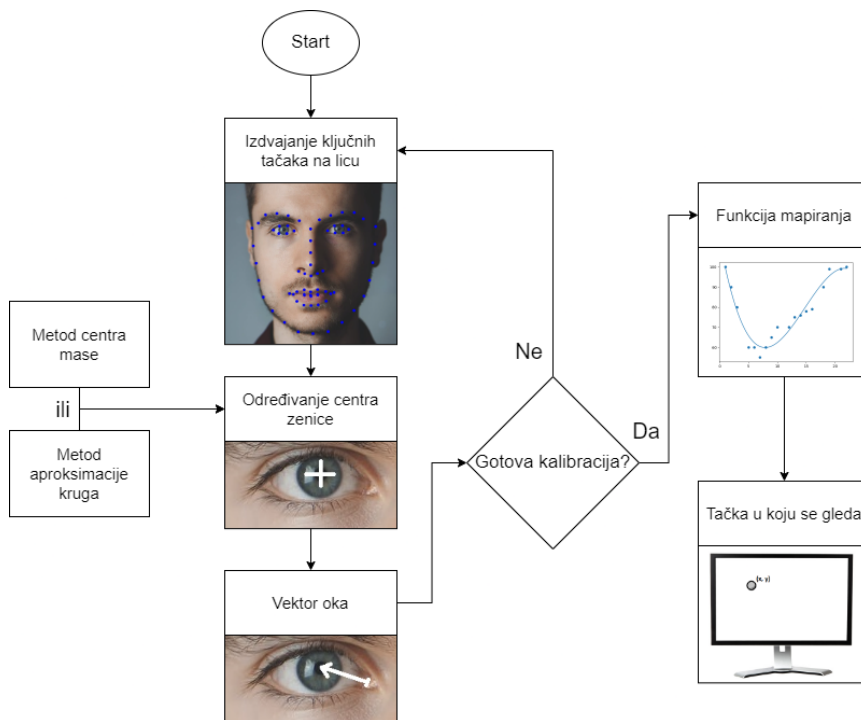
Petar Janićijević (2001), Kruševac, učenik 4. razreda Gimnazije u Kruševcu

MENTORI:

Lazar Milenković, Univerzitet u Tel Avivu

Teo Šarkić, Microsoft Development Center Serbia

Ervin Seke, Elektrotehnički fakultet Univerziteta u Beogradu



Slika 1: Grafički prikaz koraka u radu sistema

Figure 1: Graphical representation of the steps involved in the system



Slika 2: Ključne tačke na licu

Figure 2: Facial landmark points

dobiju koordinate okvira lica, a za ključne tačke se dobiju koordinate za svaku od 68 tačaka (slika 2). Od tih tačaka nama su najznačajnije tačke 36-47 tj. tačke koje obeležavaju kapke očiju.

Nalaženje centra zenice

Određivanje centra zenice je jedan od najbitnijih koraka za celokupan rad sistema jer mala greška u ovom koraku može izazvati veliku grešku u konačnom rezultatu. Zato se za ovaj korak predlažu dve metode. Prva metoda je nalaženje centra zenice korišćenjem centra mase binrizovane slike, dok je druga metoda dopuna nad prvom i pokušaj da se poboljša rezultat. Inherentni problem pristupa sa kamerom udaljenom od korisnika jeste da postoji relativno mala količina informacija u regionu oka tako da je, u

većini slučajeva, teško razlikovati samu zenicu od dužice koja je okružuje. Iz tog razloga se centar zenice aproksimira kao centar dužice.

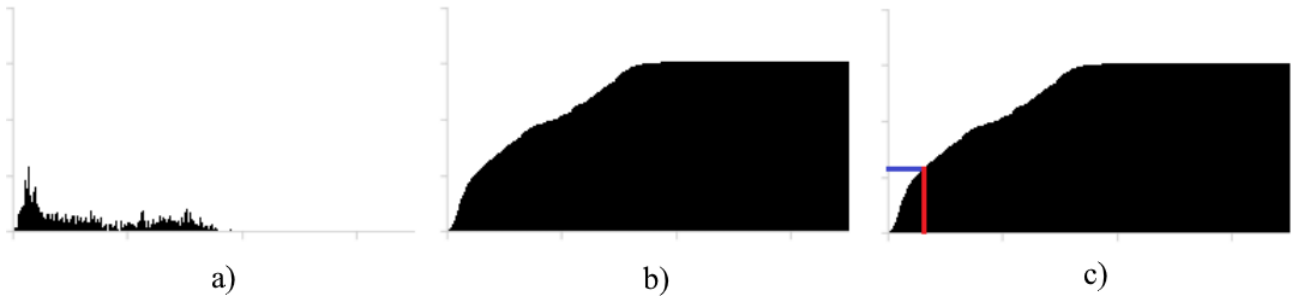
Metod centra mase. Prvobitno se izdvaja deo slike koji sadrži oko pomoću ključnih tačaka na licu (slika 3a). Nad tom isečenom slikom potrebno je uraditi binarizaciju kako bi se izdvojio tamniji deo oka koji sadrži zenicu i dužicu od svetlije beonjače (slika 3b). Binarizacija je proces kojim se slika podeli na beli i crni deo tj. bilo koji piksel koji ima vrednost veću od datog praga biće beo, a onaj koji ima vrednost manju od praga biće crn. Koliko se dobro izdvoje zenica i dužica od beonjače zavisi od praga binarizacija. Kako bi se dinamički odredio taj prag za bilo koju sliku potrebno je posmatrati histogram slike (slika 4a) tj. kumulativni histogram slike (slika 4b). Prag se postavlja tako da trećina najtamnijih piksela bude crna, a ostali da budu beli (slika 4c). Invertuje se slika kako bi pikseli zenice i dužice bili beli i zatim se primeni morfološka operacija zatvaranja za uklanjanje manjih rupa u većim celinama (slika 3c). Operacija zatvaranja se sastoji iz dilatacije, kako bi se popunile rupe, i erozije, kako bi se skupili delovi koji su nepotrebno prošireni u prvoj operaciji. Nakon toga se izračuna centar mase belog dela slike kako bi se dobio centar dužice, a samim tim približna lokacija centra zenice (slika 3d).

Pri binarizaciji velika je verovatnoća da će i deo kapaka biti kategorizovan kao deo rožnjače i



Slika 3: Faze metoda centra mase: a) Originalna slika. b) Binarizovana slika. c) Primenjena operacija zatvaranja. d) Rezultujuća koordinata

Figure 3: Different phases of the center of mass method: a) Original image. b) Binarized image. c) Closing operation applied. d) Result coordinate.



Slika 4: Određivanje praga binarizacije: a) Histogram slike. b) Kumulativni histogram slike. c) Plavo: trećina ukupnog broja piksela, crveno: Prag za binarizaciju.

Figure 4: Determining the binarization threshold: a) Histogram. b) Cumulative histogram. c) Blue denotes a third of the total number of pixels, red denotes the new binarization threshold.

time se pogoršava preciznost centra mase. Kako bi se odbacio uticaj kapaka na rezultat ovog pristupa, potrebno je izdvojiti onaj deo slike koji je ograničen kapcima. Ovo se postiže primenom proste maske definisane prethodno navedenim ključnim tačkama oko kapaka.

Metod aproksimacije kruga. Kada je pogled usmeren tako da kapci prekrivaju neki deo zenice ili dužice metod centra mase postaje manje precizan jer se ne vidi potpuna dužica, pa je centar mase pomeren u stranu. Kako bi se uticaj ovakvih slučajeva umanjio potrebno je da se dužica predstavi nekim krugom čiji bi centar odgovarao centru dužice, a time i približno centar zenice. Poluprečnik kruga koji odgovara dužici se može proceniti na osnovu razdaljine između očiju. Eksperimentalno je određeno da je poluprečnik dužice približno jednak 1/9 razdaljine između očiju u velikoj većini slučajeva, pa se ta procena primenjuje kao poluprečnik kruga. Kako bi se aproksimirani krug nalazio u takvom položaju da se njegov centar poklapa sa centrom dužice, potrebno je naći poziciju u kojoj se ivica tog kruga najbolje poklapa sa ivicom dužice (slika 5b). Za nalaženje ivice dužice pri-

menjuje se Sobel detektor za ivice koji generiše monohromu sliku na kojoj su jake ivice bele a mesta bez ivica crna (slika 5c).

Grub centar zenice dobijen iz metode centra mase se koristi kao polazna tačka za poklapanje kruga sa ivicom dužice. Posmatraju se tačke u okolini grubog centra. Svaka od tih tačaka se gleda kao mogući centar kruga. Za svaku tačku se određuje koliko se dobro krug sa centrom u toj tački poklapa sa ivicom dužice tako što se po ivici tog kruga sumiraju pikseli iz slike dobijene od Sobel detektora. Tačka u kojoj je ta suma najveća se smatra centrom dužice odnosno zenice. Ovo se efektivno može izvršiti konvolucijom sa kernelom u obliku kruga debljine jednog piksela (slika 5d). Operaciju konvolucije moguće je realizovati primenom optimizovane brze Furi-jeve transformacije, čija se implementacija nalazi u *OpenCV* biblioteci (Bradski, 2000), što čini algoritam znatno bržim nego kada se koristi iterativni pristup.

Moguća varijacija nad predstavljenom metodom jeste da se menja poluprečnik kruga. Aproksimacija za poluprečnik dužice baziran na razdaljini očiju može da varira u popu-



Slika 5: Metod aproksimacije kruga: a) Originalna slika. b) Sobel detekcija ivica. c) Kernel sa punim krugom. d) Kernel sa delom kruga koji ne uključuje kapke

Figure 5: Circle approximation method a) Original image. b) Sobel edge detection. c) Kernel with a full circle. d) Kernel with a part of the circle which doesn't include the eyelids.



Slika 6: Vektor oka

Figure 6: Eye vector

laciji za do $\pm 10\%$. Iz tog razloga, variranje poluprečnika za tu vrednost može povećati preciznost metode. Takođe, da bi se izbacio uticaj kapaka moguće je sumirati po krugu samo u intervalu $[-\pi/5, \pi/5] \cup [4\pi/5, 6\pi/5]$ tj. da se generiše kernel sa delom kruga koji pripada tom intervalu (slika 5d). Ovaj interval definiše deo kruga koji u većini slučajeva neće obuhvatiti kapke.

Vektor oka

Vektor oka predstavlja razdaljinu centra zenice od unutrašnje ivice oka (slika 6). Unutrašnja ivica je uzeta kao referentna tačka u odnosu na koju se posmatra centar zenice zato što u većini slučajeva ne menja položaj na licu kao što radi spoljašnja ivica. Vektor oka je značajna veličina jer na prost način određuje relativni položaj zenice i na osnovu koje se može posmatrati pomeranje oka i samim time kako se menja smer pogleda.

Kalibracija i funkcija mapiranja

Funkcija mapiranja je funkcija takva da sa proizvoljnim vektorom oka kao ulazom, izlaz će biti odgovarajuća tačka na ekranu u koju korisnik gleda. Ova funkcija predstavlja vezu između

vektora oka i koordinata na ekranu koja može da se opiše pomoću nekoliko različitih matematičkih modela kao što su linearna regresija, SVR (Support Vector Regression) i polinomijalna regresija. Linearna regresija ima manje zahtevnu kalibraciju, ali je zato nedovoljno precizna, dok SVR zahteva veliki broj kalibracionih tačaka i ima visoku preciznost. Polinomijalna regresija zahteva relativno mali broj tačaka za kalibraciju i ima prihvatljivu preciznost te je ona primenjena. Tačnije, primenjena je polinomijalna funkcija drugog reda koja ima sledeći oblik:

$$u_x = a_0 + a_1g_x + a_2g_y + a_3g_xg_y + a_4g_x^2 + a_5g_y^2$$

$$u_y = b_0 + b_1g_x + b_2g_y + b_3g_xg_y + b_4g_x^2 + b_5g_y^2 \quad (1)$$

gde je (u_x, u_y) tačka na ekranu u koju korisnik gleda, a (g_x, g_y) vektor oka. Koeficijenti $a_0 \dots a_5$ i $b_0 \dots b_5$ se određuju u koraku kalibracije.

Kalibracija se sastoji iz posmatranja vektora oka pri pogledu u devet tačaka na ekranu. Nakon toga se napravi relacija između koordinata prikazanih tačaka i odgovarajućih vektora oka koristeći navedenu polinomijalnu regresiju. Koeficijenti regresije se određuju metodom najmanjih kvadrata.

Velika mana navedene metode jeste to da u velikoj meri zavisi od položaja glave. Rotacija i translacija glave u bilo kom smeru izazivaju veliki pomeraj u izlaznoj tački u suprotnom smeru. Iz ovog razloga je potrebno držati glavu u jednom položaju tokom korišćenja sistema. Primenom algoritma za estimaciju položaja glave i dodavanjem pomeraja koji rezultuje pomeranjem glave na konačnu tačku na ekranu moguće je rešiti ovaj problem. U ovom radu estimacija glave nije implementirana.



Slika 7: Primer BioID baze sa obeleženim centrima zenica

Figure 7: An example of the BioID dataset with marked pupil centers

Rezultati

Ispitivanje preciznosti sistema je sačinjeno iz dva dela tj. ispitivanje preciznosti određivanja centra zenice i preciznosti određivanja smera pogleda. Za ovo se koriste, respektivno, *BioID* (Jesorsky, 2001) i *Eye of the Typer* (Papoutsaki, 2018) baze podataka.

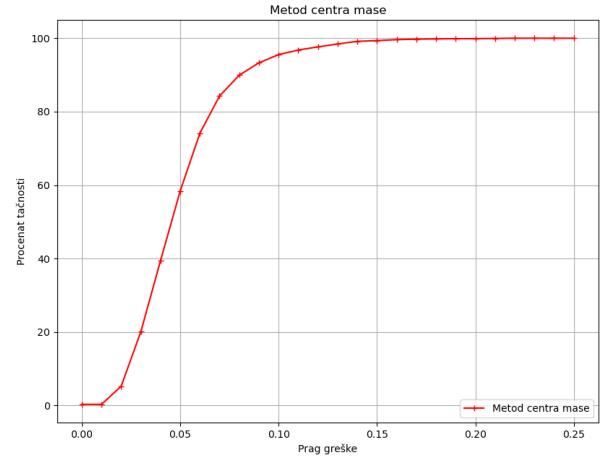
Određivanje centra zenice

Za ispitivanje preciznosti određivanja centra zenice koristi se *BioID* baza podataka. Ova baza se sastoji iz 1520 monohromih slika 23 ljudi slikanih kompjuterskom kamerom pod raznim uglovima i uslovima osvetljenja (slika 7). Takođe, baza sadrži i podatke o lokacijama centra zenice na svakoj slici.

Kako bi se odredila greška metode usvaja se normalizovana greška e

$$e = \frac{\max(d_{left}, d_{right})}{d}$$

gde su d_{left} i d_{right} udaljenosti između izračunatih centra zenice i pravih centara za levo i desno oko, a d razdaljina između očiju utvrđena na osnovu podataka iz baze. Tačnost algoritma se izražava kao procenat slika na kojima je algoritam tačno odredio centar zenice tj. na kojima je normalizovana greška e manja od nekog praga. Gleda se tačnost na različitim pragovima greške da bi se bolje razumelo koliko dobro algoritam radi za koju primenu. Dva najznačajnija praga koji se često posmatraju su manje precizan prag sa normalizovanom greskom manjom od 0.1



Slika 8: Grafik zavisnosti tačnosti od praga greške sa metodom centra mase

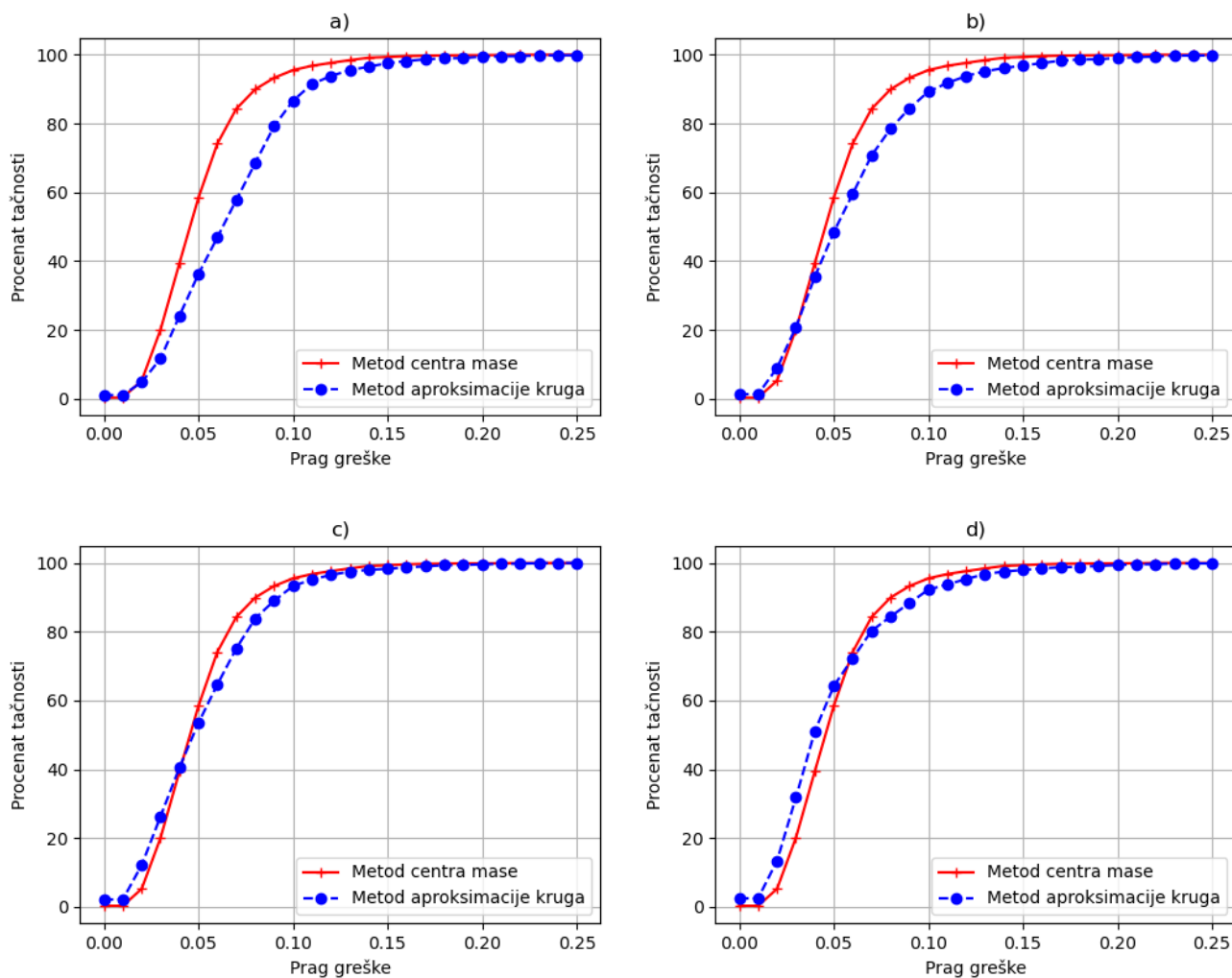
Figure 8: A plot of accuracy versus threshold for error with the center of mass method

($e < 0.1$) i više precizan prag sa greskom manjom od 0.05 ($e < 0.05$).

Metod centra mase. Korišćenjem metode centra mase, sa manje preciznim pragom greške, dobija se tačnost od 95.5%, dok se sa više preciznim pragom dobija 58.4%. Rezultati dobijeni metodom koju su predstavili Cheung i Peng nad istom bazom podataka iznose 93.4% i 87.2%, respektivno. Predstavljena metoda je tačnija od referentne pri manjoj preciznosti, ali je pri većoj znatno lošija.

Metod aproksimacije kruga. Pri ispitivanju metode aproksimacije kruga moguće je, za razliku od metode centra mase, varirati određene parametre. Jedan parametar jeste variranje poluprečnika aproksimiranog kruga za do $\pm 10\%$, dok je drugi deo kruga koji se uzima u obzir pri poklapanju sa ivicama dužice. Variranje poluprečnika kruga se radi u koracima od po 5% u intervalu $[-10\%, 10\%]$. Krug koji se koristi za poklapanje sa ivicom dužice može biti potpun ili da obuhvata samo delove koji ne uključuju kapke.

Po graphicima se vidi da, u poređenju sa metodom centra mase, najbolje rezultate daje varijacija pod d) tj. ona pri kojoj je deo kruga koji obuhvata kapke isključen i varira poluprečnik za do $\pm 10\%$, jer ona na više preciznom pragu greške daje bolje rezultate (64.2%). Ipak, i ova kombinacija nije savršena jer pri manje preciznom pragu (92.2%) još uvek daje manju



Slika 9: Upoređivanje tačnosti metode aproksimacije kruga pri menjanju parametara algoritma: a) Konstantan poluprečnik, ceo krug; b) Konstantan poluprečnik, deo kruga; c) Variranje poluprečnika, ceo krug; d) Variranje poluprečnika, deo kruga.

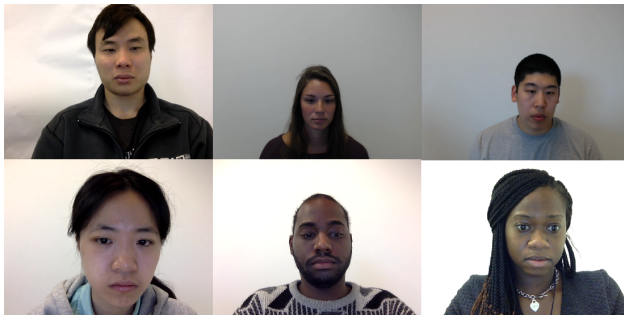
Figure 9: Comparison of the accuracy of the circle approximation method with varying parameters: a) Constant radius, full circle; b) Constant radius, partial circle; c) Varying radius, full circle; d) Varying radius, partial circle.

tačnost od prethodne metode. Ova metoda je lošija od one predstavljene u radu Cheung i Peng pri oba praga greške, ali je pritom manja razlika u tačnosti nego kod metode centra mase. Veća greška na manje preciznom pragu ukazuje na to da metod u većem broju slučajeva daje rezultate koji su više udaljeni od tačne vrednosti, ali kada je tačan biće precizniji od metoda centra mase.

Određivanje smeru pogleda

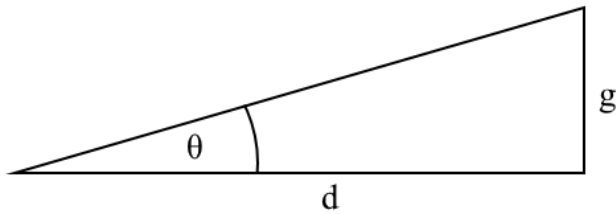
Za ispitivanje određivanja smeru pogleda se koristi baza podataka *Eye of the Typer*. Baza se sastoji iz podataka za 51 ispitanika koji su

učestvovali u istraživanju o pogledu čoveka. Istraživanje je sačinjeno iz nekoliko zadataka koji uključuju pritiskanje određenih tačaka na ekranu, čitanje i kucanje teksta. Baza sadrži snimke lica ispitanika, podatke o samim ispitanicima, dimenzije i rezolucija ekrana, udaljenost od ekrana i smer pogleda u svakom trenutku što je dobijeno uređajem za praćenje pogleda (Tobii x3-120). Poslednji zadatak koji je svaki ispitanik odradio jeste gledanje u 9 različitih tačaka na ekranu, što se koristi za kalibraciju sistema za svakog ispitanika. Od navedenih ispitanika izabrano je 6 koji nisu znatno pomerili glavu tokom istraživanja zato što bi pomeranje



Slika 10: Uzorak iz Eye Of The Typer baze podataka

Figure 10: Sample from the Eye Of The Typer database



Slika 11: Pravougli trougao za računanje ugla tačnosti

Figure 11: The right angle triangle used for calculating the angle of accuracy

glave izazvalo veliku grešku u rezultatu. Svako od njih je imao po 6 zajedničkih zadataka ne uključujući kalibraciju.

Tačnost sistema se izražava u stepenima (θ).

$$\theta = \arctan\left(\frac{g}{d}\right)$$

gde je g udaljenost tačne tačke na ekranu dobijene iz baze od procenjene tačke dobijene sistemom, a d udaljenost korisnika od ekrana (slika 11).

Kako bi se dodatno ispitale metode upoređuju se rezultati dobijeni kada se za određivanje centra zenice koristi metod centra mase i metod aproksimacije kruga. Na tabeli 1 se može videti da kada sistem primenjuje metod centra mase rezultati su znatno bolji nego sa metodom aproksimacije kruga koji u velikom broju slučajeva ima grešku veću nego kod metode centra mase. Najveća greška se javlja kod 5. ispitanika jer su oči u tom slučaju znatno zaklonjene trepavicama i nalaze se u senci zbog položaja glave pa se teže razlikuju dužica i zenica od ostatka beonjače i ostatka lica. U proseku dobija se greška od 4.46° za metodu centra mase, a 5.6° za metod

Tabela 1: Pregled uglova tačnosti po ispitanicima i po metodama za određivanje centra zenice

Table 1: An overview of the angle of accuracy for each participant and method

Ispitanik	Centar mase	Aproksimacija kruga
1.	3.06	3.71
2.	5.01	5.37
3.	4.15	5.78
4.	3.52	4.51
5.	7.58	9.89
6.	3.43	4.35
Prosek	4.46	5.60

aproksimacije kruga. U poređenju sa referentnim radom (Cheung, 2015) koji je predstavio grešku od 1.28° u slučaju bez pomeranja glave. Glavni uzrok ove greške jeste manja preciznost pri određivanju centra zenice.

Zaključak

U radu su predstavljene dve metode za nalaženje centra zenice na slici lica. Metod centra mase primenjuje prost postupak za brzo i efikasno dobijanje grubih rezultata, dok metod aproksimacije kruga nadograđuje prethodnu metodu koracima koji uključuju nekoliko parametara čiji je uticaj ispitan u radu. Metod aproksimacije kruga je prikazao bolje rezultate na preciznijem pragu, ali ima veću grešku u velikom broju slučajeva. U radu je postignuta tačnost sistema za određivanja smera pogleda od 4.46° za metod centra mase, a 5.6° za metod aproksimacije kruga na bazi *Eye Of The Typer*. Za rad sistema je potrebno držati glavu mirnom. Moguće unapređenje sistema bi bilo dodatak estimacije položaja glave kako bi se uračunao pomeraj nastao od pomeranja glave.

Literatura

- Cheung Y, Peng Q, 2015, Eye gaze tracking with a web camera in a desktop environment, Chinese University of Hong Kong
- Chennamma H.R. 2013, A survey on eye-gaze tracking techniques, Indian Journal of Computer

King D. 2009, Dlib-ml: A Machine Learning Toolkit, Journal of Machine Learning Research 10

Bradski G, 2000, The OpenCV Library. Dr. Dobb's Journal of Software Tools.

Jesorsky O, Kirchberg K. J, Frischholz R. W, 2001, Robust face detection using the Hausdorff distance, Audio Video-Based Biometric Person Authentication AVBPA 2001

Papoutsaki A, Gokaslan A, Tompkin J, He Y, Huang J. 2018, The eye of the typer: a benchmark and analysis of gaze behavior during typing, ETRA '18: Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications

Petar Janićijević

Estimating Gaze Direction With a Single Fixed Web Camera

This paper presents and evaluates a software solution for eye gaze tracking in real-time using a fixed RGB web camera, the likes of which can be found in many modern desktops, laptops and smartphones. The algorithm initially locates the face on an image and isolates the eye regions. To estimate the iris center on the eye region two methods are proposed: center of mass and circle approximation. An eye vector (Fig. 6), which describes the position of the iris center relative to the inner eye corner, is formed. A mapping function (Eq. 1) is used to relate a given eye vector to the corresponding gaze point on the user's screen, and it must be calibrated to each person. The two methods for iris center estimation are compared using the *BioID* database, where the results are presented as percentage of the images from the database for which the method is successful in estimating the iris center within a given threshold. The center of mass method performs better than the circle approximation method for a higher threshold with 95.5% compared to 92.2%, but worse for a lower threshold with 58.4% compared to 64.2%. Finally, for the comparison of the two methods when applied to the eye gaze tracking system the *EOTT* database is used. When using the center of mass method the average accuracy is 4.46° , which is better than 5.6° when circle approximation is used.