ZAŠTITA PRIVATNOSTI U MULTIMEDIJSKIM SADRŽAJIMA

Deidentifikacija ili prikrivanje otkritog

PIŠE: **SLOBODAN RIBARIĆ**

Mogućnosti i sposobnosti nadzornih sustava su drastično povećane zahvaljujući neprekidnom napretku na područjima kao što su pametne mreže kamera, bežične mreže osjetnika (senzora) u različitim dijelovima elektromagnetskog spektra, polja audiosenzora, porazdijeljena inteligencija i obrada podataka. Iako je na prvi pogled jasno da postoje opravdani razlozi za nadzor i prikupljanje multimedijskih podataka (npr. provođenje zakona, zaštita od terorističkih napada, sigurnost prometa, forenzika, predviđanje kritičnih situacija), postoji snažna potreba za zaštitom privatnosti nedužnih osoba koje su neminovno snimljene te čiji su zvučni i vizualni identiteti zabilježeni i pohranjeni.

Da bismo predočili kakav je porast broja nadzornih sustava u javnim prostorima i njihov utjecaj na privatnost treba spomenuti da se na brojnim križanjima i parkiralištima u Zagrebu nalaze kamere te da se neprekidno snima na ulazima gotovo svih javnih institucija. U svijetu je problem još izraženiji. Naprimjer, poznato je da u Ujedinjenom Kraljevstvu ima više od četiri milijuna CCTV kamera. Analize su pokazale da je prosječan stanovnik Londona dnevno zabilježen oko tristo puta. Problem zaštite privatnosti je još izraženiji ako se zna da više od 80 % nadzornih sustava koji se nalaze u poslovnom dijelu Londona ne poštuje odgovarajuće zakonske propise za zaštitu osobnih podataka. Dodatni problem zaštite privatnosti u današnjem umreženom društvu je napredak tehnologija kao što su *Google Sreet View* i *EveryScape* koje omogućuju panoramski prikaz mnogih ulica u svijetu, društvene mreže, biometrika, veliki skupovi podataka i tehnologije za njihovo pretraživanje i analizu. Sve to predstavlja dodatan prostor za prodor u privatnost pojedinaca. Julia Angwin u svojoj knjizi *Nacija pod intenzivnim elektroničkim nadzorom* analizira odnose privatnosti, sigurnosti i ljudskih sloboda u kontekstu elektroničkog nadzora i zaključuje da živimo u svijetu neselektivnog praćenja u kojemu institucije od Googlea do NSA-a (*National Security Agency)* gomilaju i pohranjuju podatke o osobama do neslućenih razmjera. Ona zaključuje da je to neselektivno i nekritičko praćenje pojačano...*tehnologijama koje mi svi tako volimo, kao što su osobna računala, prijenosna računala, pametni mobiteli i mrežni servisi.*

lako postoje opravdani razlozi za prikupljanje multimedijskih podataka, nužno je zaštititi privatnost osoba koje su nenamjerno snimljene te čiji su zvučni i vizualni identiteti zabilježeni i pohranjeni.

●●●

Na brojnim križanjima i parkiralištima u Zagrebu već se nalaze kamere, a neprekidno se snima i na ulazima javnih institucija no, to je beznačajno u usporedbi s primjerice Ujedinjenim Kraljevstvom gdje snima više od četiri milijuna CCTV kamera. Prosječan stanovnik Londona dnevno je zabilježen tristo puta.

U kontekstu tog članka možemo definirati sljedeće ključne pojmove:

- ***osobna informacija*** je bilo kakva informacija koja se odnosi na osobu
- ***osobni identifikator*** je osobna informacija koja omogućuje identifikaciju osobe.

Izrazi deidentifikacija i *anonimizacija* se vrlo često koriste kao istoznačnice, no neki ih eksperti razlikuju. Deidentifikacija se odnosi na reverzibilni ili obrativi proces uklanjanja ili prikrivanja osobnog identifikatora. Anonimizacija predstavlja proces deidentifikacije koji onemogućuje rekonstrukciju izvornog osobnog identifikatora, odnosno to je nereverzibilni postupak.

Privatnost

Nema jedinstvene definicije privatnosti. Privatnost ima duboke korijene u sociološkim i antropološkim raspravama o značenju, vrijednosti i načinima očuvanja privatnosti u različitim kulturama. S pravnog gledišta prvu su definiciju privatnosti dali Louis D. Brandeis i Samuel D. Warren prije više od 120 godina. Oni su privatnost definirali kao *the right to be let alone*, odnosno pravo osobe da bude *ostavljena na miru* u kontekstu prikupljanja i distribucije osobnih informacija, posebice od neautoriziranih natpisa, fotografija i drugih medijskih sadržaja. Također, prema Brandeisu i Warrenu osoba treba biti zaštićena od vladinih istraga i posezanja u sferu njene samoće koja se podrazumijeva razumnom i društveno prihvatljivom.

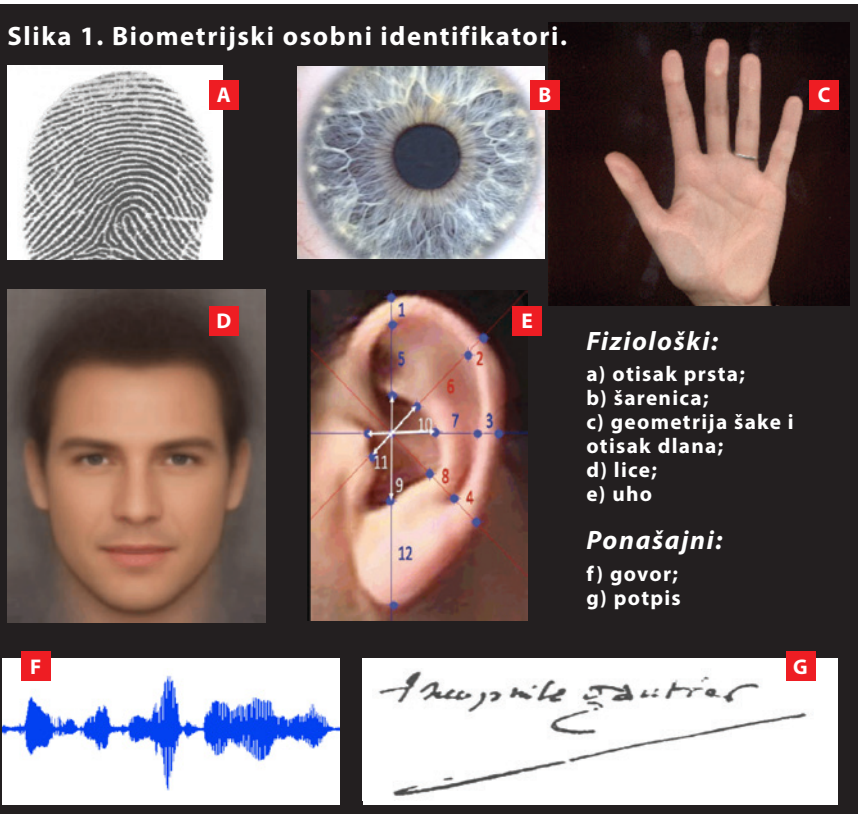
Alan F. Westin definira 2003. godine privatnost kao pravo pojedinca da određuje koju informaciju o sebi želi razotkriti drugima.

Iscrpan i sveobuhvatni uvid u teoriju privatnosti, postojeće pokušaje konceptualizacije privatnosti te različite definicije privatnosti s gledišta pravnika, filozofa i sociologa dan je u knjizi Daniela J. Solovea *Razumijevanje privatnosti*. Sedamdesetih godina prošlog stoljeća europske su zemlje započele donošenje zakona o zaštiti privatnosti: Švedska (1974.), Njemačka (1977.), Francuska (1978)... dok je temeljni okvir zaštite privatnosti i osobnih podataka u Europskoj uniji donesen 1995. godine (*The 1995 Data Protection Directive of the European Union* - Directive 95/46/EC).

Osobni identifikatori u multimedijskim sadržajima

Ovdje predložena klasifikacija osobnih identifikatora u multimedijskim sadržajima temelji se na načelima sadržanim u

Slika 1. Biometrijski osobni identifikatori.



tzv. *Sigurnoj luci* (izvorno: *Safe Harbour*) Na temelju 18 tipova osobnih identifikatora koji se navode u Sigurnoj luci, predlažemo njihovu sljedeću podjelu:

- *nebiometrijski osobni identifikatori* kao što su tekstualni sadržaji, govorni kontekst, specifični društveno-politički sadržaji, način odijevanja, frizura, registarske tablice vozila i sl.
- *biometrijski osobni identifikatori* koji su definirani kao mjerljive, nepromjenljive i jedinstvene osobne karakteristike koje se upotrebljavaju za identifikaciju osoba. Biometrijski osobni identifikatori se mogu podijeliti na *fiziološke* (npr. lice, šarenica, uho, otisak prsta ili dlana) i na *ponašajne* (npr. govor, hod, potpis, pokreti usana, gestikulacija, dinamika tipkanja)
- *neizraziti osobni biometrijski identifikatori* su fiziološke ili ponašajne značajke koje nisu nužno stalne i ne nude dovoljno visoku razinu razlikovnih sposobnosti (npr. visina, težina, boja očiju, silueta, dob, spol, rasa, pjege, madeži, ožiljci i tetovaže). Važno je naglasiti da neizraziti osobni biometrijski identifikatori ne mogu omogućiti pouzdanu identifikaciju osoba ali se mogu iskoristiti za poboljšanje performansi sustava za raspoznavanje osoba ili za razvrstavanje osoba u određene kategorije što može biti osjetljivo s gledišta privatnosti.

Na slici 1. prikazani su fiziološki i ponašajni biometrijski identifikatori koji se najčešće primjenjuju u biometrijskoj identifikaciji osoba.

Vrlo često u multimedijskim sadržajima su istodobno prisutni biometrijski, neizraziti biometrijski i nebiometrijski identifikatori te je potrebno deidentificirati sve osobne identifikatore - u tom slučaju govorimo o *višenačinskoj deidentifikaciji*.

Detekcija i prikrivanje, uklanjanje ili zamjena osobnih identifikatora u multimedijskim sadržajima je interdisciplinarni izazov koji uključuje znanstvena područja kao što su: obrada prirodnog jezika, obrada teksta, obrada slika, raspoznavanje uzoraka, strojno učenje, analiza govora, obrada signala i biometrija.

Deidentifikacija nebiometrijskih osobnih indikatora

Deidentifikacija teksta

Istraživanja u području deidentifikacije teksta zbog zaštite privatnosti osoba počela su s deidentifikacijom teksta zdravstvenih dosjea pacijenata. Pristup deidentifikaciji temeljio se na uklanjanju brojnih specifičnih kategorija informacija iz tekstualnog zapisa te na njihovoj zamjeni s nadomjesnom/lažnom informacijom. Automatska deidentifikacija teksta zdravstvenih dosjea je usredotočena na visokostrukturirane specifične tipove zapisa, ali i na slobodan tekst medicinskih zapisa s vrlo promjenljivom strukturom. Deidentifikacijske metode se temelje na predlošcima i specijaliziranom znanju potrebnom za zamjenu osobnih zdravstvenih podataka u dosjeima, a mogu se temeljiti i na složenoj kombinaciji rječnika i algoritmima za analizu teksta. U novije vrijeme pristupi deidentifikaciji teksta zasnivaju se na kombinaciji strojnog učenja, heurističkih i statističkih metoda te podudaranja s predloškom. U postupku zaštite privatnosti u tekstovima se umjesto anonimizacije koristi deidentifikacija, odnosno obrativi proces koji uz uporabu tajnog ključa dopušta prikaz izvornih osobnih podataka.

Deidentifikacija načina odijevanja i frizure

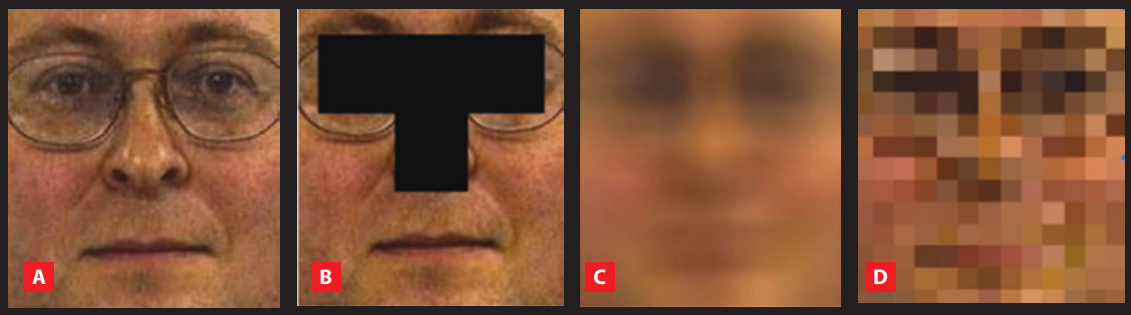
Poznato je da način odijevanja i frizura nose informaciju koja (djelomično) otkriva identitet osobe i može se koristiti za razvrstavanje ljudi u različite kategorije. Poznat je i problem *a pair-wise constraint* koji se ogleda u tome da je moguće odrediti da dva deidentificirana lica pripadaju istoj osobi u videu i to na temelju odjeće ili frizure. Uporaba govornog, specifičnog društvenog i političkog konteksta te informacije o okolini u kojoj se nalazi osoba može pomoći u otkrivanju njenog identiteta. Nažalost, vrlo je malo napravljeno na području uklanjanja ili prikrivanja takvog konteksta.

Deidentifikacija fizioloških biometrijskih osobnih indikatora

Deidentifikacija lica u slikama i videu

Lice je glavni fiziološki biometrijski identifikator u multimedijskim sadržajima. Ono osim *identifikacije*, otkriva dob, spol, emotivno raspoloženje i zdravstveno stanje i predstavlja osjetljivu značajku privatnosti. Zbog svega toga se zahtijeva deidentifikacija lica u cilju zaštite privatnosti. Rana istraživanja na području deidentifikacije lica bila su usmjerena na deidentifikaciju lica na slikama i razvijene su jednostavne tzv. *naivne metode* kao što su prekrivanje dijela lica crnim pravokutnikom ili mnogokutom (*black box method*), zamućenje (*blurring*) i pikselizacija (*pixelation*) - smanjenje rezolucije slike. Slika 2. prikazuje rezultate deidentifikacije slike lica uporabom naivnih metoda.

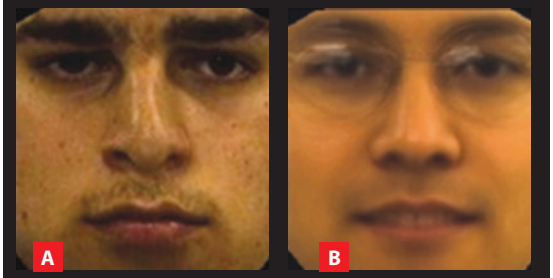
Slika 2. Primjer deidentifikacije lica (S. Ribarić, N. Pavešić, 2015.):



a) izvorna slika;
b) prekrivanje dijela lica;
c) zamućenje;
d) pikselizacija

Naivne metode deidentifikacije mogu spriječiti da osobu identificira čovjek, no ne mogu prevariti automatske sustave za raspoznavanje. Poznati su strojni postupci otkrivanja identiteta u tako deidentificiranim slikama lica koji se temelje na tzv. oponašanju postupka (*parrot recognition*). Za postizanje više razine zaštite privatnosti upotrebljavaju se metode deidentifikacije koje se temelje na uporabi svojstvenih lica (*eigenface*) u kojima se lice prikazuje s određenim, manjim brojem svojstvenih lica dobivenim Karhunen-Loeveom transformacijom (KLT). U novije vrijeme upotrebljavaju se složeniji postupci deidentifikacije lica koji se zasnivaju na postupku zamjene izvornog lica nekim srednjim licem iz galerije slika lica: *k*-istih lica (*k-Same*), *k*-istih izabranih lica (*k-Same-Select*) i *k*-istih lica temeljenih na modelu (*Model-based k-Same*). Slika 3. prikazuje rezultat deidentifikacije metodom *k*-istih lica temeljenom na modelu.

Slika 3. Deidentifikacija lica uporabom k-istih lica temeljenom na modelu (S. Ribarić, N. Pavešić, 2015.):



a) izvorna slika lica: b) deidentificirana slika lica

Poseban izazov na području zaštite privatnosti je automatska deidentifikacija lica u videu, zato što lice treba biti detektirano, lokalizirano i deidentificirano u svakom slikovnom okviru (*frame*). Lice koje nije deidentificirano samo u jednom slikovnom okviru kompromitira osobu u videozapisu i omogućuje njenu identifikaciju. Automatska deidentifikacija u videu sastoji se od detekcije i lokalizacije lica te njegovog prikrivanja.

Za detekciju i lokalizaciju lica u videu koriste se složeni postupci poznati iz računalnog vida kao što su: umjetne neuronske mreže, Schneiderman-Kanade i Viola-Jones detektori i njihove izvedenice, histogrami orijentacije gradijenta (HOG), kombinacija oduzimanja pozadine slike i vreće (*bag*) segmenata i stroja s potpornim vektorima SVM (*Support Vector Machine*).

U novije vrijeme primjenjuju se vrlo složeni postupci za detekciju i lokalizaciju lica koji koriste višestruke kanale za registraciju videa i izvode linearne i nelinearne transformacije ulazne slike (histogrami gradijenata, različiti prostori boja RGB, HSV, CIELUV, amplitude gradijenta, Gaborovi i DoG (Difference of Gaussian) filtri. Praćenje i detekcija, odnosno kombinacija prostorne i vremenske podudarnosti između slikovnih okvira, može povećati djelotvornost postupka lokalizacije lica. Nakon uspješne lokalizacije lica u slikovnom okviru slijedi njegovo prikrivanje. Mogu se rabiti i metode za deidentifikaciju lica u stacionarnim slikama. Alternativni pristup prikrivanja lica u videonadzornim sustavima temelji se na postupcima izobličenja područja lica uporabom kodiranja miješanjem (*scrambling*) koji su obrativi.

U posljednjih nekoliko godina pojavili su se videonadzorni sustavi sa zaštitom privatnosti u stvarnom vremenu. Spomenimo neke: *Respectful Cameras system* - sustav u kojem se zahtijeva da osobe čija privatnost treba biti zaštićena nose kape ili prsluke u boji. Samo njihova lica će se u videosekvencama prekriti eliptičnim bijelim područjem. *DSP-based PrivacyCam* sustav štiti privatnost osoba u videu uporabom kodiranja područja lica miješanjem. *TrustCam prototype system* - sastoji se od mreže kamera koje su opremljene posebnim modulom za kodiranje područja lica u videu. *De-identification Camera* predstavlja sklopovsko rješenje automatskog praćenja i prikrivanja lica na razini osjetnika (kamere) uporabom naivnih metoda deidentifikacije lica.

Dodatne probleme zaštiti privatnosti uzrokuju napredne tehnologije kao što su *Google Sreet View* i *EveryScape* koje omogućuju panoramski prikaz mnogih ulica u svijetu, društvene mreže, biometrika, veliki skupovi podataka te tehnologije za njihovo pretraživanje i analizu

Složeniji postupci zaštite privatnosti u videu, koji se razvijaju u posljednje vrijeme, upotrebljavaju aktivne modele izgleda (*active appearance model*) u kojima se slike lica prije deidentifikacije grupiraju prema izrazu, spolu i pozi. Svaka je takva grupa predstavljena aktivnim modelom izgleda. Izvorna se slika lica podudara sa svakim aktivnim modelom izgleda te se izvorna slika zamjenjuje u skladu sa značajkama grupe. Tako se poboljšava prirodnost i uporabljivost deidentificiranog videa.

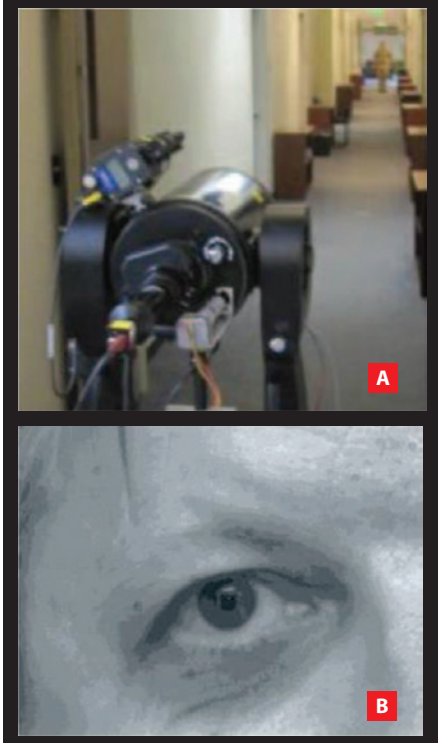
Usprkos intenzivnim istraživačkim naporima u području deidentifikacije lica u videu, ostaju još brojni neriješeni problemi i izazovi, kao što su: lokalizacija lica u različitim uvjetima osvjetljenja scene, djelomično prekrivena lica, različite poze lica, prisutnost tzv. strukturnih komponenti kao što su naočale, sunčane naočale, brada, brkovi. Poseban problem je lokalizacija i deidentifikacija lica u stvarnom vremenu za nadzorne sustave u kojima se pojavljuje veliki broj osoba (*crowd scene*).

Deidentifikacija šarenice, uha i otiska prsta

Šarenica je važan biometrijski identifikator koji zbog jedinstvenosti i male promjenjivosti s vremenom omogućuje pouzdan i neinvazivni postupak identifikacije osoba. Većina sustava za

identifikaciju na temelju šarenice - zbog njenih malih dimenzija - zahtijeva suradnju osobe i obično se šarenica snima na maloj udaljenosti (između 15 do 50 cm). Zbog toga na prvi pogled ne postoji potreba za deidentifikaciju šarenice. No, razvijeni su sustavi koji omogućuju akviziciju slike šarenice i identifikaciju na udaljenosti (*Iris-at-a-Distance - IAAD*) do 30 m i to bez suradnje osobe (slika 4). U takvim slučajevima zbog zaštite privatnosti treba deidentificirati šarenicu. Jedan od rijetkih sustava za deidentifikaciju područja očiju pa i šarenice temelji se na izobličenju područja lica uporabom kodiranja miješanjem.

Slika 4. Sustav za akviziciju i identifikaciju na temelju šarenice IAAD (De Villar et al., 2010.):



a) izgled sustava, b) primjer slike područja šarenice snimljen na udaljenosti od 30 metara.

Uz lice i šarenicu i *uho* je važan fiziološki biometrijski identifikator koji omogućuje neinvazivnu identifikaciju ili verifikaciju na udaljenosti. Mnogi su istraživači suglasni da je uho, kao biometrijska značajka, pogodno za identifikaciju osoba u videonadzornim sustavima. Autoru ovog članka nije poznato postoji li komercijalni sustav za identifikaciju na temelju uha, iako postoje brojni razvijeni eksperimentalni sustavi za identifikaciju i verifikaciju na temelju dvodimenzionalnih (2D) i trodimenzionalnih (3D) slika uha.

U bližoj budućnosti, zbog razvoja relativno jeftinih videokamera i teleskopske opreme, može se očekivati razvoj komercijalnih sustava identifikacije na temelju uha u djelomično ili potpuno neupravljivim uvjetima vanjskih scena. To će sigurno inicirati istraživanja i razvoj metoda deidentifikacije uha zbog zaštite privatnosti. Većina sustava za identifikaciju na temelju uha kombinira sliku profila lica i sliku uha tako da se zahtijeva višenačinska deidentifikacija - lica i uha.

Slike *otisaka prstiju* kao multimedijski dokumenti na prvi pogled ne zaslužuju posebnu pozornost i to zbog dva razloga: a) osoba je kooperativna i u većini slučajeva dragovoljno daje otisak prsta, b) u središtu pozornosti ovog članka je deidentifikacija multimedijskih sadržaja koji se prikupljaju na udaljenosti. No, na temelju rezultata najnovijih istraživanja (*Technology Review,* 2015.) moguće je detektirati, lokalizirati i snimiti otiske prstiju na udaljenosti od dva metra i to bez znanja i sudjelovanja osobe. Na temelju tako snimljenog otiska moguće je utvrditi identitet osobe.

Takva tehnologija predstavlja prijetnju privatnosti u bližoj budućnosti. Štoviše, otisci prstiju osim identifikacije nose vrlo osjetljive privatne informacije o osobi kao što je spol i etnička pripadnost, ali i informaciju o bolesti kao što je, naprimjer, Parkinsonova i Alzheimerova bolest. Privatnost se štiti različitim tehnikama izobličenja otisaka prstiju uporabom morfoloških postupaka i postupaka kodiranja miješanjem. Slika otiska prsta može se deidentificirati uporabom naivnih metoda koje se koriste za deidentifikaciju slika lica ili pak zamjenom slike otiska prstiju slikom sintetiziranog otiska. Jedan od zanimljivih pristupa deidentifikaciji otiska prstiju je i deidentificirani otisak koji se dobiva kombiniranjem dvaju otisaka prstiju različitih ili istih osoba. U tom slučaju deidentificirani otisak prsta izgleda poput prirodnog otiska.

Deidentifikacija ponašajnih osobnih identifikatora

Deidentifikacija glasa/govora

Za biometrijske identifikatore kao što su lice, šarenica i uho kažemo da nose *vizualni identitet* osobe dok glas nosi *zvučni identitet*. Ljudski glas je jedinstveni uzorak za svaku osobu, tj. ne postoje dvije osobe čiji glasovi zvuče identično i koristi se za automatsku identifikaciju ili verifikaciju osoba. Osim informacije o identitetu glas sadrži osjetljive privatne podatke kao što su spol, dob, emotivno stanje, zdravstveno stanje i sl. Tehnologije i servisi, kao što su audio/videonadzor, govorni servisi i sustavi za cijelodnevno praćenje i bilježenje aktivnosti, omogućuju identifikaciju osoba.

●●● Deidentifikacija se odnosi na uklanjanje ili prikrivanje osobnog identiteta snimljene osobe. Pristup deidentifikaciji temelji se na uklanjanju brojnih specifičnih kategorija informacija iz multimedijskog zapisa te na njihovoj zamjeni nadomjesnom/lažnom informacijom.

Deidentifikacija glasa se temelji na njegovoj transformaciji (*voice transformation*) koja predstavlja modifikaciju nelingvističkih značajki govornog trakta bez utjecaja na sadržaj i razumljivost govora. Značajke koje se modificiraju su, naprimjer, energija, visina glasa i vremenska skala, odnosno promjena duljine trajanja glasa. Jedan od pristupa deidentifikaciji glasa temelji se pretvorbi glasova različitih govornika u isti sintetizirani ciljni glas (*target voice*). Jedan od zanimljivih novih pristupa deidentifikaciji glasa koristi unaprijed izračunane govorne transformacije na bazi modela Gaussovih mješavina (GMM - *Gaussian Mixture Model*) za deidentifikaciju glasa novog govornika.

Na području deidentifikacije glasa u stvarnom vremenu postoje brojni izazovi kao što su deidentifikacija glasa u uvjetima okolišne buke i pozadinskog šuma te istodobni govor više osoba.

Deidentifikacija hoda i gestikulacije

Način hodanja predstavlja ponašajni biometrijski identifikator te može poslužiti za identifikaciju osoba ili za dijagnosticiranje nekih oboljenja. Uz dinamiku i način koračanja hod sadrži informaciju o izgledu osobe (silueta, duljina ruku i nogu) te informaciju o dobi i spolu. Nadzorni sustavi koji su danas neizbježni u svakodnevnom životu, zahvaljujući razvoju tehnika računalnog vida i raspoznavanja uzoraka, omogućuju raspoznavanje (identifikaciju ili verifikaciju) nekooperativnih osoba na temelju hoda.

Samo se nekoliko istraživačkih studija bavi problemom deidentifikacije hoda. Jedna opisuje deidentifikaciju koja se temelji na detekciji pokreta u videosekvencama i deidentifikaciji područja u kojemu se nalazi osoba uporabom kodiranja miješanjem na bazi diskretne kosinusne

transformacije (DCT). Drugi pristup se temelji na zamućenju volumnih dijelova x,y,t (*voxel*) koji su definirani kao područja u kojima je detektirana osoba u slijedu slikovnih okvira.

Jedan od glavnih problema u deidentifikaciji hoda u videonadzornim sustavima je kako prikriti značajke hoda, a istodobno očuvati uporabnost i prirodnost deidentificiranog videozapisa.

Gestikulacija se može definirati kao pokreti dijelova tijela (prstiju, šake, ruke, glave ili lica) ili pak cijelog tijela s namjerom ili bez namjere da imaju neko značenje. Poznata je činjenica da se gestikulacija razlikuje od čovjeka do čovjeka te da može biti korištena za raspoznavanje osoba. Razvijeni su prototipovi sustava za raspoznavanje osoba na temelju gestikulacije rukama, ali koliko je autoru ovog članka poznato još ne postoje sustavi za deidentifikaciju gestikulacije rukama.

Deidentifikacija neizrazitih biometrijskih osobnih identifikatora

Neizraziti biometrijski osobni identifikatori, kao što su naprimjer: visina, težina, spol, boja očiju, rasa, silueta, madeži, tetovaže i sl. izravno ne omogućuju identifikaciju osoba, ali dopuštaju razvrstavanje u različite kategorije.

Spomenimo tri glavna načina korištenja neizrazitih biometrijskih osobnih identifikatora:

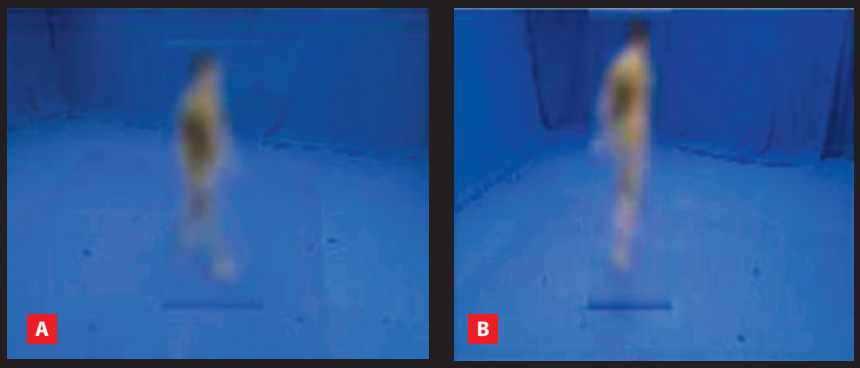
- identifikacija ili verifikacija uporabom verbalnog opisa neizrazitih biometrijskih identifikatora
- identifikacija ili verifikacija uporabom fuzije biometrijskih i neizrazitih osobnih identifikatora u cilju povećanja točnosti biometrijskog sustava
- pretraživanje velikih biometrijskih baza.

Bez obzira na navedene načine uporabe neizrazitih biometrijskih osobnih identifikatora jasno je da oni sadrže osjetljive informacije o privatnosti te da trebaju biti deidentificirani u multimedijskim dokumentima.

Deidentifikacija siluete tijela

Silueta tijela je važna neizrazita biometrijska značajka koja pomaže u procesu raspoznavanja osobe. Ona se koristi sama ili u kombinaciji s drugim biometrijskim identifikatorima. Nadalje, predstavlja značajku koja omogućuje tzv. reidentifikaciju, odnosno praćenje osoba u nadglednim sustavima s više kamera čija se vidna polja ne prekrivaju. Objavljeno je samo nekoliko radova koji se odnose na deidentifikaciju siluete tijela. Deidentifikacija se temelji na dilataciji i zamućenju siluete, ili na kombinaciji linearne integralne konvolucije (LIC) i eksponencijalnog zamućenja polja slikovnih elemenata koje odgovara mjestu siluete u videu. Slika 5. prikazuje primjer deidentifikacije siluete tijela u videu koji snima aktivnost osobe.

Slika 5. Deidentifikacija osobe u videu pri čemu je sačuvana informacija o vrsti aktivnosti osobe (Ivasic-Kos et al., 2014.):



a) hodanje; b) skakanje.

Jedan zanimljiv pristup deidentifikaciji siluete tijela koristi zamjenu osobe s drugom osobom iz zadane galerije slika.

Uvjet uspješne deidentifikacije siluete tijela je dobra segmentacija slike na pozadinu i objekt od interesa, međutim zbog složenosti okoline i nestacionarnosti pozadine, promjena uvjeta osvjetljenja i trešenja kamere, detekcija siluete je daleko od zadovoljive točnosti. Uz sve to javlja se i problem prikrivanja siluete uz uvjet očuvanja prirodnosti deidentificiranog videa.

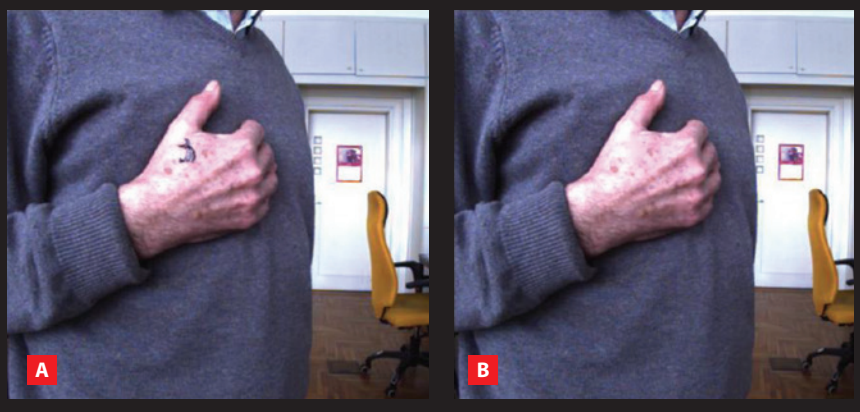
Deidentifikacija spola, rase i etničke pripadnosti

U znanstvenoj literaturi mnogo je radova koji se odnose na automatsko raspoznavanje spola, dobi, rase i etničke pripadnosti, no vrlo malo je napravljeno na njihovoj deidentifikaciji. Informacija o navedenim neizrazitim biometrijskim identifikatorima dobivaju se iz slika lica, glasa, hoda i siluete tijela i njihovom kombinacijom. U literaturi se navodi da je prikrivanje rase ili spola težak problem ako se želi sačuvati prirodnost i uporabnost deidentificiranog videa. Naprimjer, pokušaj prikrivanja boje kože - koja je u uskoj vezi s rasom - uporabom različitih transformacija boje, uzrokuje smanjenje prirodnosti deidentificiranog videa.

Deidentifikacija ožiljaka, madeža i tetovaže (*SMT - Scars, Marks, Tattoos*)

Ožiljci, madeži i tetovaže su oznake na koži koje daju vredniju informaciju potrebnu za identifikaciju osoba negoli je to dob, spol, visina ili rasa. Istraživanja su pokazala da oznake na licu kao što su pjege, madeži i ožiljci mogu poboljšati performansu sustava za automatsko raspoznavanje lica ili sustava za pretraživanje baze lica. Razvijeni su i postupci za detekciju ožiljaka, madeža i tetovaže za slike osoba koje su snimljene u stvarnim uvjetima i ti se postupci koriste u različi-

Slika 6. Deidentifikacija tetovaže (D. Marčetić, S. Ribarić, et al., 2014.):



a) izvorna (sirova) slika; b) Deidentificirana tetovaža - prekrivena tetovaža i njeno mjesto

tim forenzičkim scenarijima.

Tetovaža se prvenstveno koristi u automatskim sustavima za pretraživanje slika CBIR (*Content-Based Image Retrieval*) u kriminalistici i forenzici. Izgled tetovaže i njeno mjesto na tijelu mogu se upotrijebiti za identifikaciju osobe te ona predstavlja osjetljivu značajku privatnosti. Osim jednog rada (D. Marcetić, S. Ribarić et al., 2014.) nema objavljenih radova o deidentifikaciji ožiljaka, pjega, madeža i tetovaže. Prototip sustava za deidentifikaciju tetovaže koji je razvijen na FER-u Sveučilišta u Zagrebu sastoji se od modula za detekciju kože i područja interesa ROI (*Region of Interest*) - potencijalnog područja tetovaže, modula za izlučivanje značajki, baze tetovaža, modula za podudaranje i detekciju tetovaže te modula za prikrivanje tetovaže i njenog mjesta na koži.

Otvorena pitanja

Deidentifikacija je jedna od glavnih metoda zaštite privatnosti u umreženom društvu. Istraživanja u području deidentifikacije multimedijskih sadržaja su još u povojima i pred znanstvenicima s tog područja stoje mnogi izazovi kao što su deidentifikacija fizioloških i ponašajnih biometrijskih identifikatora u kombinaciji s nebiometrijskim identifikatorima, kao što su govorni kontekst, specifični društveno-politički sadržaji, način odijevanja i frizura.

I u području deidentifikacije biometrijskih osobnih identifikatora postoje brojni problemi kao što su njihova detekcija, lokalizacija, praćenje u stvarnom vremenu te prikrivanje, uklanjanje ili zamjena originalnih identifikatora. Ti su problemi posebno izraženi za nadzorne video/audiosustave u kojima se pojavljuje više ljudi. Spomenimo još neka otvorena pitanja: Kakvu metriku koristiti za mjerenje uspješnosti zaštite privatnosti postupkom deidentifikacije? Kako mjeriti prirodnost i upotrebljivost deidentificiranih multimedijskih sadržaja? Koje su deidentifikacijske metode primjenljive u stvarnom vremenu? ■

Autor **Slobodan Ribarić** redoviti je profesor na Fakultetu elektrotehnike i računarstva Sveučilišta u Zagrebu. Voditelj je COST Action projekta *Deidentifikacija za zaštitu privatnost u multimedijskim sadržajima* (http://www.cost.eu/COST_Actions/ict/IC1206) i 6733 DePPSS istraživačkog projekta Hrvatske zaklade za znanost koja je financirala rad na temelju kojega je pripremljen ovaj članak.

Profesor Ribarić je voditelj međunarodne specijalne sekcije BiForD *(Biometrics & Forensics & De-identification and Privacy Protection)* na skupu MIPRO. Autor se zahvaljuje prof. Nikoli Pavešiću s Fakulteta za elektrotehniko u Ljubljani i prof. Aladdinu Ariyaeeiniaju, *University of Hertfordshire* (UK) za korisne savjete tijekom pisanja članka.