

PREPOZNAVANJE EMOCIJA SA SLIKA LICA PRIMENOM MULTISENZORSKIH SISTEMA
FACIAL EMOTION RECOGNITION USING MULTISENSOR SYSTEMS

Ranko Petrović, Miloš Pavlović, Branka Stojanović, Snežana Puzović

REZIME: Istraživanja na temu kako mi kao ljudska bića reagujemo na svet oko sebe kao i kako intereagujemo sa njim i dalje su jedno od najvećih naučnih izazova. Interakcija između čoveka i računara bila bi prirodnija kada bi računari bili u stanju da registruju i reaguju na ljudske emocije. U ovom radu su predstavljeni neki od sistema za prepoznavanje emocija u interakciji čoveka i računara koji se zasnivaju na izrazima lica, glasu, rukopisu kao i upotrebi miša i tastature. Takođe, u radu su predstavljeni i praktični rezultati sistema za prepoznavanje četiri različita emocionalna stanja (neutralno emocionalno stanje, radost, iznenađenost, ljutnja) sa slika lica uslikanim termalnom kamerom i kamerom vidljivog spektra, a koji je baziran na histogramu orijentisanih gradijenata i metodi nosećih vektora.

KLJUČNE REČI: interakcija čoveka i računara, prepoznavanje emocija, HOG karakteristike, SVM klasifikator

ABSTRACT: Research on how we, as human beings, react with the world around us and how we interact with it remains one of the greatest challenges in science. Human-computer interaction would be more natural if computers were able to register and respond to human emotions. This paper presents some of the systems for emotions recognition in human-computer interaction, that are based on facial expressions, voice, handwriting, as well as the use of mouse and keyboard. Also, the practical results of a system for recognizing four different emotional states (neutral emotional state, joy, surprise, anger) from face images taken with a thermal camera and a visible light camera based on the histogram of oriented gradients and the support-vector machine classifier are presented.

KEY WORDS: human-computer interaction (HCI), emotion recognition, HOG features, SVM classifier

1. UVOD

Emocija predstavlja psihološko stanje ljudskog uma [1]. Čovekovo emocionalno stanje utiče na to kako će se ponašati i kako će odreagovati u različitim situacijama. Tako na primer, pozitivne emocije omogućavaju ljudima da budu kreativni, da rešavaju kompleksne probleme, dok ih sa druge strane negativne emocije potiskuju u usko, fokusirano razmišljanje, pa tako neke vrlo jednostavne probleme nisu u stanju da reše ako su na primer uplašeni [2].

Novi trendovi u interfejsima u okviru interakcije čoveka i računara koji su evoluirali od upotrebe miša i tastature do sistema za automatsko prepoznavanje govora i specijalnih interfejsa dizajniranih za hendikepirane ljude ne koriste u potpunosti prednosti ovih dragocenih komunikativnih sposobnosti, što rezultira još uvek manjim oblikom interakcije nego što je sama prirodna interakcija između dva živa bića. Ako bi računari bili u mogućnosti da prepoznaju određeni emocionalni model, mogli bi onda i da omogućе određenu i odgovarajuću pomoć korisnicima na način koji je više u skladu sa njihovim potrebama i željama. Opšte je prihvaćeno iz psihološke teorije da se ljudske emocije mogu svrstati u šest tipičnih emocionalnih grupa: iznenađenost, strah, odvratnost, ljutnja, radost i tuga [3].

Automatsko prepoznavanje emocionalnih stanja je važna istraživačka tema u sistemima za video nadzor. Mogućnost automatske procene dosade, nepažnje i stresa u situacijama u kojima je čvrsta pažnja od suštinske važnosti, kao što su pilotiranje avionom, rad u kontroli leta ili u nuklearnom postrojenju je od izuzetnog značaja kada je u pitanju razvoj sistema nove generacije u interakciji čoveka i računara [4]. Sistemi za prepoznavanje emocija se koriste u okviru automobila gde se informacije o mentalnom stanju vozača mogu analizirati kako bi se on održavao budnim tokom vožnje. Ovo pomaže da se

izbegnu nesreće prouzrokovane mentalnim zdravstvenim stanjem vozača [5]. Mehanizmi za otkrivanje scena ili frejmova koji prikazuju slike lica sa izrazima bola, besa i straha mogli bi da omogućе indeksiranje filmova koji obiluju nasilnim scenama. Generalno kompjuterska analiza emocionalnih stanja značajno bi mogla da olakša istraživanje u oblastima kao što su nauka o ponašanju, antropologija, neurologija, psihijatrija u kojoj su pouzdanost, osetljivost i preciznost naročito gorući problemi [4, 6].

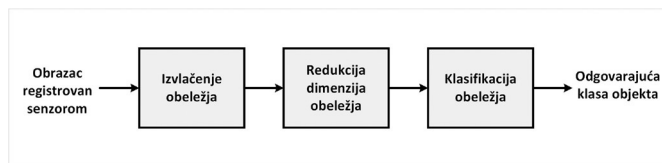
Posebnu pažnju treba obratiti na veliki tehnološki napredak i velike investicije koje se dešavaju u sklopu realizacije pametnih gradova (eng. *smart cities*) čiji je najvažniji cilj poboljšanje efikasnosti i kvaliteta života građana. Internet Stvari (eng. *Internet of Things, IoT*) se koriste za komunikaciju sa ljudima i za nadgledanje infrastrukture i servisa. Informaciono-komunikacione tehnologije, pomoću sistema za praćenje u realnom vremenu koji podrazumeva veliki broj senzora, omogućavaju prikupljanje i obradu svih vrsta informacija kako bi građanima pružili efikasnije usluge. U suštini jedan od glavnih ciljeva pametnih gradova je sveukupno poboljšanje života, a budući da je zdravlje važan faktor u tom pogledu, pametni zdravstveni sistem mora da bude nezaobilazan deo svakog pametnog grada. Pomoću sistema za elektronsko zdravstvo (eng. *eHealth*) registrovani lekari imaju mogućnost da stalno daljinski nadgledaju emocionalna stanja pacijenata koja mogu da ukažu na potencijalne zdravstvene probleme, a onda u slučaju potrebe i da pravovremeno reaguju. Primer jednog ovakvog sistema je sistem za monitorisanje Parkinsonove bolesti na osnovu emocionalnog stanja građana koje se registruje iz govornog signala prikupljenog različitim sensorima postavljenim širom grada [7, 8].

Rad je organizovan u šest poglavlja. U poglavlju dva su opisane različite tehnike koje se koriste u prepoznavanju emo-

cija i dat je pregled postojeće literature na tu temu. Poglavlje tri definiše metodologiju rada sa detaljnim objašnjenjem histograma orijentisanih gradijenata (eng. *Histogram of Oriented Gradients, HOG*) kao i metode nosećih vektora (eng. *Support Vector Machine, SVM*) koji su korišćeni u ovom radu. Poglavlje četiri opisuje realizovanu arhitekturu sistema za prepoznavanje emocija sa slika lica. Takođe, u ovom poglavlju je opisana i baza slika lica koja je korišćena za treniranje i testiranje. Poglavlje pet donosi diskusiju i prikaz rezultata koji su dobijeni testiranjem, dok su u poglavlju šest izneti zaključci i predlozi za buduće istraživanje.

2. TEHNIKE PREPOZNAVANJA EMOCIJA I PREGLED POSTOJEĆE LITERATURE

Obrazac (eng. *pattern*) je reprezentativni potpis određenog objekta pomoću kojeg ga je moguće prepoznati. Proces prepoznavanja obrazaca u suštini podrazumeva mapiranje skupa obrazaca u jednu od nekoliko klasa objekata. Ponekad je obrazac predstavljen vektorom koji sadrži obeležja objekta. U skladu sa tim opšti algoritam za prepoznavanje obrazaca se može predstaviti u tri koraka: izdvajanje obeležja, izbor obeležja i klasifikacija, kao što je prikazano na slici 1 [1].



Slika 1. Osnovni koraci u prepoznavanju obrazaca

Proces izdvajanja obeležja podrazumeva upotrebu jednog ili više senzora za merenje reprezentativnih obeležja. Blok za izbor obeležja vrši selekciju najznačajnijih obeležja. Blok za klasifikaciju klasifikuje odabrana obeležja u jednu od nekoliko klasa objekata. Kod problema prepoznavanja obrazaca mogu se primeniti dva tipa klasifikacije podataka: nadgledana klasifikacija ili diskriminacija, i nenadgledana klasifikacija ili grupisanje. U nadgledanoj klasifikaciji, obično se za skup trening instanci ili podataka unapred zna kojoj klasi pripadaju. Podaci iz trening skupa obuhvataju skup merenja o svakom objektu i prosleđuju se klasifikatoru na obučavanje. Nakon procesa obučavanja klasifikatoru se prosleđuju podaci čija klasa još uvek nije određena, radi mapiranja novih podataka u odgovarajuće klase. U nenadgledanoj klasifikaciji oznake klasa nisu unapred poznate. Sistem učenja deli kompletan skup podataka na po mogućstvu nepreklapajuće podskupove na osnovu neke određene mere sličnosti koja je karakteristična za svaki podskup, koji se naziva još i klasa ili klaster [1].

Prepoznavanje emocija takođe predstavlja problem koji se svodi na prepoznavanje obrazaca. Tako na primer prepoznavanje emocija zasnovano na izrazu lica zahteva izdvajanje skupa crta lica iz izraza lica posmatrane osobe. Prepoznavanje emocija se ovde svodi na klasifikaciju crta lica u jednu od nekoliko klasa emocija. Obično se nadgledani pred-trenirani klasifikator, koji koristi emocionalna obeležja kao ulaz u sistem i klase emocija kao izlaz iz sistema, koristi za rešavanje ovog tipa

problema. Emocionalno stanje ljudskog uma se izražava na različite načine uključujući izraze lica, glas, držanje, rukopis, pa čak i način upotrebe tastature i miša. Kada se za prepoznavanje emocija koristi jedan od navedenih načina govori se o unimodalnom pristupu. Prepoznavanje emocija iz više modaliteta predstavlja multimodalni pristup [1].

2.1 Prepoznavanje emocija na osnovu glasa

Informacije iz spektralnih elemenata glasa, kao što su brzina govora, visina tona, energija i intenzitet govora, pravilnost ritma, itd. koriste se u prepoznavanju emocija kod ljudi. Standardna procedura u prepoznavanju emocija iz govornog signala, podrazumeva izdvajanje karakterističnih obeležja koja različiti klasifikatori svrstavaju u različite klase tako što ih porede sa odgovarajućim skupom karakteristika koje su korišćene za treniranje [9].

Obeležja govornog signala se mogu koristiti za prepoznavanje razlike između nekoliko emocionalnih stanja. Za izgradnju sistema za prepoznavanje emocija veoma je važno izdvojiti različite prozodijske tj. akustičke i spektralne karakteristike iz govornog signala. Neka od glavnih obeležja govornog signala su visina tona, amplituda i formanti. Visina tona je posledica treperenja glasne žice i može se meriti promenom frekvencije. Vreme trajanja između dva maksimuma u govornom signalu naziva se period, dok broj vibracija u jedinici vremena predstavlja osnovnu frekvenciju tona. Kada se govori o spektralnom domenu, obeležja koja se mogu koristiti u prepoznavanju emocija su linearni predikcioni koeficijenti (eng. *Linear Prediction Coefficient, LPC*), Mel frekvencijski kepsralni koeficijenti (eng. *Mel Frequency Cepstral Coefficient, MFCC*) i linearni predikcioni kepsralni koeficijenti (eng. *Linear Prediction Cepstral Coefficient, LPCC*). Za prepoznavanje emocija najvrednije informacije su sadržane u spektralnom obliku vokalnog trakta. Obeležja kao što su visina tona, amplituda, modulacija i formanti se koriste u dijagnozi depresije. Visina tona i energija se koriste u detekciji emocionalnih stanja kao što su radost, ljutnja, tuga i neutralno stanje. Spektralna obeležja se koriste u analizi emocionalnog stanja kao što je uzbuđenje [10-13].

2.2 Prepoznavanje emocija na osnovu rukopisa

jedan od najjednostavnijih i najintuitivnijih oblika interakcije čoveka i računara jeste prepoznavanje rukopisa, odnosno njegovo pretvaranje u kucani tekst. Najznačajnije informacije u rukopisu nisu u obliku slova već u načinu na koji su ona napisana [2]. Karakteristike rukopisa poput početne linije, nagiba, pritiska olovke, veličine i margine mogu pomoći u prepoznavanju osoba koje su emocionalno uznemirene ili depresivne i potrebna im je psihološka pomoć u savladavanju takvih negativnih emocija. Osoba koja se oseća napeto, anksiozno ili depresivno prikazaće drhtave poteze kod određenih slova, dok će osobe koje se osećaju srećno i uzbuđeno imati jasne i glatke poteze [9].

Nagib je dominantni ugao poteza olovke prema gore ili prema dole, odnosno relativno poravnanje slova prema osnovnoj liniji. Linija na kojoj se nalazi većina slova nečijeg rukopisa, naziva se osnovna linija. Postoje četiri najčešća pravca osnovne linije, uzlazni, silazni, ravni i talasni, pri čemu svaki od njih odgovara određenom emocionalnom stanju. Pritisak olovke definisan je kao sila koju primenjujemo na površinu za pisanje instrumentom za pisanje. Iznenađna transformacija veličine nečijeg rukopisa iz normalne veličine u mikroskopsku obično ukazuje na uznemireno ili depresivno stanje pisca. Uglavnom sve nabrojane karakteristike jednog rukopisa kao što su nagib, osnovna linija, pritisak, veličina slova itd. su istovremeno i karakteristična obeležja koja se prosleđuju nekom od klasifikatora koji na osnovu poređenja sa odgovarajućim trening skupom kao rezultat daje emocionalno stanje pisca [9].

2.3 Prepoznavanje emocija na osnovu upotrebe miša i tastature

Analiza dinamike pritiska i pomeranja miša nije nova ideja, jer su u pitanju uobičajene biometrijske tehnike koje se koriste u polju bezbednosti računarskih sistema. Biometrijske karakteristike povezane sa kucanjem na tastaturi se mogu predstaviti pomoću obeležja koja se zasnivaju na vremenskom trajanju između pritisaka tastera, brzini kucanja, učestalosti korišćenja "backspace" tastera, itd. Kada je u pitanju korišćenje miša, biometrijske karakteristike se zasnivaju na obeležjima kao što su brzina miša, ubrzanje, preciznost, pravac pomeranja, broj klikova itd [14].

Obeležja koja se dobijaju kucanjem na tastaturi se mogu podeliti na vremenska i frekvencijska obeležja. Vremenska obeležja predstavljaju trajanja pojedinačnih pritisaka na taster i trajanja između pritisaka na različite tastere, dok frekvencijska obeležja predstavljaju meru učestalosti pritiskanja određenog tastera kao što su "backspace" taster, taster za brisanje, numerički tasteri itd. Kako ova ponašanja mogu biti pod uticajem emocija, parametri pritiska na tastere mogu biti pokazatelji određenih emocionalnih stanja. Obeležja koja se dobijaju upotrebom miša se zasnivaju na načinu pomeranja pokazivača, upotrebi klika ili skrolovanja [14].

Međutim, glavni problem kada je u pitanju prepoznavanje emocija pomoću miša i tastature je činjenica da način na koji emocije utiču na pritisak tastera i upotrebu miša zavisi od same osobe. Tako na primer, jedna osoba će u slučaju stresa brže kucati na tastaturi, dok će neka druga osoba praviti pauze u stresnim situacijama. Zbog toga se preporučuje izgradnja individualnog klasifikatora za svakog korisnika posebno [14].

2.4 Prepoznavanje emocija na osnovu slika lica

Najčešći modalitet za prepoznavanje emocija je analiza izraza lica. Tradicionalno postoje dve glavne grupe tehnika za predstavljanje izraza lica i izdvajanje odgovarajućih karakterističnih obeležja. Prvu grupu čine geometrijske tehnike koje se oslanjaju na geometrijska obeležja u vidu parametara karak-

terističnih crta lica kao što su oči, usta i nos. Sa druge strane postoje tehnike koje se zasnivaju na izgledu lica i na njega gledaju kao na niz vrednosti različitog intenziteta koje su prikladno obrađene u pretprocesiranju [1, 15]. Hibridne tehnike podrazumevaju kombinaciju obeležja ovih tehnika.

Konvencionalni pristupi za prepoznavanje emocija na osnovu slika lica koji podrazumevaju upotrebu obeležja i klasifikatora su i dalje veoma aktuelni, naročito kada je u pitanju prepoznavanje emocija u realnom vremenu, jer od sve popularnijeg pristupa koji se zasniva na dubokom učenju, razlikuju se po znatno manjoj složenosti a postižu takođe visok stepen tačnosti [16]. Obeležja koja se koriste za prepoznavanje emocija na osnovu slika lica su lokalni binarni obrazac (eng. *Local Binary Pattern, LBP*) kao najpopularnije obeležje koje se zasniva na teksturi [17, 18], lokalni gradijentni kod (eng. *Local Gradient Code, LGC*) koji poredi susedne piksele i po horizontali i po vertikali i po dijagonali [19], Gabor filter koji transformiše sliku u male talase koji olakšavaju prepoznavanje željenih obeležja [20], kao i histogram orijentisanih gradijenata, na kojem se i zasniva algoritam predložen u ovom radu. Fuzija pomenutih obeležja povećava uspešnost prepoznavanja [17].

Za klasifikaciju izdvojenih obeležja se takođe mogu koristiti različite tehnike poput metode nosećih vektora, fuzzy klasifikatora, klasifikatora koji se zasniva na sakrivenom Markovljevom modelu (eng. *Hidden Markov Model, HMM*), kao i različiti tipovi neuralnih mreža [1].

3. METODOLOGIJA RADA

Algoritam korišćen u ovom radu se zasniva na HOG obeležjima i SVM klasifikatoru. Skupovi podataka za trening i testiranje su formirani pomoću multisenzorskog sistema koji se sastoji od jedne kamere koja radi u vidljivom opsegu i jedne termalne kamere. U kontrolisanim uslovima pouzdanost sistema za prepoznavanje emocija sa slika lica je znatno veća nego u nekontrolisanim uslovima kao što su na primer uslovi slabog i promenljivog osvetljenja, dok u uslovima potpunog mraka sistem postaje potpuno neupotrebljiv. Upotrebom termalnog senzora u okviru multisenzorskog sistema koji je invarijantan čak i na ekstremne promene osvetljenja kao i na uslove potpunog mraka, može se prevazići opisan problem, jer se za razliku od senzora vidljive svetlosti koji se zasnivaju na reflektovanoj svetlosti, termalni senzori zasnivaju na emitovanoj radijaciji.

3.1 Histogrami orijentisanih gradijenata

Histogrami orijentisanih gradijenata (eng. *Histogram of Oriented Gradients, HOG*) su obeležja koja generalno imaju dobre karakteristike kao deskriptori koji se koriste u prepoznavanju objekata, a posebno u sistemima za prepoznavanje lica. HOG se koristi u slikama vidljivog opsega za identifikaciju ivica lica koje su posledica nejednake refleksije svetlosti sa lica. Distribucija intenziteta osvetljenosti ili pravac lokalnih ivica u slici se opisuje pomoću HOG obeležja. Kada je u pita-

nju termalna slika, HOG se koristi za identifikaciju ivica lica koje su posledica nejednake raspodele toplote po licu. HOG radi na nivou lokalnih ćelija, pa je invarijantan u odnosu na geometrijske i fotometrijske transformacije što su i glavne prednosti ovog deskriptora [21-23].

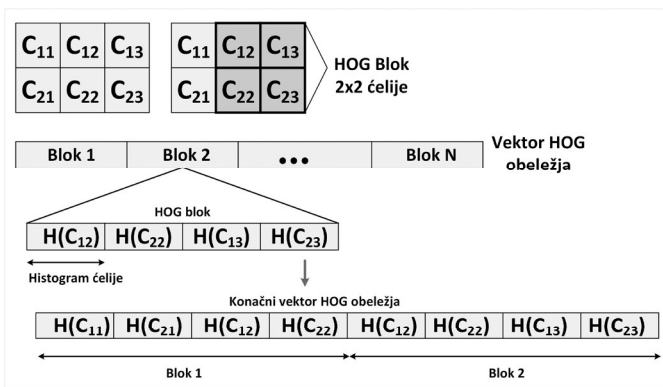
Gradijent u digitalnoj obradi slike predstavlja pravac najvećeg povećanja osvetljenosti u svakom pikselu slike. Kombinovane promene osvetljenja duž horizontalnih i vertikalnih pravaca u posmatranim pikselima odgovaraju vrednostima gradijenta i predstavljene su matricom gradijentnih modula G . Takođe se formira i matrica orijentacije gradijenata za svaki piksel čije su vrednosti uglovi θ .

$$\nabla f(x, y) = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix}^T = [G_x \ G_y]^T$$

$$G = |\nabla f| = \sqrt{G_x^2 + G_y^2}$$

$$\theta = \arctg\left(\frac{G_y}{G_x}\right)$$

Cilj proračuna HOG obeležja je da se dobiju diskretni histogrami gradijenata koji odgovaraju ćelijama, jedinicama na koje je slika podeljena a koje se sastoje od grupe piksela. Da bi se obuhvatile prostorne informacije većih razmera trebalo bi povećati veličinu ćelija, ali povećanjem veličine ćelija može doći do gubitka prostornih informacija sitnijih razmera, odnosno detalja. Za izračunavanje HOG obeležja potrebno je diskretizovati moguće gradijentne vrednosti, zbog čega se definišu dozvoljeni pravci koje gradijent može imati u svakom pikselu. Gradijent se projektuje u dva najbliža, susedna diskretizovana pravca. Pomenuta projekcija podrazumeva preraspodelu vrednosti gradijentnog modula na dozvoljene, diskretne pravce i to u zavisnosti od ugaonog rastojanja gradijenta od svakog od njih. Pravci mogu biti potpisani ($0^\circ \leq \theta \leq 360^\circ$) ili nepotpisani ($0^\circ \leq \theta \leq 180^\circ$). U slučaju nepotpisanih pravaca, nema potrebe za više od devet različitih diskretnih pravaca [24, 25].



Slika 2. Proračun vektora HOG obeležja

Histogrami se zatim izračunavaju brojanjem mogućih pravaca u svakoj ćeliji, odnosno sumiranjem vrednosti gradijen-

tnih modula koji odgovaraju posmatranom pravcu i to za sve piksele u ćeliji i za svaki diskretni pravac. Vrednosti histograma opisuju prisustvo dominantnih pravaca u posmatranoj ćeliji i na taj način lokalno prisustvo ivica u slici za više susednih ćelija, što se može videti i na slici 3 [26].

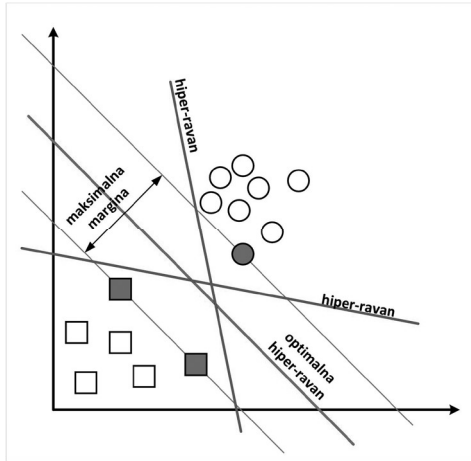


Slika 3. Primer HOG obeležja na slici lica

Lokalna normalizacija vrednosti histograma za svaku od ćelija je neophodna, jer vrednosti histograma između prostorno bliskih ćelija mogu značajno varirati. Normalizacija se vrši po blokovima koji obuhvataju nekoliko ćelija. Pošto su blokovi preklapajući, pojedinačne ćelije mogu ući u sastav više blokova, što znači da svaki histogram kojim je svaka ćelija opisana ulazi u sastav konačnog HOG vektora više puta. Velike vrednosti preklapanja mogu obuhvatiti više informacija, ali na taj način se značajno povećava i veličina vektora [23, 24].

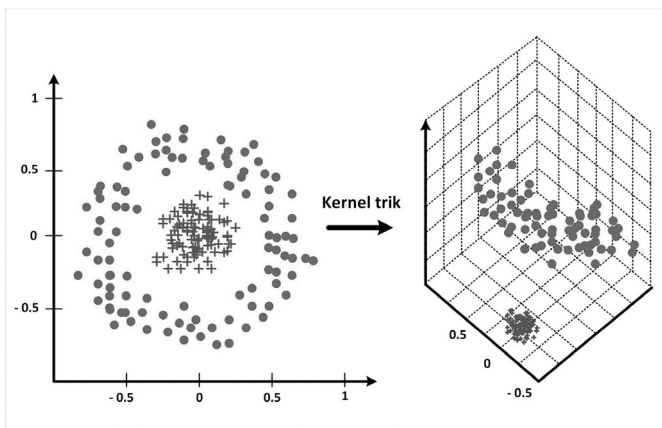
3.2 Metoda nosećih vektora

Metoda nosećih vektora (eng. *Support Vector Machine, SVM*) je klasična metoda mašinskog učenja za klasifikaciju podataka i regresiju koja ima veliku primenu u oblasti kao što je prepoznavanje lica, jer postiže veoma dobre rezultate kada je u pitanju klasifikacija velike količine podataka. SVM predstavlja tzv. model nadgledanog učenja koji generiše ulazno-izlazne funkcije mapiranja na osnovu skupa označenih podataka za treniranje. Podaci za treniranje se mapiraju iz nekog ulaznog prostora u neki novi prostor veće dimenzionalnosti u kojem pomenuti podaci iz trening skupa mogu biti linearno separabilni. Zadatak faze obuke sistema je da pronade optimalno separabilnu hiper-ravan u tom novom prostoru (eng. *feature space*) gde su obeležja predstavljena na takav način da svi podaci iz jedne klase mogu biti sa iste strane te hiper-ravni. Optimalna hiper-ravan je ona koja razdvaja podatke za obučavanje tj. podatke iz trening skupa sa maksimalnom mogućom marginom, a koja u stvari predstavlja rastojanje između hiper-ravni i najbližeg podatka za obučavanje svake klase. SVM određuje optimalnu ravan razdvajanja tako što maksimizira rastojanje između razdvajajuće hiper-ravni i podataka koji su najbliži potencijalnoj liniji razdvajanja. Ti podaci se nazivaju noseći vektori. Tako je u suštini separabilna hiper-ravan određena malim brojem podataka različitih klasa koji su joj najbliži, tzv. nosećim vektorima (eng. *support vectors*) [27, 28].



Slika 4. Optimalna hiper-ravan sa maksimalnom marginom

Ukoliko skup podataka za treniranje nije linearno separabilan potrebno je mapirati ulazni vektorski prostor u novi višedimenzionalni prostor pomoću kernel funkcije. To je takozvani kernel trik (eng. *kernel trick*) kojim se omogućava rešavanje nelinearnih problema [28].



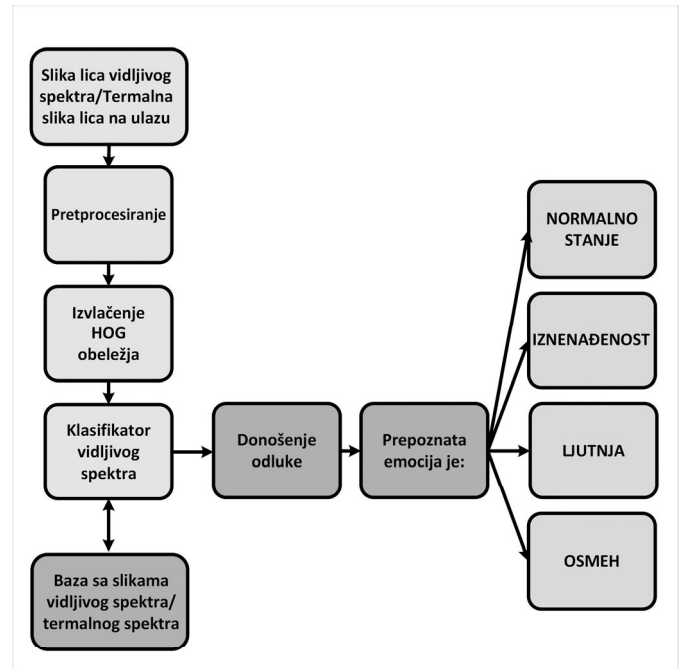
Slika 5. Mapiranje ulaznog prostora u prostor veće dimenzionalnosti

4. ALGORITAM ZA PREPOZNAVANJE EMOCIJA SA SLIKA LICA

Ovaj rad se fokusira na algoritam koji traži i uči karakteristična obeležja koja se međusobno razlikuju za različita emocionalna stanja, ali istovremeno uči i obeležja, koja minimiziraju razlike između slika lica, koja pripadaju različitim osobama sa istim emocionalnim stanjem. Izbor, priprema i izdvajanje kvalitetnih obeležja je od ključnog značaja za dalji tok algoritma. Testirano je prepoznavanje četiri različita emocionalna stanja: neutralno emocionalno stanje, iznenađenje, sreća i ljutnja. Algoritam je primenjen na slike lica koje su uslikane pomoću multisenzorskog sistema zbog čega pripadaju različitim spektralnim domenima, domenu vidljive svetlosti i termalnom domenu. Razlog korišćenja termalnih slika lica za prepoznavanje emocija leži u činjenici da u uslovima slabe osvetljenosti, kao i u uslovima potpunog mraka, nije moguće napraviti dovoljno kvalitetnu sliku lica vidljivog spektra na osnovu koje bi moglo da se obavi prepoznavanje emocija [29-31].

4.1 Arhitektura sistema

Blok dijagram sistema je prikazan na slici 6. Sistem je u mogućnosti da obrađuje slike lica uzete termalnim senzorom i da ih poredi sa bazom popunjenom termalnim slikama kao i da obrađuje slike lica uzete senzorom vidljivog spektra i da ih poredi sa bazom popunjenom slikama koje pripadaju opsegu vidljive svetlosti [32, 33].



Slika 6. Blok dijagram predloženog sistema za prepoznavanje emocija

Na ulazu u sistem se može naći termalna slika lica ili slika lica koja pripada opsegu vidljive svetlosti sa kojih je potrebno prepoznati jednu od četiri emocije kao što su neutralno stanje, iznenađenje, sreća ili ljutnja.

Gotovo svi algoritmi za prepoznavanje emocija zahtevaju prolazak kroz iste faze koje su neophodne u procesu prepoznavanja. Da bi se postigla veća efikasnost algoritma, potrebno je na ulaznim slikama lica primeniti određene tehnike pred-obrađe. Uklanjanje mrtvih piksela i lokalnih varijacija koje mogu da budu posledica promene raspodele toplote u termalnim slikama lica, odnosno, promene raspodele osvetljenja u slikama lica koje pripadaju spektru vidljive svetlosti, rešava se u ovoj fazi. Uklanjanje šuma koji može da bude posledica samog senzora ili uslova okruženja u kojem je slika uslikana i koji može značajno da degradira performanse algoritma se takođe rešava blokom za preprocesiranje. Filtriranje slike predstavlja najčešći oblik preprocesiranja, a izbor odgovarajućeg filtra zavisi od vrste izvora šuma kao i od primene samog sistema u kojem je potrebno izvršiti filtriranje. Najčešći filteri koji se koriste u obradi slike, Gausov niskopropusni filter i medijan filter su veoma uspešni u uklanjanju šuma ali po cenu zamućenja (blurovanja) ivica. U slučaju opisanog algoritma to nije dobro rešenje, s obzirom da se predloženi sistem za prepoznavanje lica zasniva na deskriptorima koji nose informaciju o ivicama. Blurovanjem ivica pomoću navedenih filtera gube se informacije od značaja koje su neop-

hodne u procesu prepoznavanja emocija. Kao rešenje nameće se upotreba bilateralnog filtra koji se sastoji od dve komponente, prve koja obavlja klasično usrednjavanje piksela i druga koja je odgovorna za očuvanje ivica [32, 34-36].

Nakon prolaska kroz blok za pretprocesiranje izračunavaju se HOG obeležja za ulaznu sliku, bilo da je u pitanju termalna slika ili slika vidljivog spektra. Izdvajanje HOG obeležja je implementirano na ćelijama od 8x8 piksela. Dimenzije preklapajućih blokova za normalizaciju su 2x2 ćelije sa 2 preklapajuće ćelije između susednih blokova. Za slike lica rezolucije 320x240 piksela koje su korišćene u ovom radu izdvajano je po 40716 HOG obeležja. Takođe, implementirani su nepotpisani pravci sa devet diskretnih pravaca [26, 32, 36].

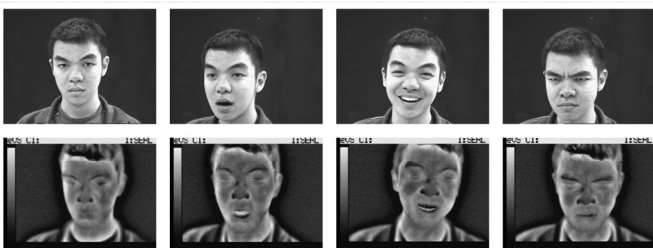
Izvučena HOG obeležja se prosleđuju SVM klasifikatoru. Višeklasni SVM realizovan u obliku „jedan protiv svih“ sa Gausovim kernel trikom je korišćen za klasifikaciju. Nakon obrade dobijenih vektora obeležja, SVM klasifikator daje rezultate poklapanja (eng. *matching scores*). Rezultat primene SVM klasifikatora je vektor sa rezultatima poklapanja za svaku od četiri emocije pojedinačno [32, 36].

Za realizaciju predloženog sistema za prepoznavanje emocija korišćen je Matlab programski paket.

4.2 Korišćena baza slika lica

Baza koja je korišćena prilikom implementacije opisanog sistema za prepoznavanje emocija sastoji se od slika lica vidljivog spektra i ogovarajućih slika lica termalnog spektra prikupljenih od 30 različitih osoba. Svaki par slika, jedna koja pripada spektru vidljive svetlosti i druga koja pripada termalnom spektru, je uslikan istovremeno u različitim položajima lica, pod različitim uslovima osvetljenja (dnevna svetlost, potpuni mrak, dodatno osvetljenje) i u različitim emocionalnim stanjima (normalno stanje, iznenađenost, sreća i ljutnja), ukupno 176-250 slika po osobi. Dimenzije svih slika u bazi je 320x240 piksela [37-39].

Skup podataka za trening se sastoji od 348 slika, pri čemu je za svaki izraz lica od svake osobe uzeto po 3 fotografije. Skup podataka za testiranje se sastoji od 116 slika pri čemu je za svaki izraz lica od svake osobe uzeta po jedna fotografija. Odabrane slike su uslikane frontalno kako orijentacija lica ne bi uticala na performanse prepoznavanja. Prilikom podele slika na skupove za trening i testiranje izvršena je deterministička podela pri čemu 75% slika pripada skupu za trening, a 25% skupu za testiranje.



Slika 7. Skup podataka za testiranje za četiri različita emocionalna stanja (neutralno emocionalno stanje, iznenađenost, radost, ljutnja)

Skup podataka za trening korišćen u ovom istraživanju sadrži za sva četiri tipa emocija po tri slike vidljivog i termalnog spektra za svaku osobu koje su uslikane u uslovima dnevnog svetla u različitim pozama. Za skup podataka za testiranje od svake osobe uzeta je po jedna slika lica iz oba domena za svaki od četiri tipa emocija, koje nisu korišćene u okviru skupa podataka za trening i koje se dovode na ulaz u sistem. Na slici 6 prikazana su dva različita skupa podataka za test, jedan za domen vidljive svetlosti, drugi za termalni domen, pri čemu su u oba slučaja pokrivena četiri tipa emocionalnog stanja: neutralno stanje, iznenađenost, radost i ljutnja.

5. REZULTATI I DISKUSIJA

U slučaju opisanog sistema za prepoznavanje emocija moguće je primeniti matricu konfuzije, kako bi se ocenile performanse algoritima i to za četiri različite klase pri čemu svaka klasa predstavlja jedno od četiri emocionalna stanja. U tabelama ispod predstavljene su konfuzione matrice pri čemu jedna odgovara termalnim slikama lica, a druga slikama lica koje pripadaju spektru vidljive svetlosti. Dijagonala matrice predstavlja tačno pozitivno (eng. *true positive*) prepoznate klase emocija. Ostala polja matrice predstavljaju netačne klasifikacije predloženog sistema. Tako na primer u slučaju termalnih slika lica, kao što se može videti iz matrice konfuzije, 25 klasifikovanih slika lica sa neutralnim emocionalnim stanjem su stvarno bile slike lica sa neutralnim emocionalnim stanjem. Međutim, 2 slike lica koje su klasifikovane kao emocionalno stanje „ljutnja“ su u stvari bile slike lica sa neutralnim emocionalnim stanjem.

Tabela 1. Matrica konfuzije za termalne slike lica

	Neutralno emocionalno stanje (stvarno)	Iznenađenost (stvarno)	Radost (stvarno)	Ljutnja (stvarno)
Neutralno emocionalno stanje (klasifikovano)	25	1	2	2
Iznenađenost (klasifikovano)	1	20	6	1
Radost (klasifikovano)	0	5	17	2
Ljutnja (klasifikovano)	2	5	4	18

Kada je reč o slikama lica koje pripadaju opsegu vidljive svetlosti, 15 slika lica koje su klasifikovane kao emocionalno stanje „radost“, stvarno su bile slike lica koje odgovaraju tom emocionalnom stanju. Međutim, osam slika lica koje su klasifikovane kao emocionalno stanje „radost“ u stvari su bile slike lica koje odgovaraju emocionalnom stanju „iznenađenost“.

Tabela 2. Matrica konfuzije za slike lica vidljivog spektra

	Neutralno emocionalno stanje (stvarno)	Iznenadenost (stvarno)	Radost (stvarno)	Ljutnja (stvarno)
Neutralno emocionalno stanje (klasifikovano)	24	2	1	3
Iznenadenost (klasifikovano)	0	20	4	3
Radost (klasifikovano)	0	8	15	5
Ljutnja (klasifikovano)	0	1	5	22

Tačnost predloženog sistema koja se dobija kada se broj tačno klasifikovanih emocionalnih stanja (zbir po dijagonali) podeli sa ukupnim brojem odrađenih klasifikacija (zbir cele matrice) iznosi 72% za termalne slike lica i 71% za slike lica koje pripadaju spektru vidljive svetlosti. Rezultati donekle zavise i od sposobnosti subjekata da u momentu uzimanja uzoraka prikažu zahtevano emocionalno stanje.

6. ZAKLJUČAK I BUDUĆI RAD

Dizajn sistema sledeće generacije za interakciju čoveka i računara mora da uključi suštinu emocionalne inteligencije – sposobnost prepoznavanja afektivnih stanja korisnika kako bi bili efektivniji i efikasniji. Ipak, dizajn i razvoj automatizovanog sistema koji je sposoban za tako nešto je prilično težak.

U ovom radu je dat pregled postojećih tehnika za prepoznavanje različitih emocionalnih stanja na osnovu karakterističnih obeležja izdvojenih iz govornog signala, rukopisa, načina upotrebe tastature i miša, kao i iz izraza lica. Prikazani su i eksperimentalni rezultati za prepoznavanje četiri različita emocionalna stanja na osnovu slika lica uz upotrebu HOG obeležja i SVM klasifikatora. Eksperiment je rađen na slikama lica koje pripadaju opsegu vidljive svetlosti, kao i na termalnim slikama lica, pri čemu je postignuta tačnost prepoznavanja preko 70%.

Budući rad bi podrazumevao upotrebu neuralnih mreža i dubokog učenja u procesu klasifikacije različitih emocionalnih stanja, kao i realizaciju multimodalnog sistema za prepoznavanje emocija koji bi podrazumevao izdvajanje karakterističnih obeležja iz slika lica kao i iz govornog signala, nakon čega bi se donosila konačna odluka o vrsti prepoznatog emocionalnog stanja. Takođe fuzija informacija koje se dobijaju iz slika lica koje pripadaju različitim spektralnim domenima sa ciljem da se poveća tačnost prepoznavanja emocija, će takođe biti tema nekih budućih istraživanja.

LITERATURA

- [1] A. Konar and A. Chakraborty, Emotion recognition - A pattern analysis approach, John Wiley & Sons, Inc., 2015.
- [2] A. Dix, J. Finlay, G. Abowed and R. Beale, Human-computer interaction, Pearson Education Limited, 2004.
- [3] C. Busso, Z. Deng, S. Yildirim and M. Bulut, "Analysis of emotion recognition using facial expressions, speech and multimodal information," in International Conference on Multimodal Interfaces, ICMI, 2004.
- [4] M. Pantic and L. Rothkrantz, "Toward an affect-sensitive multimodal human-computer interaction," Proceedings of the IEEE, vol. 91, no. 9, pp. 1370-1390, 2003.
- [5] S. Krothapalli and S. Koolagudi, Emotion recognition using speech features, Springer, 2013.
- [6] Affective computing - Focus on Emotion Expression, Synthesis and Recognition, I-Tech Education and Publishing, 2008.
- [7] G. Muhammad and et al., "A Facial-Expression Monitoring System for Improved Healthcare in Smart Cities," IEEE Access , vol. 5, pp. 10871-10881, 2017.
- [8] M. Alhussein, "Monitoring Parkinson's Disease in Smart Cities," IEEE Access, vol. 5, pp. 19835-19841, 2017.
- [9] S. Kedar, D. Bormane, A. Dhadwal and R. Agarwal, "Automatic Emotion Recognition through Handwriting Analysis: A Review," in International Conference on Computing Communication Control and Automation, Pune, India , 2015.
- [10] S. Basu, J. Chakraborty, A. Bag and M. Aftabuddin, "A review on emotion recognition using speech," in International Conference on Inventive Communication and Computational Technologies (ICICCT), Coimbatore, India , 2017.
- [11] P. Chandrasekar, S. Chapaneri and D. Jayaswal, "Automatic Speech Emotion Recognition: A survey," in International Conference on Circuits, Systems, Communication and Information Technology Applications (CSCITA), Mumbai, India, 2014.
- [12] S. Lugovic, I. Dundjer and M. Horvat, "Techniques and Applications of Emotion Recognition in Speech," in International Convention on Information and Communication Technology, Electronics and Microelectronics MIPRO, 2016..
- [13] Z. Zeng, M. Pantic, G. Roisman and T. Huang, "A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, no. 1, pp. 39-58, 2009.
- [14] A. Kolakowska, "A review of emotion recognition methods based on keystroke dynamics and mouse movements," in International Conference on Human System Interactions (HSI), Sopot, Poland , 2013.
- [15] E. Sariyanidi, H. Gunes and A. Cavallaro, "Automatic Analysis of Facial Affect: A Survey of Registration, Representation, and Recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence , vol. 37, no. 1, pp. 1113-1133, 2015.
- [16] C. K. Byoung, "A Brief Review of Facial Emotion Recognition Based on Visual Information," Sensors, vol. 18, no. 2, 2018.
- [17] J. Kumari, E. Rajesh and A. Kumar, "Fusion of features for the effective facial expression recognition," in 2016 International Conference on Communication and Signal Processing (ICCSP), Melmaruvathur, India , 2016.
- [18] C. Shan, S. Gong and P. W. McOwen, "Facial expression recognition based on Local Binary Patterns: A comprehensive study," Image and vision computing, vol. 27, no. 6, pp. 803-816, 2009.
- [19] Y. Tong, R. Chen and Y. Cheng, "Facial expression recognition algorithm using LGC based on horizontal and diagonal prior principle," Optik, vol. 125, no. 16, pp. 4186-4189, 2014.
- [20] N. U. Khan, "A comparative analysis of facial expression recognition techniques," in 3rd IEEE International Advance Computing Conference (IACC), Ghaziabad, India , 2013.
- [21] A. Kartali and et al., "Real-time algorithms for facial emotion : A comparison of different approaches," in Symposium on Neural Networks and Applications (NEUREL), Belgrade, 2018.

- [22] M. Pavlović, B. Stojanović, N. Vlahović and S. Stanković, "Comparison of Feature Descriptors for Visible and Thermal Face Recognition," in 18th International Symposium INFOTEH, Jahorina, 2019.
- [23] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, 2005.
- [24] O. Deniz, G. Bueno, J. Salido and F. De la Torre, "Face recognition using histograms of oriented gradients," Pattern Recognition Letters, vol. 32, no. 12, pp. 1598-1603, 2011.
- [25] K. R. Kakkirala, S. R. Chamala and S. R. Jami, "Thermal Infrared Face Recognition: A review," in UKSim-AMSS 19th International Conference on Modelling & Simulation, 2017.
- [26] C. Shu, X. Ding and C. Fang, "Histogram of the oriented gradient for face recognition," Tsinghua Science & Technology, vol. 16, no. 2, pp. 216-224, 2011.
- [27] S. Čabarkapa, N. Kojić, V. Radosavljević and B. Reljin, "An implementation of SVM in CBIR system," in TELFOR, Belgrade, 2008.
- [28] G. M. Fung and O. L. Mangasarian, "Multicategory proximal support vector machine classifiers," Machine learning, vol. 59, no. 1-2, pp. 293-300, 2005.
- [29] G. Bebis, A. Gyaourova, S. Singh and I. Pavlidis, "Face recognition by fusing thermal infrared and visible imagery," Image and vision computing, vol. 24, no. 7, pp. 727-742, 2006.
- [30] R. Gade, M. T. B., S. Z. Nielsen, H. S.-Petersen, H. J. Andersen, K. Basselbjerg, H. T. Dam, O. B. Jensen, A. Jørgensen, H. Lahrmann, T. K. O. Madsen, E. S. Bala and B. Ø. Povey, "Thermal imaging systems for real-time applications in smart cities," International Journal of Computer Applications in Technology, vol. 53, no. 4, pp. 291-308, 2016.
- [31] G. Hermosilla, J. Ruiz-del-Solar, R. Verschae and M. Correa, "A comparative study of thermal face recognition methods in unconstrained environments," Pattern recognition, vol. 45, no. 7, pp. 2445-2459, 2012.
- [32] M. Pavlović, B. Stojanović, R. Petrović and S. Stanković, "Fusion of visual and thermal imagery for illumination invariant face recognition system," in NEUREL, Belgrade, 2018.
- [33] W. R. Schwartz, H. Guo, J. Choi and L. S. Davis, "Face identification using large feature sets," IEEE transactions on image processing, vol. 21, no. 4, pp. 2245-2255, 2012.
- [34] Stojanović, B., Puzović, S., Vlahović, N., Petrović, R. and Stanković, S., 2018, November. Real-time multi-sensor infrared imagery enhancement. In 14th Symposium on Neural Networks and Applications (NEUREL) (pp. 1-6). IEEE, 2018.
- [35] R. Petrović, B. Stojanović, M. Pavlović and S. Stanković, "Thermal Imagery Noise Influence on Face Recognition Performance," in 18th International Symposium INFOTEH, Jahorina, 2019.
- [36] R. Petrović, B. Stojanović, M. Pavlović and S. Stanković, "Thermal-to-visible face recognition for illumination invariant systems," in IcETRAN, Palic, Serbia, 2018.
- [37] "IEEE OTCBVS WS Series Bench: DOE University Research Program in Robotics under grant DOE-DE-FG02-86NE37968; DOD/TACOM/NAC/ARC Program under grant R01-1344-18; FAA/NSSA grant R01-1344-48/49; Office of Naval Research under grant #N000143010022," [Online]. Available: <http://vci-pl-okstate.org/pbvs/bench/>.
- [38] B. Hernandez, G. Olague, R. Hammoud, L. Trujillo and E. Romero, "Visual learning of texture descriptors for facial expression recognition in thermal imagery," Computer vision and image understanding, vol. 106, no. 2-3, pp. 258-269, 2007.
- [39] S. M. Desa and S. Hati, "IR and visible face recognition using fusion of kernel based features," in 19th International Conference on Pattern Recognition, Tampa, FL, USA, 2008.



Ranko Petrović, Vlatacom institut visokih tehnologija, Beograd

Kontakt: ranko.petrovic@vlatacom.com

Oblast interesovanja: Biometrija, Papirozanvanje lica, Multimedije, Digitalna forenzika



Miloš Pavlović, Vlatacom institut visokih tehnologija, Beograd

Kontakt: milos.pavlovic@vlatacom.com

Oblast interesovanja: Biometrija, Papirozanvanje lica, Multimedije, Digitalna forenzika



Branka Stojanović, JOANNEUM RESEARCH DIGITAL, Graz, Austria

Kontakt: branka.stojanovic@joanneum.at

Oblast interesovanja: Informaciona bezbednost, Biometrija, Obrada signala, Digitalna forenzika



Snežana Puzović, Vlatacom institut visokih tehnologija, Beograd

Kontakt: snezana.puzovic@vlatacom.com

Oblast interesovanja: Digitalna obrada slike, Procena kvaliteta poboljšanja slike, Mašinsko ucenje

