

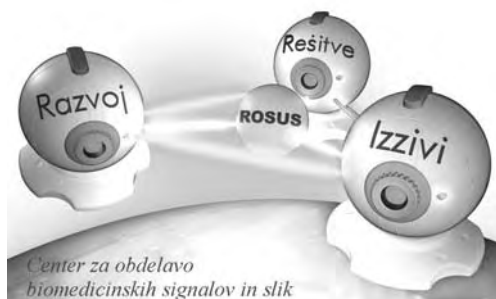


Univerza v Mariboru

*Fakulteta za elektrotehniko,
računalništvo in informatiko*



Inštitut za
računalništvo



ROSUS 2013

Računalniška obdelava slik in njena uporaba v Sloveniji 2013

Zbornik 8. strokovne konference

Urednik
Božidar Potočnik

Maribor, 21. marec 2013

Zbornik prispevkov strokovne konference
ROSUS 2013 Računalniška obdelava slik in njena uporaba v Sloveniji 2013
Maribor, 21. marec 2013

Elektronska izdaja

CIP - Kataložni zapis o publikaciji
Univerzitetna knjižnica Maribor

004.932 (082)

STROKOVNA konferenca ROSUS (8 ; 2013 ; Maribor)

ROSUS 2013 [Elektronski vir] : računalniška
obdelava slik in njena uporaba v Sloveniji 2013 :
zbornik 8. strokovne konference, Maribor, 21.
marec 2013 / urednik Božidar Potočnik. - El.
zbornik. - Maribor : Fakulteta za elektrotehniko,
računalništvo in informatiko, Inštitut za
računalništvo, 2013

Dostopno tudi na:

<http://cobsis.uni-mb.si/rosus2013/>

ISBN 978-961-248-376-0

1. Potočnik, Božidar

COBISS.SI-ID 73372673

ISBN 978-961-248-376-0



Programski odbor ROSUS 2013:

dr. Božidar Potočnik
dr. Damjan Zazula
dr. Boris Cigale
dr. Nikola Guid
dr. Baldomir Zajc
dr. Tomaž Tomažič
dr. Slobodan Ribarić
dr. Veljko Vlasisavljević
dr. Saša Divjak

dr. Stanislav Kovačič
dr. Zdravko Kačič
dr. Aleš Leonardis
dr. Janez Perš
dr. Mitja Lenič
dr. Franc Solina
dr. Andrej Šoštarič
dr. Franjo Pernuš
dr. Eldar Gadžijev

Organizacijski odbor 2013:

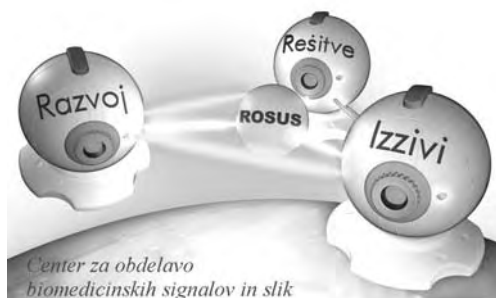
dr. Božidar Potočnik
dr. Damjan Zazula
dr. Danilo Korže
dr. Boris Cigale
dr. Smiljan Šinjur
Jurij Munda
Vojko Glaser
Sebastijan Šprager
Gašper Sedej

Spoštovani!

Po sedmih uspešnih konferencah ROSUS 2006–2012 želimo s konferenco ROSUS 2013 spet promovirati pomembnost ekonomske koristnosti računalniške obdelave slik na področjih industrije, biomedicine in drugih poslovnih procesov. Vezi, ki smo jih na prejšnjih konferencah stkali med raziskovalci, razvijalci, ponudniki rešitev ter uporabniki računalniške obdelave slik v slovenskem prostoru, želimo še dodatno okrepiti, ob tem pa nuditi tudi možnosti sklepanja novih sodelovanj in svetovanja pri razreševanju konkretnih poslovnih oziroma raziskovalnih problemov.

Tudi letos namenjamo glavni poudarek aplikacijam s področja računalniške obdelave slik, ki so že integrirane oziroma pripravljene za integracijo v poslovne procese. S tem želimo ne le demonstrirati, da takšna avtomatska obdelava lahko zaradi svoje natančnosti in hitrosti prinaša velike ekonomske koristi, ampak tudi dokazati, da aplikacije računalniške obdelave slik nudijo nove priložnosti za uveljavitev na trgu visokih tehnologij. Seveda ne smemo pozabiti na možnost diskusije ter predstavitev konkretnih problemov in potreb, ki se porajajo pri uporabnikih, s katerimi bomo računalniško obdelavo slik in njeno koristnost še bolj približali avditoriju.

dr. Božidar Potočnik
predsednik konference
ROSUS 2013



POKROVITELJI



ROSUS 2013
<http://cobsis.uni-mb.si/rosus2013/>

KAZALO

VABLJENA PREDAVANJA.....	1
Mihaela Triglav Čekada, GI, Ljubljana <i>Uporaba merskih in nemerskih fotografij v fotogrametrične namene.....</i>	<i>3</i>
INDUSTRIJSKE APLIKACIJE	19
Martin Šavc, Božidar Potočnik, Damjan Zazula, FERI, Maribor <i>Ocenitev spektra osvetlitve z uporabo linearnih modelov logaritemskih vrednosti hiperspektralnih slik.....</i>	<i>21</i>
Lovro Kuščer, Janez Diaci, FS, Ljubljana <i>Računalniški vid pri karakterizaciji gibanja oddaljenih objektov</i>	<i>32</i>
Peter Lepej, Peter Berk, Jurij Rakun, Miran Lakota, FERI, FKBV, Maribor <i>Prednosti in slabosti digitalne obdelave slik na poljskem robotu</i>	<i>39</i>
MEDICINSKE IN BIOMEDICINSKE APLIKACIJE.....	49
Simon Zelič, Matjaž Divjak, Aleš Holobar, FERI, Maribor <i>Automatic calibration of video-based gaze tracking during robot-assisted gait rehabilitation</i>	<i>51</i>
Jurij Jemec, Miran Bürmen, Boštjan Likar, Franjo Pernuš, FE, Ljubljana <i>Modeli širjenja svetlobe v optično gostih medijih in njihova uporaba pri določanju optičnih lastnosti</i>	<i>58</i>
Žiga Lesjak, Žiga Špiclin, Boštjan Likar, Franjo Pernuš, FE, Ljubljana <i>Avtomatska zaznava časovno pogojenih morfoloških sprememb v MR slikah glave</i>	<i>65</i>
Robert Korez, Boštjan Likar, Franjo Pernuš, Tomaž Vrtovec, FE, Ljubljana <i>Razgradnja medicinskih slik hrbtenice</i>	<i>72</i>

DRUGO	81
Domen Tabernik, Luka Čehovin, Matej Kristan, Marko Boben, Aleš Leonardis, FRI, Ljubljana	
<i>VICOS EYE – Spletna storitev za kategorizacijo vizualnih objektov.....</i>	<i>83</i>
Žiga Stopinšek, Gregor Berginc, Miran Erič, Franc Solina, FRI, XLAB, ZVKDS, Ljubljana	
<i>Uvajanje 3D tehnologij pri varstvu kulturne dediščine.....</i>	<i>89</i>
ŠTUDENTSKE APLIKACIJE	99
Jaka Pohar, Jernej Bule, Peter Peer, FRI, Ljubljana	
<i>Ujemanje prstnih odtisov na podlagi grebenov.....</i>	<i>101</i>
Uroš Mlakar, Božidar Potočnik, FERI, Maribor	
<i>Razpoznavanje človeških emocij na osnovi digitalnih posnetkov.....</i>	<i>110</i>
Janez Cesar, Borut Batagelj, FRI, Ljubljana	
<i>Določanje centra očesa s pomočjo krivulj enake intenzitete</i>	<i>122</i>
Jernej Kranjec, Božidar Potočnik, FERI, Maribor	
<i>Razvijanje 360° panoramske slike iz leče s sferičnim zrcalom.....</i>	<i>129</i>

VABLJENA PREDAVANJA

UPORABA MERSKIH IN NEMERSKIH FOTOGRAFIJ V FOTOGRAMETRIČNE NAMENE

dr. Mihaela Triglav Čekada

Geodetski inštitut Slovenije
E-pošta: mihaela.triglav@gis.si
URL: <http://www.gis.si>

POVZETEK: *Na Geodetskem inštitutu Slovenije imamo dolgoletno fotogrametrično tradicijo že od ustanovitve leta 1953 dalje. Fotogrametrija omogoča izmero 3D-podatkov o prostoru na osnovi različnih fotografij, prvenstveno so se uporabljale merske fotografije velikega formata. Danes štejemo fotogrametrijo med ostale vede daljinskega zaznavanja, kamor sodijo še satelitsko zaznavanje in lasersko skeniranje terena. Kombinacija laserskega skeniranja in fotogrametričnih metod nam omogoča tudi uporabo nestandardnih nemerskih fotografij v merske namene. Fotogrametrija nam omogoča poleg sprotnega izvedenja sprememb v prostoru tudi pogled v preteklost. Najprej bomo predstavili nekaj klasičnih fotogrametričnih uporab: topografske karte velikih meril, kulturna dediščina in naravne nesreče. Omenili bomo še dva primera uporabe laserskega skeniranja: podrobne 3D-modele mest in preučevanje višine snežne odeje. Na koncu pa bomo predstavil še nekaj zanimivih nekonvencionalnih fotogrametričnih primerov, od analize 3D-obsega Triglavskega ledenika skozi čas do obdelave arhivskih merskih posnetkov za potrebe preučevanja 3D-sprememb v prostoru.*

1. UVOD

Fotogrametrija je znanost in umetnost kako določiti lokacijo in obliko objekta iz fotografij. Osnovne principe določitve objektov v postoru iz dveh fotografij preko preseka žarkov je podal že polkovnik Laussedat leta 1859. Pulfrich je v letu 1901 predstavil principe stereofotogrametrije, ki so omogočili izdelavo prvega stereo-avtografa von Orelu, s katerim so lahko neprekinjeno risali plastnice in lokacij objektov v ravnini iz fotografij narejenih s tal t.i. terestričnih fotografij. Bauersfeld je v letu 1923 v skonstruiral stereoplanigraf za fotografije narejene iz zraka t.i. aerofotografije, ki je uporabljal iste principe stereofotogrametrije kot prej omenjenji stereo-avtograf [1].

Fotogrametrija je bila dolgo časa edina veda, ki je lahko hitro podala 3D-podatke o prostoru in omogočila izdelavo podrobnih topografskih kart, zato jo štejemo kot eno izmed osnovnih merskih metod geodezije [2]. Z razvojem stroke in lažjo dostopnostjo računalnikov, pa se je fotogrametrija razvila tudi na druga netopografska področja, kot so kulturna dediščina, industrijske meritve, medicina, biologija in drugo [2, 3]. Z pospešenim razvojem satelitskih sprejemnikov v zadnjih 20ih letih se je fotogrametrija nekako zlila med ostale vede daljinskega zaznavanja, saj so poleg klasičnega

aerofotografiranja prišli v ospredje različni visokoločljivostni satelitski senzori - z velikostjo slikovnega elementa 0,5 m - katerih rezultati so čisto primerljivi z rezultati aerofotogrametrije. Kot konkurenco fotogrametriji moramo omeniti še velik razvoj laserskega skeniranja v zadnjih desetih letih, ki lahko poda podobne ali celo natančnejše rezultate fotoklasične fotogrametrične metode. Lasersko skeniranje lahko deloma štejejo tudi med napredne fotogrametrične metode, saj v veliki večini primerov zaradi lažje interpretacije uporabljamo laserske podatke v kombinaciji z ortofotografijami.

V Sloveniji je bil pobudnik za uvedbo fotogrametrije v visokošolski študij takoj po drugi svetovni vojni prof. Čuček. Prof. Čuček je bil tudi ustanovitelj in dolgoletni direktor Inštituta za geodezijo in fotogrametrijo, ki je bil ustanovljen leta 1953. V prvih letih delovanja Inštituta se je le-ta predvsem ukvarjal z izdelavo fotogrametričnih instrumentov in pripomočkov od fototeodolita do fotoperisovalnika [4]. Inštitut za geodezijo in fotogrametrijo se je na osnovi Zakona o geodetski dejavnosti (U.l. 8/2000) ter sklepa Vlade Republike Slovenije (U.l. 84/2000) v letu 2000 statusno preoblikoval v Geodetski inštitut Slovenije [5]. Na Geodetskem inštitutu Slovenije smo še vedno zvesti dolgoletni fotogrametrični tradiciji in opravljamo vrsto klasičnih fotogrametričnih del ter tudi uvajamo nove metode, ki omogočajo uporabo nemerskih fotografij v merske namene. Poleg fotogrametrije pa na Geodetskem inštitutu Slovenije pokrivamo še ostala geodetska področja od osnovnega geodetskega sistema, kartografije do nepremičnin.

V pričujočem prispevku bomo predstavili različne primere uporabe fotogrametrije, ki jih izvajamo na Geodetskem inštitutu Slovenije.

2. KLASIČNA FOTOGRAMETRIJA

Fotogrametrijo delimo glede na postavitev stojišča fotoaparata na aerofotogrametrijo in bližnjelikovno oz. terestrično fotogrametrijo. V aerofotogrametriji uporabljamo posnetke narejene iz zraka predvsem za izdelavo različnih obsežnejših topografskih prikazov. Terestrična fotogrametrija pa se ukvarja z manjšimi območji oz. podrobnostmi na objektih, zato so tudi področja njene aplikacije lahko zelo široka.

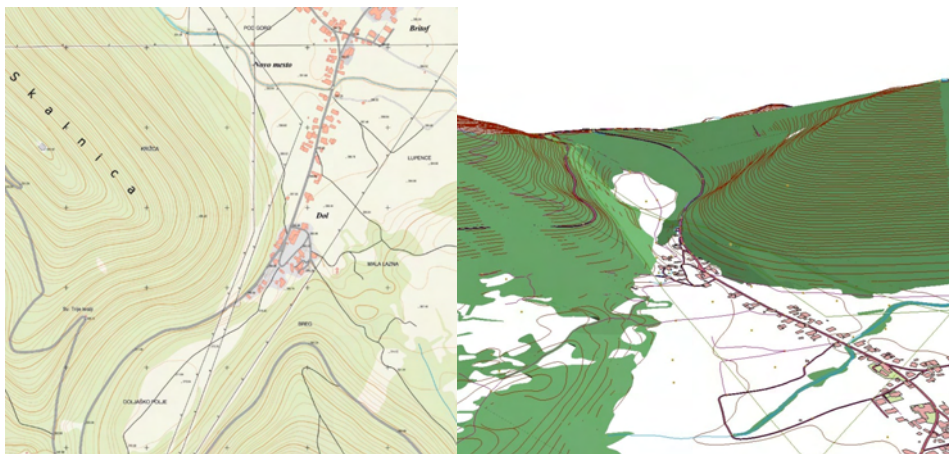
V fotogrametriji večinoma merimo 3D-podatke na osnovi stereoskopskega gledanja dveh fotografij, čemur pravimo stereoizmera. Takima fotografijama pravimo stereopar. Fotografiji stereopara prikazujeta isti objekt, imata vzporedni osi snemanja in zamaknjen projekcijski center.

Najbolj splošno poznani in razširjeni fotogrametrični izdelek je ortofotografija, ki nam poda 2D-podatke o objektih v prostoru. Za izdelavo ortofotografije potrebujemo digitalni model višin in fotografijo. Digitalni model višin (DMV) pa izdelamo s pomočjo stereoizmere, v novejšem času tudi s pomočjo podatkov laserskega skeniranja.

V nadaljevanju bomo predstavili nekaj tipičnih fotogrametričnih aplikacij, kjer fotogrametrični stereopari služijo kot osnova za izmero 3D-podatkov o prostoru.

2.1 Topografske karte velikih meril

Geodetski inštitut Slovenije je vpet v izdelavo različnih topografskih kart na osnovi stereoizmere že od njegovega nastanka dalje. Sodelovali smo pri razvoju novih topografskih izdelkov od začetkov na Temeljnih topografskih kartah v merilih 1:5000 do 1:10 000 (TTN-5 in TTN-10) v 1970-ih letih do digitalnih državnih topografskih kart meril 1:5000 (DTK5) in topografske baze konec 1990-ih let [4]. Topografske karte velikih meril glede na ostale topografske karte vsebujejo veliko večje število višinskih podatkov, od plastnic pa do posameznih kot, ki prikazujejo nadmorske višine sprememb v reliefu. Pred nekaj več kot desetimi leti je končni izdelek DTK5 prešel iz analogne tiskane karte v digitalno obliko. To je pri izdelavi topografskih kart velikih meril odprlo nove možnosti zapisov vsebine in prikazov. Tako naše digitalne DTK5 lahko opišemo kot 2,5D karte, saj imajo vsi objekti poleg lokacije zapisano v atributu tudi nadmorsko višino. Nekateri objekti, kot so stavbe, imajo v atributu zapisani tudi več nadmorskih višin: višina slemena, kapi, stika s terenom. Takšne topografske karte lahko pokažemo v 2D kot klasično karto ali v 3D kot 3D-prikaz vsebine (slika 1).



Slika 1: Prikaz izseka nadgrajene DTK5: a) klasičen kartografski 2D-prikaz, b) 3D-prikaz.

Na Geodetskem inštitutu Slovenije že od vzpostavitve sistema DTK5 sodelujemo pri kontroli izdelave DTK5, ter občasno sodelujemo tudi pri njeni izdelavi. Za potrebe nekaterih občin pa izdelujemo tudi nadgrajeno različico DTK5, ki je v primerjavi z državno predpisano vsebino DTK5 vsebuje še več dodatnih vsebin.

Za izdelavo DTK5 izvajamo stereoizmero na osnovi Cikličnega aerofotografiranja Slovenije (CAS) ali pa posebnih aerofotografiranj posameznih občin. V Sloveniji se lahko pohvalimo z dolgoletnim projektom CAS, ki ga izvaja Geodetska uprava Republike Slovenija (GU RS) že od leta 1975 naprej [6]. CAS omogoča pokritje Slovenije v 3-4 letnem ciklu na merskih aerofotografijah velikega formata. Same fotografije so izdelane v merilu, ki omogoča izdelavo topografskih prikazov v merilu 1:5000. V letu 2006 je CAS v celoti prešel iz analognih aerofotografij velikega formata

na digitalne aerofotografije, ki omogočajo popolnoma digitalno fotogrametrično linijo na digitalnih fotogrametričnih postajah, brez vmesnega skeniranja analognih aeroposnetkov.

V zadnjih letih pa smo posamezne liste DTK5 na Geodetskem inštitutu Slovenije izdelali tudi že na osnovi podatkov zračnega laserskega skeniranja v kombinaciji z ortofotografijami, ki so služile za lažjo interpretacijo.

2.2 Kulturna dediščina

Dokumentacija stanja objektov kulturne dediščine je največji uporabnik bližnjelikovne fotogrametrije. Sama fotogrametrična izmera sledi istemu principu stereoizmere, kot ga uporabljamo pri aerofotogrametriji, le da so osi slikovnega koordinatnega sistema postavljene drugače. Pri uporabi fotogrametrije v namene dokumentiranja kulturne dediščine, predvsem skrajšamo trajanje potrebnih terenskih del, saj s klasičnimi geodetskimi metodami (tahimetrija, GNSS-izmera) pomerimo samo oslonilne točke, vse ostale podrobnosti na objektu pa fotogrametrično izmerimo v pisarni. Fotografije same imajo v tem primeru še zelo pomemben dokumentacijski vidik, saj lahko določene podrobnosti na njih izmerimo šele veliko kasneje, ko se bo mogoče zato pojavila potreba (npr. porušitev objekta, restavracija detajlov, ki so odpadli...).

Med najbolj pomembne fotogrametrične rezultate dokumentacije kulturne dediščine štejemo:

- načrte v zelo velikih merilih nekje od 1:100 do 1:10,
- ortofotografije, ki zaradi projekcije omogočajo izmero detajlov, ki so prikazani na fotografijah
- merski 3D-modeli in animacije
- mapa fotografij preventivnega snemanja z orientacijskimi podatki (koordinate oslonilnih točk)

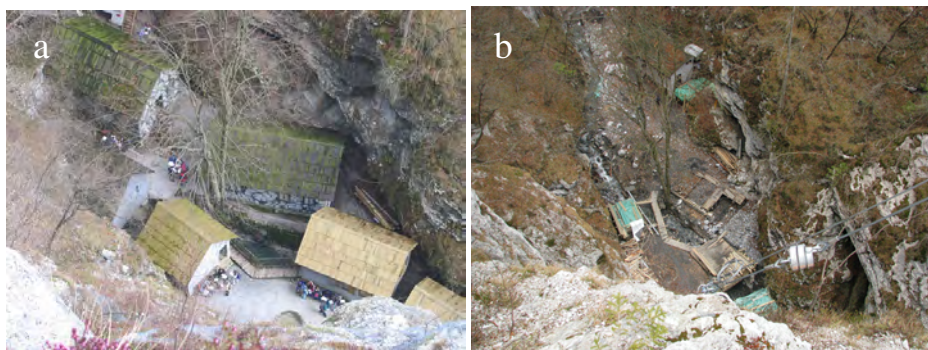
Posamezen objekt kulturne dediščine je lahko predstavljen samo z enim ali z več tipi dokumentacije.

Kot enega izmed prvih primerov s tega področja, ki je bil izveden še na takratnem Inštitutu za geodezijo in fotogrametrijo, lahko omenimo fotogrametrične meritve **Ljubljanskega gradu**, ki so se izvajale že leta 1969 [4]. Od takrat naprej je Geodetski inštitut Slovenije sodeloval pri izdelavi različne fotogrametrične dokumentacije za različne objekte kulturne dediščine. Intenzivno smo bili vključeni v projekt 'Izmere' objektov in območij kulturne dediščine, ki ga je financiralo Ministrstvo za kulturo Republike Slovenije in se je izvajal od leta 1993 naprej [9]. Omenimo samo nekatere bolj znane objekte kulturne dediščine za katere smo pripravili različne tipe dokumentacije na osnovi fotogrametrične izmere: Prešernova rojstna hiša v Vrbi na Gorenjskem, Blejski otok z okolico, grad Socerb, grad Viltuš, grad Gradac, grad Snežnik (slika 2), Stolnica v Ljubljani (cerkev Sv. Nikolaja), Krstilnica sv. Janeza Krstnika v Kopru, območje Auerspergove železarne, partizanska bolnišnica Franja... [7, 8, 9].



Slika 2: 3D-prikaz gradu Snežnik.

Prav preventivna fotogrametrična dokumentacija **Bolnice Franje**, izdelana v okviru projekta 'Izmere' v letu 2002, je omogočila njeno natančno rekonstrukcijo po poplavih v septembru 2007, ki so uničile skoraj vse originalne barake (slika 3).



Slika 3: Bolnica Franja: a) pred nesrečo v letu 2002, b) po nesreči v 2008.

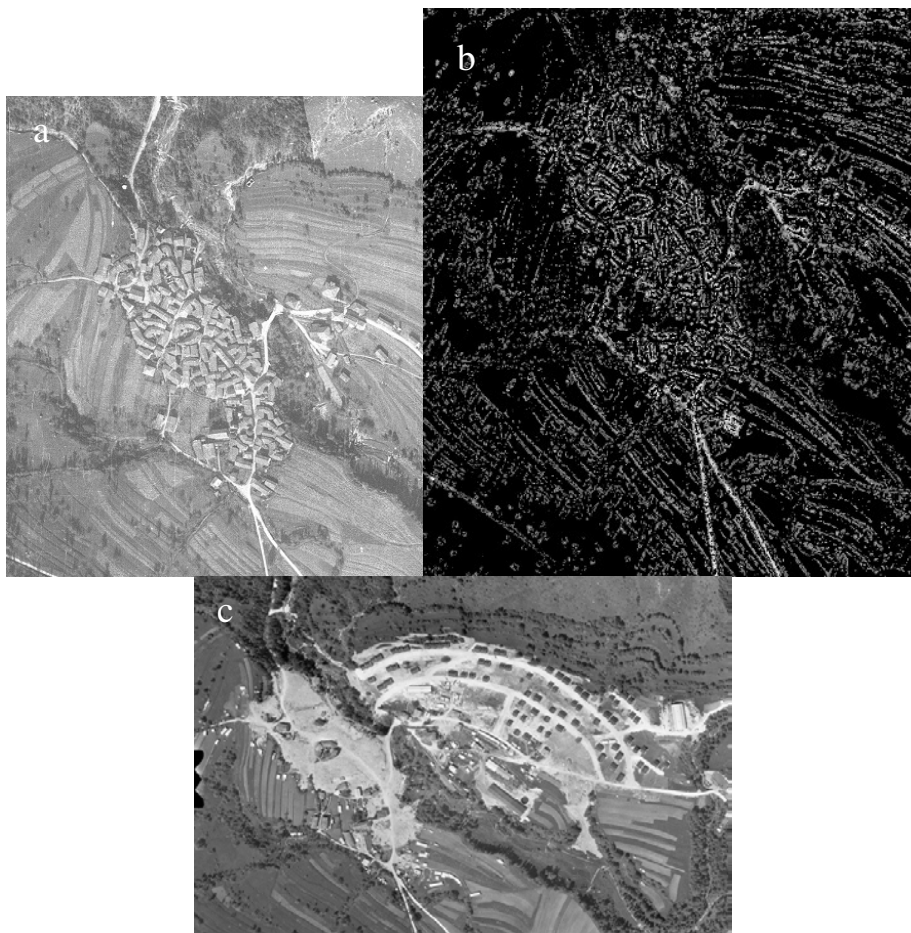
Danes v primeru dokumentacije kulturne dediščine klasično bližnjleslikovno fotogrametrijo vse bolj izpodriva terestrično lasersko skeniranje. Vendar moramo poudariti, da moderni terestrični laserski skenerji vsebujejo poleg laserskega merilnika še merski fotoaparati. Zato dobimo kot rezultat terestričnega laserskega skeniranja, poleg oblaka laserskih točk, še merske fotografije umeščene v prostor.

2.4 Naravne nesreče

Fotogrametrija se uporablja tudi za interventno pridobivanje podatkov o posledicah naravnih nesreč, saj je v veliko primerih dostop na prizadeta območja otežen in stanje lahko popišemo le iz zraka. Stanje v prostoru pred nesrečo lahko hitro izrednotimo iz že omenjenih aeroposnetkov CAS. Stanje po nesreči pa se v večini primerov zelo hitro spreminja, zato moramo samo naročilo in izvedbo interventnega snemanja izvesti zelo hitro. Na primeru poplav vidimo, da že nekaj dni po sami poplavi na fotografijah zelo težko določimo do kam je segala voda. Pri organizaciji interventnih snemanj so tako pomembni dnevi.

Geodetski inštitut Slovenije je bil udeležen pri izvedbi in obdelavi fotogrametričnih posnetkov posledic **drobirskega toka v Logu pod Mangartom v letu 2000**. Že dva dni po nesreči se je izvedno interventno helikoptersko fotografiranje z merskim fotoaparatom srednjega formata. Sam fotograf se je v času snemanja nagibal čez vrata helikopterja. Z obdelavo obstoječih posnetkov CAS obravnavanega območja in novih fotogrametričnih posnetkov, smo zelo hitro izdelali prve načrte in prečne prereze nanosov, izračunali prostornino prestavljenega materiala in izdelali še animacijo obstoječega stanja. Sama obdelava posnetkov je bila takrat izvedena še na analitičnem fotogrametričnem inštrumentu v katerega vstavimo originalne negative fotografskega filma. Druge inštitucije so kasneje izvedle še podrobne analize dogodka na osnovi posebnega aerofotografiranja, ki je bilo izvedeno nekaj dni po našem helikopterskem snemanju ter z obdelavo satelitskih slik [10].

Tekoče fotografije CAS in arhivski fotografije CAS ter fotografije posebnih snemanj predstavljajo neizmiren vir 3D-podatkov o spremembah v prostoru. Stare analogne fotografije se skladiščijo v arhivu, pred njihovo uporabo v digitalnih fotogrametričnih postajah pa jih je potrebno skenirati s fotogrametrično ločljivostjo na posebnem fotogrametričnem skenerju. V sklopu arhiva Geodetskega inštituta Slovenije posedujemo tudi tak skener in skeniramo arhivske fotografije CAS. Kot primer uporabe arhivskih fotografij za potrebe preučevanja posledic naravnih nesreč na objektih kulturne dediščine si lahko ogledamo možnost digitalne rekonstrukcije **vasi Breginj, ki so jo porušili kmalu po katastrofalnem potresu v 1976**. Na Geodetskem inštitutu Slovenije smo na osnovi arhivskih aerofotografij izdelali digitalni model površja (DMP) po vrhovih stavb vasi Breginj pred potresom (slika 4).



Slika 4: Vas Breginj pred potresom na izseku aerofotografije iz leta 1956, a) DMP po vrhovih stavb pred potresom in b) izsek aerofotografije iz leta 1977 po poružitvi vasi.

3. LASERSKO SKENIRANJE

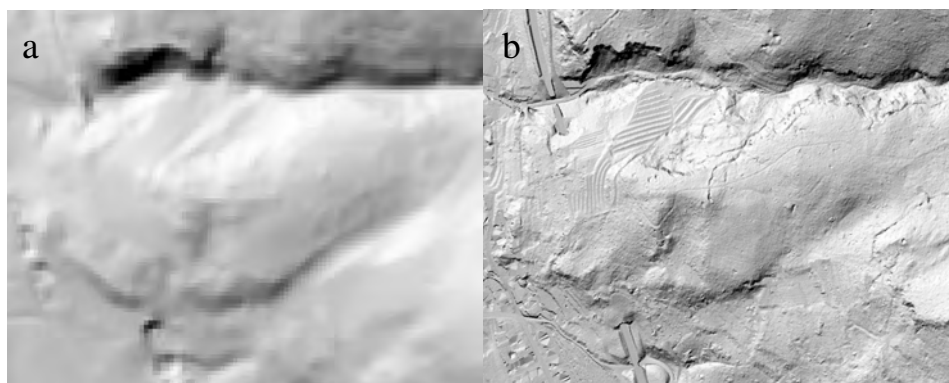
Lasersko sneniranje lahko glede na postavitev stojišča, podobno kot klasično fotogrametrijo, delimo na zračno lasersko skeniranje in terestrično lasersko skeniranje. Zračno lasersko skeniranje pa delimo še po vrstah aplikacije na topografsko in batimetrično lasersko skeniranje, ki se ločita po valovni dolžini uporabljenega laserja. Podatke topografskega zračnega laserskega skeniranja lahko uporabimo za različne prostorske ter geomorfološke študije [11]. Terestrično lasersko skeniranje pa po vrsti aplikacije in tipu inštrumenta delimo na tri razrede inštrumentov [12]:

- kratkega dosega do 25 m, ki so primerni za zajem manjših objektov kot so posamezni arhološki ostanki in industrijske aplikacije
- srednjega dosega do 250 m, ki so primerni za zajem gradbenih konstrukcij, objektov kulturne dediščine in arheoloških najdišč
- dolgega dosega nad 250 m, ki so primeri za zajem obsežnejših objektov, manjših plazov

Medtem, ko je najširše uporabljeni fotogrametrični izdelek ortofotografija, lahko pri laserskem skeniranju rečemo, da so to različni digitalni modeli višin:

- digitalni model reliefa (DMR), ki prikazuje tla brez vegetacije in objektov
- različni digitalni modeli površja (DMP), ki prikazujejo višine npr. vegetacije, stavb

Digitalni modeli višin pridobljeni iz topografskih laserskih podatkov večinoma po ločljivosti močno prekašajo take izdelane na osnovi fotogrametrije, npr. iz stereoparov CAS (slika 5).



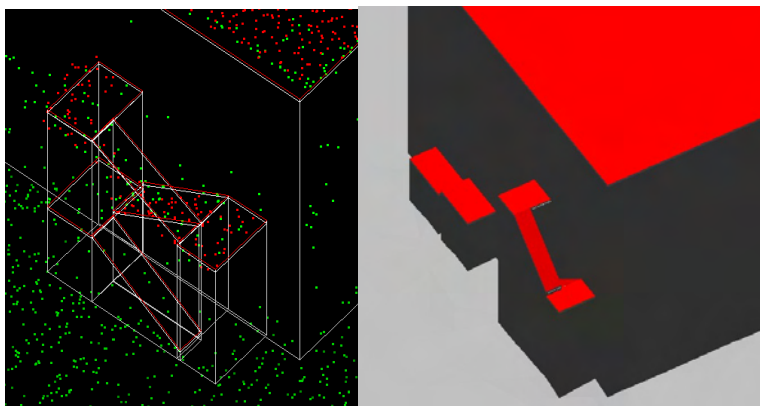
Slika 5: Senčna reliefa fotogrametričnega DMR 5 m x 5 m CAS2006 a) in lidarskega DMR 0,1 m x 0,1 m 2009 istega območja b).

Ker bodo podatki zračnega laserskega skeniranja postali skozi državna snemanja slejkoprej širše dostopni različnim raziskovalcem in drugim inštitucijam v Sloveniji [13], se bomo v nadaljevanju osredotočili na dve zanimivi aplikaciji, ki smo ju izvedli na Geodetskem inštitutu Slovenije.

3.1 Izdelava podrobnih 3D-modelov mest

V okviru mednarodnega čezmejnega projekta INTERREG IIIA: HARMO-GEO smo med drugim v letih 2000-2006 preučevali tudi primernost zelo gostih laserskih podatkov za izdelavo zelo podrobnih 3D-modelov mest. Uporabljeni laserski podatki z gostoto 20 točk/m² omogočajo podrobno izmero nadstreškov, dimnikov, strešnih oken, balkonov, napuščev ter celo klimatskih naprav [14]. Na osnovi laserskega skeniranja izdelan 3D-

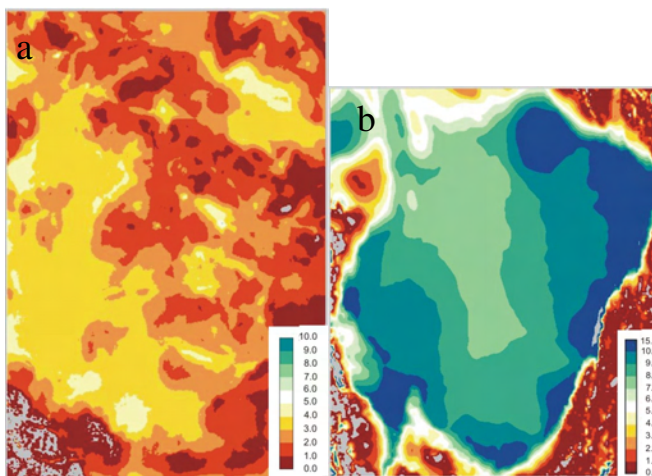
model mesta **Nove Gorice** je bil na koncu oblepljen še z terestričnimi fotografijami, s čimer smo dobili zelo realističen in podroben prikaz mesta. Sam 3D-model je merski, kar pomeni, da bi v njem lahko tudi merili razdalje med objekti.



Slika 6: Detajli stopnišča v: a) oblak laserskih točk z žičnim 3D-modelom in b) ploskovni 3D-modelu istega detajla.

3.2 Analiza višine snežne odeje

Podrobne digitalne modele višin izdelane na osnovi podatkov zračnega skeniranja izvedenih v več časovnih presekih lahko uporabljamo za analizo premikov zemljine od antropogenih zasipanj kraških kotanj do preučevanja premikanj zemeljskih plazov. V okviru čezmenjnega projekta SI-AT: NH-WF smo v letu 2012 izvedli prvo analizo višine snežne odeje na treh tesnih območjih. Prvo zasneženo lasersko skeniranje smo izvedli v maju, drugo brez snega pa v septembru. Z primerjavo DMP, ki prikazuje višino snežne odeje v maju, in DMR izdelanim na osnovi septemberskega snemanja smo določili lokalne razlike v višini snežne odeje. Oba DMV-ja imata velikostjo celice $1\text{ m} \times 1\text{ m}$ in sta bila izdelana na osnovi laserskih podatkov z povprečno gostoto 8 točk/m^2 . Na sliki 7 sta prikazani najbolj zanimivi področji Triglavskega ledenika in ledenika pod Skuto. Vidimo lokane spremembe v višini snežne odeje, ki so pomembne za razlago obstanka obeh slovenskih ledenikov. Takšni podatki o višini snežne odeje nam lahko služijo za ugotavljanje območjih z napihanim snegom (zameti), za modeliranje nevarnosti pred snežnimi plazovi ter za preučevanje vodnega ekvivalenta snega. Za razliko od meteoroloških meritev višine snega, ki so točkovne narave, je taka primerjava ploskovna in veliko pove o lokalnih topografskih trendih, ki botrujejo različnim višinam snega [15].



Slika 7: Višina snežne odeje v maju 2012 na Triglavskem ledeniku a) in na ledeniku pod Skuto b).

4. UPORABA NEMERSKIH FOTOGRAFIJ V MERSKE NAMENE

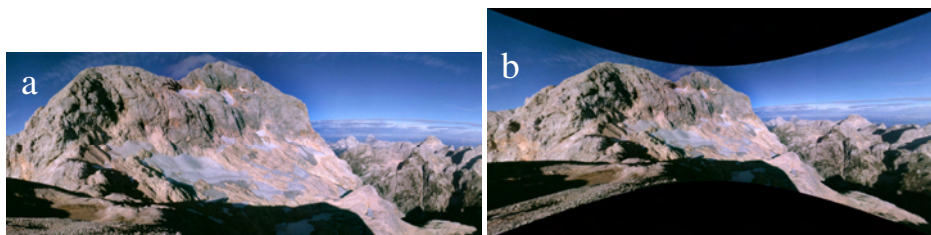
Rezultate prej omenjenih fotogrametričnih metod in laserskega skeniranja, lahko uporabimo tudi za obdelavo nemerskih fotografij. Kot najbolj uporaben se je izkazal podroben digitalni model višin, ki ga lahko uporabimo pri obdelavi različnih nemerskih fotografij od preučevanja zmanjševanja Triglavskega ledenika do analize poplav [16, 17, 18, 19].

4.3 Triglavski ledenik

Z fotogrametrično in geodetsko izmero Triglavskega ledenika smo pričeli sodelovati z Geografskim inštitutom Antona Melika ZRC SAZU v letu 1999. Od takrat naprej smo najprej na dve leti, po letu 2007 pa celo vsako leto meril Triglavski ledenik s kombinacijo različnih fotogrametričnih in klasičnih geodetskih metod izmere. Najprej smo na ledeniku merili oslonilne točke za posebna helikopterska fotogrametrična snemanja ledenika. V letu 2005 smo izvedli klasično aerofotografiranje ledenika in širšega območja Triglavskih podov z fotogrametrično kamero velikega formata. Od takrat naprej smo izvajali predvsem terestrična fotogrametrična snemanja samega ledenika v kombinaciji s tahimetrično izmero njegovega oboda [20]. Nazadnje smo v letu 2012 izvedli še dvakratno zračno lasersko skeniranje širšega območja okoli ledenika, ki smo ju omenili že v prejšnjem poglavju.

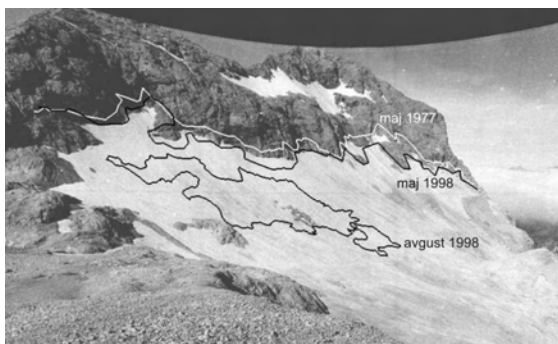
Sama fotogrametrična in geodetska izmera Triglavskega ledenika je namenjena izdelavi načrtov ledenika v merilu 1:1000 in preučevanju zmanjševanja površine in prostornine ledenika. Dolgoročno preučevanje ledenikov nam veliko pove o samih lokalnih in globalnih spremembah klime.

S pomočjo pridobljenih 3D-podatkov o ledeniku, predvsem s pomočjo podrobnega DMR z velikostjo celice 2 m×2 m izdelanega na osnovi aerofotogrametričnega snemanja v letu 2005, pa smo uspeli obdelati tudi starejše arhivske fotografije ledenika.



Slika 8: Horizontova fotografija Triglavskega ledenika posneta konec talilne dobe v letu 2010: a) original b) razpačena v centralno projekcijo.

Podrobno smo preučili panoramske fotografije narejene s fotoaparatom Horizont. Ta fotoaparat uporabljajo za vsakomesečno fotografiranje ledenika z dveh stalnih stojišč že od leta 1976 naprej. Fotoaparat ni merski, poleg tega pa ima zaradi panoramske projekcije zelo velike distorzije. Fotoaparat smo najprej kalibrirali in s tem določili parametre distorzije fotografije. Z odstranitvijo distorzij smo fotografije pretvorili v centralno projekcijo, potem pa smo uporabili DMR za preučevanje spreminjanja oboda ledenika skozi čas (slika 8). Osnovna predpostavka, ki nam je omogočila uporabo moderenega DMR za preučevanje starih stanj ledenika je, da se DMR na tem območju v 30 letih ni bistveno spreminjal in da se zato stik terena in ledenika vedno lahko opiše na modernem DMR. To nam je omogočilo, da smo na osnovi modernega DMR določili obseg ledenika v 3D za njegova stanja v preteklosti, ki so dokumentirana na Horizontovih fotografijah. 3D-obod ledenika smo uporabili za izračun površine in teroretične prostornine ledenika. Teoretično prostornino dobimo kot razliko med ploskvijo, ki povezuje 3D-obod ledenika in DMR iz leta 2005 [16]. Samo zgornji 3D-rob ledenika pa smo uporabili za določitev višine snežne odeje v posameznem mesecu. Z pridobitvijo še podrobnejšega DMR narejenega na osnovi laserskega skeniranja v letu 2012, bomo lahko analizo ponovili na še bolj natančnih podatkih.



Slika 9: Višina zgornjega roba ledenika v maju 1997 in maju 1998 prikazana na stanju ledenika 26. 8. 1977. Sklenjena črta prikazuje še celoten obod ledenika v avgustu 1998.

3.3 Naravne nesreče – poplave 2010 in 2012

Kot smo že prej omenili, ocenjujemo obsege poplav večinoma na izdelkih interventnih snemanj. Odvisno od obsežnosti poplave njen obseg najhitreje in najlažje določimo na ortofotografijah narejenih na osnovi aerofotogrametrije ali visokoločljivostnih satelitskih posnetkov. Na Geodetskem inštitutu Slovenije pa smo pričeli razvijati metodo za obdelavo nemerskih fotografij poplav s pomočjo DMR.

Prve teste uporabnosti metode z uporabo DMR $5\text{ m} \times 5\text{ m}$, narejenim na osnovi fotogrametričnega zajema v okviru CAS, smo izvedli na primeru poplav leta 2010. Obdelali smo posnetke poplav v Dobropolju in delu Ljubljanskega Barja [17, 18]. Poplave na Ljubljanskem barju so služile tudi kot testni primer na katerem smo primerjali rezultate dobljene na osnovi fotogrametričnega DMR in DMR, narejenega na osnovi zračnega laserskega skeniranja. DMR narejen na osnovi laserskega skeniranja omogoča določanje višine poplavne vode v rangu nekaj decimetrov, kar fotogrametrični DMR ne omogoča [18].





Slika 10: Poplave 5. 11. 2012 v Šoštanju: a) rezultat izmere obsega poplav, b) ena izmed 3 fotografij uporabljenih za izmero obsega poplav v Šoštanju (foto: Dejan Tonkli).

Ob poplavah v novembru 2012 pa smo na Geodetskem inštitutu Slovenije izvedli akcijo kolaborativnega zbiranja različnih fotografij poplav, ki smo jo razširili preko spletnih medijev. Lahko rečemo, da lahko približno 50 % pridobljenih fotografij s strani javnosti takoj uporabimo za evidentiranje obsega poplav. 3D-rob poplav smo izmerili na osnovi fotogrametričnega DMR 5 m × 5 m izdelanega v okviru CAS. Podrobni rezultati so objavljeni na naši spletni strani [21].

4. ZAKLJUČEK

Skupni imenovalec našega fotogrametričnega delovanja na Geodetskem inštitutu Slovenije je kako iz različnih merskih ali celo nemerskih fotografij pridobiti čimbolj natančne merske podatke. V predstavitvi smo najprej prikazali nekaj klasičnih nalog fotogrametrije: topografske karte velikih meril in dokumentiranje objektov kulturne dediščine. Omenili smo tudi lasersko skeniranje, kot metodo, ki v nekaterih aplikacijah počasi nadomešča klasično fotogrametrijo. Potem pa smo se predvsem ustavili na primerih, kjer smo kombinirali različne merske podatke in nemerske fotografije ter tako pridobili merske podatke: analiza arhivskih posnetkov Triglavskega ledenika in preučevanje poplav 2010, 2012.

LITERATURA

1. K. Kraus (1993) Photogrammetry. Volume 1: Fundamentals and Standard Processes, Dümmler/Bonn.
2. T. Luhmann, S. Kyle, I. Harley (2006) Close Range Photogrammetry, Whittles Publishing.
3. J. Fryer, H. Mitchell, J. Chandler (2007), Applications of 3D Measurements from Images, Whittles Publishing.

4. D. Radovan (1998), Zbornik ob 45-letnici Inštituta za geodezijo in fotogrametrijo FGG, Inštitut za geodezijo in fotogrametrijo FGG, Ljubljana.
5. <http://www.gis.si/sl/o-nas/zgodovina>
Zgodovina Geodetskega inštituta Slovenije
6. K. Oštir (2006), Daljinsko zaznavanje, Založba ZRC, Ljubljana.
7. D. Mravlje (1981), Prostorska izmera spomeniških zgradb, *Geodetski vestnik*, 3, str. 150-154.
8. K. Oven (2001), Rotunda Carmine – kulturna dediščina in fotogrametrija, *Geodetski vestnik*, 45/1&2, str. 72-82.
9. Grobovšek, J. (2002), Preventivno fotogrametrično snemanje gradu Snežnik v okviru nacionalnega projekta 'Izmere', *Geodetski vestnik*, 46/4, str. 379-390.
10. M. Triglav Čekada, D. Radovan (2010), Interventna in preventivna snemanja pri naravnih nesrečah, *GIS v Sloveniji 2009-2011*, str. 189-196.
11. M. Triglav Čekada (2011), Možnosti uporabe zračnega laserskega skeniranja (LIDAR) za geomorfološke študije, *Geografski vestnik* 83/2, str. 81-93.
12. V. Bric, M. Triglav Čekada, M. Bitenc (2012), Uporaba laserskega skeniranja pri zaščiti in reševanju ter vojaških aktivnostih, *Geoprostorska podpora obrambnemu sistemu Republike Slovenije*, str. 181-204.
13. M. Triglav Čekada, V. Bric, K. Oven (2012), Prvo vsedržavno lasersko skeniranje Slovenije, *GIS v Sloveniji 2011-2012*, str. 191-196.
14. B. Barborič, M. Triglav Čekada, S. Berk, M. Kosmatin Fras (2008), Metodologija za izboljšavo in harmonizacijo geodetskih podlag za potrebe čezmejnega prostorskega načrtovanja, *GIS v Sloveniji 2007-2008*, str. 273-286.
15. M. Triglav Čekada, V. Bric, M. Klanjšček, B. Barborič, M. Pavšek (2013), Zračno lasersko skeniranje in sneg, *Raziskave s področja geodezije in geofizike 2012 : zbornik predavanj*, str. 85-93.
16. M. Triglav Čekada, D. Radovan, M. Gabrovec, M. Kosmatin Fras (2010), Preučevanje Triglavskega ledenika s pomočjo arhivskih Horizontovih fotografij. *GIS v Sloveniji 2009-2010*, str. 147-153.

17. M. Triglav Čekada, M. Zorn (2012), Uporabnost nemerskih fotografij za preučevanje poplav - primer poplav na Dobropolju septembra 2010. *GIS v Sloveniji 2011-2012*, str. 55-62.
18. M. Triglav Čekada, M. Zorn (2012), Poplave septembra 2010 - obdelava nemerskih fotografij s fotogrametričnim DMR in lidarskimi podatki, *Geodetski vestnik*, str. 802-813.
19. M. Triglav Čekada, M. Klanjšček, M. Zorn (2013), Preučevanje poplav novembra 2012 na podlagi nemerskih posnetkov z vključevanjem javnosti. *Raziskave s področja geodezije in geofizike 2012 : zbornik predavanj*, str. 75-83.
20. Triglav Čekada, M., Zorn, M., Kaufmann, V, Lieb, K.G. (2012), Merjenje malih Alpskih ledenikov: primeri iz Slovenije in Avstrije, *Geodetski vestnik*, 56/3, str. 443-461.
21. <http://www.gis.si/sl/poplave-2012>
Poplave 2012 - zbiranje fotografskega gradiva za potrebe raziskav

INDUSTRIJSKE APLIKACIJE

OCENITEV SPEKTRA OSVETLITVE Z UPORABO LINEARNIH MODELOV LOGARITEMSKIH VREDNOSTI HIPERSPEKTRALNIH SLIK

Martin Šavc, Božidar Potočnik, Damjan Zazula

Fakulteta za elektrotehniko, računalništvo in informatiko

E-pošta: martin.savc@uni-mb.si

POVZETEK: Meritve barve v sliki so močno odvisne od spektra svetlobe, s katero je opazovana površina osvetljena. Pri primerjanju z meritvami iste površine zajete v drugačnih osvetlitvenih pogojih, je potrebno upoštevati ali odstraniti lastnosti različnih spektrov svetlobe. Informacija o osvetlitvi običajno ni na voljo, zato le-to ocenimo z algoritmi barvne vztrajnosti. V tem delu predstavimo metodo, ki uporabi linearne modele za opis logaritmiranih meritev v hiperspektralni sliki in z njimi oceni spekter svetila. Rezultati pokažejo bistvene prednosti predstavljene metode pred drugimi preprostimi algoritmi barvne vztrajnosti.

1. UVOD

Uporaba barvnih slik v računalniškem vidu nudi številne prednosti pred sivinskimi slikami. Izmerjena vrednost barve je odvisna od reflektivnosti opazovane površine in spektra svetlobe, s katero je osvetljena. To povzroča težave, kadar primerjamo vrednosti slik iz različno osvetljenih scen.

Če poznamo lastnosti svetlobe, lahko izmerjene vrednosti pikselov preslikamo v takšne, ki bi jih izmerili pod svetlobo nevtralnega spektra. V računalniškem vidu lastnosti svetlobe ocenimo z algoritmi barvne vztrajnosti (*ang. colour constancy algorithms*). Ti algoritmi ocenjujejo barvno vrednost osvetlitve z uporabo določenih predpostavk o vsebini scene. Pogosto so uporabljene preproste predpostavke o povprečni barvni vrednosti ali povprečni vrednosti robov v sceni. Pregled različnih metod za barvno vztrajnost je na voljo v [1]. Kljub uporabljenim preprostim predpostavkam, ki pogosto celo niso izpolnjene v sceni, pa takšni algoritmi, v primerjavi z veliko kompleksnejšimi algoritmi, dosegajo dobre rezultate [2].

Pristopi k barvni vztrajnosti, opisani v [3, 4, 5], predlagajo uporabo linearnih modelov za predstavitev spektra svetila in reflektivnosti površin. S pomočjo S meritev reflektivnosti površin, ki so osvetljene z V spektri svetil, sestavimo sistem bilinearnih enačb in poiščemo analitično rešitev. Natančen pregled potrebnih in zadostnih pogojev za iskanje unikatne rešitve takšnega sistema najdemo v [4, 5]. V praksi imajo te metode številne pomanjkljivosti. Prva težava je zagotoviti ujemanje iste površine v različnih slikah pod različnimi osvetlitvami. Drug problem je izbira števila baznih funkcij, ki jih uporabimo za opis linearnega modela reflektivnosti površin in spektra svetila. V [6] so pokazali, da

potrebujemo minimalno 3 bazne funkcije za opis spektrov dnevne svetlobe (več, če uporabljamo fluorescentna ali svetila LED) ter 6 do 8 baznih funkcij za natančen opis reflektivnosti površin. Z eksperimenti smo preverili, da je kljub temu srednja kvadratična napaka med dejanskim spektrom svetila oz. reflektivnostjo površine in njuno predstavitevjo s koeficienti linearnega modela še vedno precejšnja. Pri analitičnem reševanju enačb zatorej pride do prevelikega prileganja k izmerjenim vrednostim in posledično do slabših rezultatov.

Večina obstoječih algoritmov je namenjenih uporabi na slikah s tremi barvnimi kanali. Nekateri raziskave pa uporabljajo tudi hiperspektralne slike [7, 8]. Takšne slike imajo močno povečano število barvnih kanalov. Njihov izhod je interpretiran kot meritev spektra svetlobe v vsaki točki slike. Ker preprosti algoritmi (npr. "Gray World" ali "Gray Edge" algoritem) opazujejo vsak kanal posebej, jih lahko uporabimo tudi na teh slikah. Seveda pa ti algoritmi ne bodo izkoristili vseh informacij, ki jih vsebujejo hiperspektralne slike.

V našem delu se osredotočimo na uporabo linearnih modelov za ocenitev spektra osvetlitve v hiperspektralnih slikah. Vrednosti hiperspektralne slike bomo logaritmirali ter s tem množenje reflektivnosti površine in spektra svetila v modelu spremenili v seštevanje. Logaritmirane spektre in reflektivnosti bomo nato z linearnimi modeli razbili na bazne funkcije in koeficiente. S poznavanjem baznih funkcij bomo nato za poljuben izmerjen spekter svetila lahko ocenili koeficiente spektra svetila in reflektivnosti površin, iz katerih bomo rekonstruirali reflektivnost opazovane površine in spekter svetila, ki jo osvetljuje.

Na koncu podajmo še strukturo tega prispevka. V drugem poglavju opišemo našo metodo barvne vztrajnosti na osnovi uporabe linearnih modelov v hiperspektralnih slikah. V tretjem poglavju predstavimo izvedeni eksperiment in dobljene rezultate. Delo zaključimo z diskusijo ter s kratkim sklepom.

2. LINEARNI MODELI V LOGARITEMSKEM PROSTORU

Odzivi senzorjev vsakega kanala hiperspektralne slike so ozki, kar pomeni, da združujejo le majhno okolico sosednjih valovnih dolžin v eno meritev. V modelu odziva lahko zato zanemarimo funkcijo občutljivost senzorja (tj. kamere), ob tem pa tudi ni potrebno izvesti integracije preko območja vse valovnih dolžin. Predpostavimo, da meritve posameznega kanala pripadajo valovni dolžini, na katero so senzorji tega kanala maksimalno občutljivi. Za kanal chn izmerjeno vrednost v v točki \mathbf{x} definiramo kot

$$v_{chn}(\mathbf{x}) = k(\mathbf{x}) R(\lambda_{chn}, \mathbf{x}) I(\lambda_{chn}) \quad , \quad (1)$$

kjer je λ_{chn} valovna dolžina, ki smo jo izmerili s senzorji kanala chn . Funkcija R opisuje reflektivnost površine v točki \mathbf{x} , funkcija I pa opisuje spekter svetila, s katerim je osvetljena scena. Funkcija k je množilni faktor, ki opisuje geometrijsko razmerje med svetilom in površino. Geometrijsko razmerje je odvisno od kota med vpadno svetlobo in

normalo površine v točki \mathbf{x} in je opisano z Lambertovim kosinusnim zakonom.

S faktorjem k dodatno opišemo tudi družino vseh reflektivnosti površin in spektrov svetil, ki lahko vrnejo enake meritve v_{chn} za vse kanale. V nadaljevanju podajamo dodatno razlago te ideje. Kadar imamo dve različni površini, katerih reflektivnost se razlikuje le za faktor k_i in ju osvetlimo z dvema svetiloma, katerih spektra se razlikujeta za faktor $1/k_i$, bodo izmerjeni vrednosti obeh površin enaki. Ker takšnih kombinacij ne moremo razločiti, opišemo obe reflektivnosti in oba spektra in vse njihove mnogokratnike z enakimi funkcijami R in I . Razlike v intenziteti, bodisi spektra ali reflektivnosti pa opišemo v skupnem faktorju k .

V enačbi (1) sta pomnožena spekter svetila in reflektivnost površine. Znano je, da z logaritmiranjem množenje spremenimo v seštevanje. Če ta pristop uporabimo pri našem reševanju, se s tem problem ločevanja spektra svetila in reflektivnosti površin ustrezno poenostavi. Logaritmirano enačbo (1) zapišemo v naslednji obliki

$$\log(v_{chn}) = \log(k) + \log(R(\lambda_{chn})) + \log(I(\lambda_{chn})) \quad , \quad (2)$$

kjer smo izpustili argument pozicije točke \mathbf{x} v funkcijah v_{chn} , k in R . Odvisnost od točke \mathbf{x} sliki v nadaljevanju ne igra pomembne vloge, zato jo obravnavamo kot implicitno.

Za funkcije spektra svetila in reflektivnosti površin v logaritemskem prostoru uporabimo linearna modela. Enačba za model reflektivnosti površin je

$$\log(R(\lambda_{chn})) = \sum_{n=1}^N r_n R_n^l(\lambda_{chn}) \quad , \quad (3)$$

kjer so r_n koeficienti reflektivnosti površin, R_n^l pa so splošne bazne funkcije linearnega modela za logaritmirano reflektivnost površin (nadpisan l označuje logaritemski prostor). Koeficienti so odvisni od točke \mathbf{x} , bazne funkcije pa so za vse točke enake. Število vseh baznih funkcij za reflektivnost površin je N . Podobna enačba velja za model spektra svetila, ki je podan kot

$$\log(I(\lambda_{chn})) = \sum_{m=1}^M i_m I_m^l(\lambda_{chn}) \quad , \quad (4)$$

kjer so i_m koeficienti spektra svetila, I_m^l pa so splošne bazne funkcije za logaritmiran spekter svetila. Število vseh baznih funkcij za spekter svetil je M .

Za iskanje baznih funkcij imamo na voljo različne metode multivariantne analize (npr. metodi PCA in ICA). Podrobnosti izbire baznih funkcij bodo opisane v naslednjem poglavju.

Z vstavljanjem enačb (3) in (4) v enačbo logaritemskega modela za hiperspektralne slike (2) dobimo naslednji izraz

$$\log(v_{chn}) = \log(k) + \sum_{n=1}^N r_n R_n^l(\lambda_{chn}) + \sum_{m=1}^M i_m I_m^l(\lambda_{chn}) \quad , \quad (5)$$

ki opisuje model logaritma izmerjene barvne vrednosti z uporabo linearnih modelov logaritma reflektivnosti površin in spektra svetil. Formule vseh kanalov lahko zapišemo v matrični obliki kot

$$\mathbf{v} = \mathbf{B} \mathbf{w} \quad , \quad (6)$$

$$\mathbf{v} = [\log(v_1) \quad \cdots \quad \log(v_p)]^T \quad , \quad (7)$$

$$\mathbf{B} = \begin{bmatrix} 1 & I_1^l(\lambda_1) & \cdots & I_M^l(\lambda_1) & R_1^l(\lambda_1) & \cdots & R_N^l(\lambda_1) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & I_1^l(\lambda_p) & \cdots & I_M^l(\lambda_p) & R_1^l(\lambda_p) & \cdots & R_N^l(\lambda_p) \end{bmatrix} \quad , \quad (8)$$

$$\mathbf{w} = [\log(k) \quad i_1 \quad \cdots \quad i_M \quad r_1 \quad \cdots \quad r_N]^T \quad . \quad (9)$$

Sistem linearnih enačb bo rešljiv, kadar bodo bazne funkcije linearnih modelov linearno neodvisne, skupno število koeficientov M in N pa bo manjše od števila vseh opazovanih kanalov (oz. valovnih dolžin) P . Število kanalov P je v hiperspektralni sliki običajno veliko večje od smiselnega števila baznih funkcij. Nekolinearnost baznih funkcij pa lahko preverimo pred njihovo uporabo. To nam omogoča, da koeficiente reflektivnosti površin in spektra svetil ocenimo za vsako točko posebej.

Koeficiente poiščemo z minimizacijo regularizirane napake vsote kvadratov, kar zapišemo kot

$$E_{lsq} = \sum_{p=1}^P (\mathbf{v}_p - \mathbf{B}_p \cdot \mathbf{w})^2 + \|\mathbf{w}\|^2 \quad (10)$$

kjer je \mathbf{v}_p p -ti element v vektorju \mathbf{v} , \mathbf{B}_p pa p -ta vrstica matrike \mathbf{B} . Minimizacija napake vsote kvadratov vodi v rešitev pseudo inverza. Regularizacija, uporabljena v postopku minimizacije, pa preprečuje, da bi koeficienti v vektorju \mathbf{w} postali preveliki.

Ko z minimizacijo ocenimo vrednosti koeficientov lahko z enačbama (3) in (4) rekonstruiramo originalno reflektivnost površin in spekter svetila.

3. EKSPERIMENT IN REZULTATI

Ekspersment smo izvedli v programskem okolju iPython [9] s knjižnicami za matrično računanje NumPy in SciPy [10], knjižnico za strojno učenje MLPy [11], knjižnico za obdelavo podatkov MDP [12] ter knjižnico za prikaz diagramov in slik Matplotlib [13].

3.1 Podatki in nastavitve

V eksperimentu smo uporabili izmerjene spektre izbranih svetil in reflektivnosti površin ter hiperspektralne slike naravnih scen, ki so bile zajete z meritvijo referenčne površine, pri čemer so bile meritve popravljene na nevtravno osvetlitev. Izmerjene spektre in reflektivnosti smo dobili iz podatkovne zbirke opisane v [14], hiperspektralne slike pa smo dobili iz zbirk opisanih v [15, 16].

V zbirki [14] je vsak spekter svetila in reflektivnosti površine izmerjen z vzorčenjem spektra svetlobe na območju valovnih dolžin med 380 nm in 780 nm v razmaku 4 nm. Te meritve torej vsebujejo 102 vzorca. Hiperspektralne slike iz zbirke [15, 16] pa imajo spektre vpadne svetlobe izmerjene v območju valovnih dolžin med 410 nm in 710 nm v razmaku 10 nm, kar pomeni, da smo imeli na razpolago 31 vzorcev na meritev. Razlike v vzorčenju so povzročile, da so bili podatki ene zbirke opisani v drugačnem vektorskem prostoru kot podatki druge zbirke. Zato smo med prostoroma obeh meritev najprej opravili naslednjo transformacijo. Najprej smo vsak spekter svetila oz. reflektivnosti površin linearno interpolirali na območje 380 nm do 780 nm v razmaku 1 nm. Nato smo interpolirane meritve zgladili z Gaussovimi jedrom velikosti 41 nm in standardnim odklonom 5 nm. Iz zglajenih meritev smo obdržali le tiste vzorce, ki so sovpadali z meritvami hiperspektralnih slik.

Meritve spektrov svetil in reflektivnosti površin smo uporabili za iskanje baznih funkcij. Za iskanje baznih funkcij smo uporabili metodo ICA [17], implementirano v knjižnici MDP. Metoda v podatkih poišče statistično neodvisne bazne funkcije. Za to metodo smo se odločili, ker je porazdelitev koeficientov, ki so določeni z baznimi funkcijami, v primeru ICA podobna Laplace-ovi porazdelitvi. Absolutna vsota takšnih koeficientov je majhna, kar je v skladu z regularizacijo, ki jo uporabljamo pri iskanju rešitve. Na voljo smo imeli 1995 meritev reflektivnosti različnih površin, s katerimi smo poiskali bazne funkcije. Za spektre svetil smo uporabili zbirko „`measured_with_sources.illum`“, ki vsebuje 102 meritvi. Te smo razdelili na dva dela. Prvih 16 meritev smo uporabili za iskanje baznih funkcij, ostalih 86 meritev pa smo skupaj s slikami uporabili kot neznana svetila v sceni.

Pri izbiri števila baznih funkcij smo se držali ugotovitev iz [6]. Poiskali smo 5 baznih funkcij za spektre svetila. Za spektre dnevne svetlobe bi sicer zadostovale zgolj 3 bazne funkcije, a ker imamo v naši bazi tudi umetne vire svetlobe, smo zato izbrali še 2 dodatni bazni funkciji. Za reflektivnosti površin smo poiskali 8 baznih funkcij.

V glavnem delu eksperimenta smo uporabili 16 hiperspektralnih slik in 86 meritev spektrov svetil, ki niso bile uporabljene pri določanju baznih funkcij spektrov. Za vsako kombinacijo hiperspektralne slike in spektra svetila smo izvedli postopek, opisan v naslednjem odstavku. Kot rešitev smo podali oceno izbranega spektra svetila. Predpostavili smo, da je scena osvetljena z uniformno osvetlitvijo.

V nadaljevanju podrobneje opisujemo postopek eksperimentiranja. Iz slike smo naključno izbrali 1 % točk (oz. pikslov). Meritve slike v teh točkah smo pomnožili z

izbranim spektrom svetila. Za vsako pomnoženo meritev smo poiskali ocene koeficientov linearnih modelov. Iz ocen koeficientov spektra različnih točk smo nato podali dve končni oceni spektra svetila. Prva ocena je bila izračunana iz povprečne vrednosti najdenih ocen koeficientov. Druga ocena pa je bila izračunana iz ocen z maksimalno frekvenco. Te so bile ocenjene s histogramom, diskretiziranim na 1000 vrednosti med minimalno in maksimalno oceno posameznega koeficienta.

Napako ocene spektra smo za posamezen par slike in spektra svetila izračunali s kotno napako med ocenjenim in uporabljenim spektrom. Enačba za kotno napako je:

$$E_{kot} = \arccos \left(\frac{\mathbf{I}_{ref}}{\|\mathbf{I}_{ref}\|} \cdot \frac{\mathbf{I}_{oc}}{\|\mathbf{I}_{oc}\|} \right), \quad (11)$$

kjer vektor \mathbf{I}_{ref} vsebuje vrednosti uporabljenega spektra svetila, vektor \mathbf{I}_{oc} pa vrednosti ocenjenega spektra. V rezultatih podamo povprečno vrednost, mediano in standardni odklon te napake za vse kombinacije slik in spektrov.

Da bi lažje ovrednotili učinkovitost našega algoritma, smo le-tega primerjali z učinkovitostjo metode opisane v [18]. Ta metoda modelira družino algoritmov barvne vztrajnosti, med katere sodijo algoritmi GWA (*ang. Gray World Assumption*), WPR (*ang. White Patch Retinex*), GGW (*ang. General Gray World*) in GE (*ang. Grey Edge Algorithm*). Metodo smo implementirali tako, da smo z njo lahko ocenili vrednost spektra svetila v hiperspektralni sliki. Ker metoda obravnava barvne kanale neodvisno, smo z modifikacijo samo povečali število barvnih kanalov na 31. Vse piksele slike smo torej pomnožili z izbranim spektrom svetila in to sliko nato obdelali z algoritmi GWA, WPR, GGW stopnje 2 ter GE stopnje 1 in 2, v nadaljevanju označenima kot GE_1 in GE_2 .

Naj še omenimo, da v eksperimentih na navadnih barvnih slikah (glej [2]) dosegata metodi GE_1 in GE_2 veliko boljše rezultate od metod GWA, WPR in GGW ter sodita med modernejšee metode barvne vztrajnosti. Metode GWA, WPR in GGW so preproste metode, se pa velikokrat uporabljajo v praktičnih aplikacijah in v primerjavah z novimi algoritmi.

3.2 Rezultati

V tabeli 1 so zbrani kvalitativni rezultati. V stolpcih LLM_{MEAN} in LLM_{MAP} so rezultati naše predstavljene metode (tj. Metode na osnovi logaritemskega linearnega modela – LLM). Z LLM_{MEAN} je označen pristop, kjer je bila končna ocena spektra svetila določena kot povprečje ocen obdelanih točk, z LLM_{MAP} pa je označen pristop, kjer je bila končna ocena spektra svetila ocenjena z maksimalno frekvenco med obdelanimi točkami. Za primerjavo so podani tudi rezultati metod GWA, WPR, GGW, GE_1 in GE_2 .

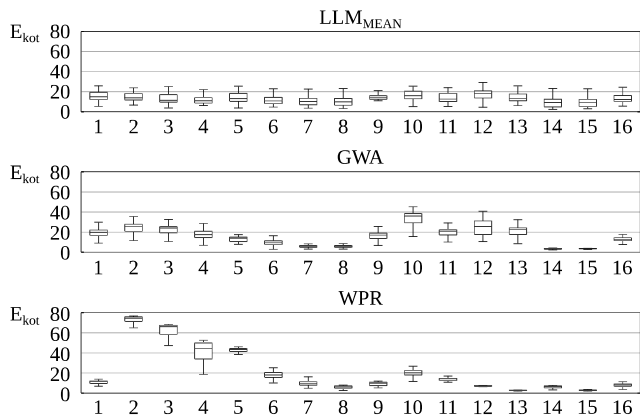
Rezultati metode WPR so imeli v določenih slikah veliko večje odstopanje od ostalih metod. Pregled teh slik je pokazal, da imajo v kanalih pri nizkih valovnih dolžinah večja

odstopanja. Na ta odstopanja pa je metoda WPR silno občutljiva. Posamezne kanale hiperspektralne slike smo zato zgladili z Gaussovimi jedrom standardnega odklona 2 in ponovno aplicirali metodo WPR. Rezultati te metode so prikazani v tabeli v stolpcu z oznako WPR2.

Tabela 1: Kotna napaka ocene spektra svetila za 12 hiperspektralnih slik in 86 spektrov svetil, za primerjane metode barvne vztrajnosti (stolpci tabele).

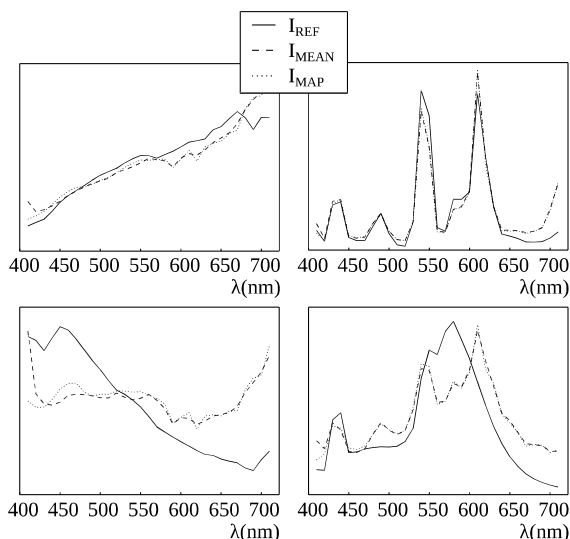
	LLM _{MEAN}	LLM _{MAP}	GWA	WPR	WPR2	GGW	GE ₁	GE ₂
povprečje	13,18	13,06	15,54	20,54	10,72	14,63	26,12	25,45
odklon	5,39	5,95	9,53	21,44	8,63	8,82	11,64	11,23
mediana	12,61	12,63	14,61	11,01	7,76	13,15	24,82	23,99

Na sliki 1 so prikazani diagrami napak posameznih slik za metode LLM_{MEAN}, GWA in WPR. Rezultati ostalih metod so bili podobni in zato niso prikazani. Iz grafov je razvidna bistvena prednost metode LLM_{MEAN}. Metodi GWA in WPR imata zelo dobre rezultate na nekaterih slikah, medtem ko je na ostalih slikah njun rezultat slabši. Metoda LLM_{MEAN} pa za vse slike dosegla podobne (stabilne) rezultate.



Slika 1: Statistika kotnih napak za metode LLM_{MEAN}, GWA in WPR po posameznih slikah. Na abscisni osi so navedene zaporedne številke slik.

Naši metodi, tj. metodi LLM_{MEAN} in LLM_{MAP} , pa ima tudi slabost. To postane očitno, ko primerjamo spekter uporabljenega svetila z ocenama spektra svetila, ki ga vrnete naši metodi. Na sliki 2 so prikazane izbrane ocene LLM_{MEAN} in LLM_{MAP} za štiri spektre svetila na isti hiperspektralni sliki. Kvaliteta rezultata naših metod je močno odvisna od spektra svetila. Metodi dobro rekonstruirata spektre umetnih svetil, ki imajo zelo značilno obliko (slika 2, graf desno zgoraj). Veliko težav pa imajo s spektri dnevne svetlobe, ki bi jih lahko dobro predstavili tudi z baznimi funkcijami reflektivnosti površin (slika 2, graf levo spodaj). Težave pa povzročajo tudi spektri z značilno obliko, ki niso bili uporabljeni pri iskanju baznih funkcij spektra svetila (slika 2, graf desno spodaj).



Slika 2: Štiri ocene spektra svetila, izračunani z metodama LLM_{MEAN} in LLM_{MAP} . Na grafu so le-ti označeni kot I_{MEAN} in I_{MAP} . Uporabljeni oz. dejanski spektri svetila so označeni z I_{REF} . Prikazane krivulje so normirane, tako da je njihova Evklidska norma enaka 1.

4. DISKUSIJA IN ZAKLJUČEK

Predstavili smo metodo za oceno spektra svetila v hiperspektralnih slikah. Rezultate smo primerjali s preprostimi metodami barvne vztrajnosti, ki dosegajo dobre rezultate na navadnih barvnih slikah. Pokazali smo, da naša metoda dosega dosti boljše rezultate od uveljavljenih metod barvne vztrajnosti, kot sta algoritma GWA in WPR.

S slabimi rezultati sta presenetili metodi GE_1 in GE_2 , ki se na barvnih slikah izkažeta veliko bolje od metod GWA in WPR. Razlog je najverjetneje omejeno število testnih slik. Ker se večina modernih metod barvne vztrajnosti opira na predpostavke o vsebini

slike, so njihovi rezultati močno odvisni od te vsebine in skoraj neodvisni od spektra svetila, ki je bilo v sliki prisotno. To je razvidno iz diagramov na sliki 1, kjer v določenih slikah metodi GWA in WPR dosegata skoraj popolne rezultate za vse uporabljene spektre. Naš eksperiment je majhno učno množico slik povečal z različnimi spektri svetil, kar pa ni imelo bistvenega vpliva na uveljavljene algoritme. Množica slik je bila premajhna, da bi njihovi rezultati bili reprezentativni.

Naša metoda poda oceno spektra svetila in reflektivnosti površine za posamezno točko hiperspektralne slike. Te ocene smo združili v oceno spektra uniformnega svetila. Predlagali smo dva načina združitve rezultatov različnih točk, in sicer povprečenje ocen spektra ali izbira ocene z največjo frekvenco ocenjeno s histogramom. Izkazalo se je, da so bile končne ocene obeh načinov zelo podobne.

Bistvena prednost predstavljene metode pred primerjanimi metodami je njena neodvisnost od vsebine slike. Moderne metode barvne vztrajnosti se močno opirajo na predpostavke o vsebini scene in ne upoštevajo relacij med različnimi valovnimi dolžinami. Naša metoda izkoristi dodatno informacijo v hiperspektralnih slikah in uspe ločiti spekter svetila in reflektivnost površine v vsaki točki posebej. Četudi bi ti oceni za posamezno točko bili napačni, je združitev ocen več točk v sliki omogočila dobro oceno spektra svetila.

Slabost naše metode je njena močna odvisnost od vrste uporabljenega spektra svetila. Metoda daje odlične rezultate za spektre značilnih oblik, kot je spekter fluorescentnih svetilk. Težave ji povzročajo spektri dnevne svetlobe, ki nimajo hitrih prehodov in jih lahko dobro opišemo tudi z baznimi funkcijami reflektivnosti površin. S predlagano metodo je tudi nemogoče popravljati vpliva svetlobe, ki ne prihaja naravnost iz svetila, ampak se je najprej odbila z obarvane površine. Želen rezultat v tej točki bi bila ocena vpadne svetlobe in reflektivnosti opazovane površine. Vpadna svetloba bi imela spekter, ki bi bil odvisen od svetila in površine, s katere se je svetloba najprej odbila. Naša metoda bi v najboljšem primeru lahko ločila spekter originalnega svetila, v reflektivnosti pa bi združila vpliv prve in druge površine.

S tem delom smo tudi pokazali, da razmerja med barvnimi kanali so pomembna in uporabna za barvno vztrajnost. Nadaljnji razvoj algoritmov barvne vztrajnosti za hiperspektralne slike bi pomagal vzpodbuditi uporabo hiperspektralnih kamer. Te so trenutno precej drage in se uporabljajo za zelo specifične aplikacije. V prihodnosti bomo naše raziskave usmerili tudi v uporabo te metode in pridobljenega znanja za razvoj algoritmov barvne vztrajnosti za navadne tribarvne slike.

ZAHVALA

Raziskavo sta finančno podprla Ministrstvo za izobraževanje, znanost, kulturo in šport in Evropski sklad za regionalni razvoj s pogodbo št. 3211-10-000464 (kompetenčni center Biomedicinska tehnika).

LITERATURA

1. M. Ebner (2007), Color constancy, John Wiley & Sons Ltd, West Sussex, England.
2. A. Gijzenij, T. Gevers, J. van de Weijer (2011), Computational Color Constancy: Survey and Experiments, *IEEE Transactions on Image Processing*, vol. 20, str. 2475–2489.
3. M. Tsukada, Y. Ohta (1990), An approach to Color Constancy Using Multiple Images, *Proceedings IEEE ICCV 1990*, str. 385 – 389.
4. M. D'Zmura, G. Iverson (1993), Color constancy. I. Basic theory of two-stage linear recovery of spectral descriptions for lights and surfaces. *Journal of the Optical Society of America A* 10, str. 2148-2165.
5. M. D'Zmura, G. Iverson (1993), Color constancy. II. Results for two-stage linear recovery of spectral descriptions for lights and surfaces. *Journal of the Optical Society of America A* 10, str. 2166-2180.
6. D.H. Foster (2011), Color constancy, *Vision Research*, vol. 51, str. 674-700.
7. Th. Gevers, H. M. G. Stokman, J. van de Weijer (2000), Colour Constancy from Multispectral Data, *The Eleventh British Machine Vision Conference*, University of Bristol.
8. A. Abrardo, L. Alparone, I. Cappellini, A. Prosperi (1999), Color constancy from multispectral images, *Proceedings ICIP 1999*, vol. 3, str. 570–574.
9. F. Pérez, B. E. Granger (2007), IPython: A System for Interactive Scientific Computing, *Computing in Science and Engineering*, vol. 9, str. 21-29.
10. E. Jones, T. Oliphant, P. Peterson et al (2001), SciPy: Open source scientific tools for Python, '<http://www.scipy.org/>'.
11. D. Albanese, R. Visintainer, S. Merler, S. Riccadonna, G. Jurman, C. Furlanello (2012), mlpy: Machine Learning Python.
12. T. Zito, N. Wilbert, L. Wiskott, P. Berkes, (2009). Modular toolkit for Data Processing (MDP): a Python data processing frame work, *Frontiers in Neuroinformatics*.
13. J. D. Hunter (2007), Matplotlib: A 2D graphics environment, *Computing In Science & Engineering*, vol. 9, str. 90-95.

14. K. Barnard, L. Martin, B. Funt, A. Coath (2002), A Data Set for Color Research, *Color Research and Application*, vol. 27, str. 148–152.
15. S.M.C. Nascimento, F. Ferreira, D.H. Foster (2002), Statistics of spatial cone-excitation ratios in natural scenes, *Journal of the Optical Society of America A*, vol. 19, str. 1484-1490.
16. D.H. Foster, S.M.C. Nascimento, K. Amano(2004), Information limits on neural identification of colored surfaces in natural scenes, *Visual Neuroscience*, vol. 21, str. 331-336.
17. A. Hyvärinen, J. Hurri, P. O. Hoyer (2009), *Natural Image Statistics: A Probabilistic Approach to Early Computational Vision*, Springer-Verlag London.
18. J. van de Weijer, T. Gevers, A. Gijsenij (2007), Edge-Based Color Constancy, *IEEE Transactions on Image Processing*, vol. 16, str. 2207-2214.

RAČUNALNIŠKI VID PRI KARAKTERIZACIJI GIBANJA ODDALJENIH OBJEKTOV

Lovro Kuščer, Janez Diaci

Laboratorij za digitalne sisteme in elektrotehniko
Fakulteta za strojništvo, Univerza v Ljubljani
E-pošta: lovro.kuscerc@fs.uni-lj.si

POVZETEK: *V prispevku sta predstavljeni dve izvedbi eksperimentalnih merilnih sistemov za določanje geografske lokacije oddaljenih objektov. Funkcionalnost ene izmed izvedb merilnega sistema smo povečali z integracijo računalniškega vida, ki omogoča samodejno sledenje premikajočim se objektom ter ponavljajoče se merjenje njihovega položaja v prostoru. Rezultati eksperimentov kažejo, da je z razvitim merilnim sistemom mogoče izvesti karakterizacijo gibanja ljudi in vozil pri oddaljenostih do 1 km.*

1. UVOD

V sodobnem času je natančno določanje položaja na zemeljski površini postalo nepogrešljivo zlasti na področju transporta, logistike, geodezije, reševanja v nujnih primerih, kmetijstva ter tudi v okoljskih in vojaških aplikacijah [1, 2, 3]. Razvoj navigacijskih sistemov in naprav je v zadnjih desetletjih doživel skokovit napredek in s tem odprl možnosti njihove uporabe na novih področjih. Vendar pa v določenih situacijah natančno poznavanje zgolj lastnega položaja, orientacije in hitrosti v nekem globalnem koordinatnem sistemu ne zadostuje. Pogosto želimo hitro pridobiti informacije o relativnem ali absolutnem položaju oziroma hitrosti nekega oddaljenega objekta, ki je lahko nedostopen, nevaren ali preveč oddaljen. V takšnih primerih se pojavi potreba po namenski napravi za oddaljeno merjenje geografske lokacije objektov in karakterizacijo njihovih gibanj.

Za namen izvedbe raziskav na tem področju smo razvili dva eksperimentalna merilna sistema na osnovi komercialno dostopnih gradnikov. Prvi predstavlja lažjo, prenosno izvedbo, ki je namenjena izvajanju meritev na težko dostopnih terenih, medtem ko je drugi prirejen za namestitev na vozilo.

Osnovno funkcionalnost razvitih eksperimentalnih merilnih sistemov smo nadgradili z integracijo računalniškega vida, ki omogoča ponavljajoče se merjenje položaja premikajočih se objektov ter s tem tudi karakterizacijo njihovega gibanja. Merilni sistemi z opisanimi zmogljivostmi so uporabni za namene, kot sta na primer nadzor vozil in plovil v prometu ali spremljanje gibanja pešcev v različnih okoljih.

2. EKSPERIMENTALNA MERILNA SISTEMA

Prenosno izvedbo merilnega sistema (slika 1) sestavljajo laserski razdaljemer, elektronski kompas in inklinomer ter modul za globalno satelitsko navigacijo (GPS). Merilnemu sistemu je dodana digitalna video kamera, ki je nameščena na okular laserskega razdaljemera in služi predvsem za dokumentacijo meritev ter zmanjšanje možnosti napačne identifikacije merjenih objektov. Kot sistemski računalnik za krmiljenje delovanja merilnih naprav je uporabljeno vezje FPGA, ki med drugim skrbi tudi za zajem in prikaz slike ter ostalih informacij na vgrajenem prikazovalniku LCOS.



Slika 1: Prenosna izvedba merilnega sistema za določanje položaja oddaljenih objektov.

Pri meritvi položaja oddaljenega objekta je najprej potrebno izmeriti položaj merilnega sistema, kar se izvede z modulom GPS. Nadalje se izvede meritev relativnega položaja oddaljenega objekta z uporabo bliskovnega laserskega razdaljemera ter elektronskega kompasa in inklinomera. Na osnovi omenjenih meritev je možno določiti položaj oddaljenega objekta v nekem globalnem koordinatnem sistemu. Glede na zmogljivosti laserskega razdaljemera je meritve možno izvajati pri oddaljenostih do 20 km.

Druga, mobilna izvedba merilnega sistema (slika 2) je osnovana na enakem laserskem razdaljemeru, medtem ko za določitev položaja in orientacije merilnega sistema uporablja drugačne senzorje. Bistvena razlika je uporaba zmogljivega GPS modula z dvema antenama, ki omogoča natančno določitev položaja merilnega sistema ter služi tudi za določanje azimuta. S tem je izločen vpliv nezanesljivosti elektronskega kompasa, čemur je podvržena prenosna izvedba merilnega sistema. Nadalje je v mobilni izvedbi uporabljeno tudi žiroskopsko stabilizirano vrtljivo podnožje, na katerem sta nameščena laserski razdaljemer in video kamera. Vrtljivo podnožje vsebuje tudi optična zaznavala zasuka, kar je še zlasti pomembno pri določanju orientacije kamere v prostoru.

V osnovni izvedbi mobilne različice merilnega sistema [4] je uporabljena video kamera, ki je namenjena iskanju oddaljenih objektov, za dokumentacijo meritev in zmanjšanje možnosti napačne razpoznave objektov na večjih oddaljenostih.



Slika 2: Mobilna izvedba merilnega sistema: (1) ohišje z laserskim razdaljemerom in kamero, (2) motorizirano podnožje, (3) elektronski inklinomer, (4) antena GPS, (5) močnostni in komunikacijski vmesnik in (6) prenosni računalnik.

3. INTEGRACIJA RAČUNALNIŠKEGA VIDA

Razvita prenosna in mobilna različica merilnih sistemov potrebujeta operaterja, ki izbere oddaljeni objekt in izvede meritev. Če se objekt premika, mu mora operater slediti z ročnim usmerjanjem merilnega sistema, kar pa je zamudno, zahtevno ter večinoma težko izvedljivo opravilo. Zato želimo realizirati samodejno sledenje izbranemu premikajočemu se objektu, kar bo omogočalo ponavljajoče se merjenje njegovega položaja in s tem tudi določitev trajektorije gibanja. Za ta namen je smiselna uporaba računalniškega vida na osnovi zmogljive video kamere. V obeh različicah merilnih sistemov sta sicer že uporabljeni video kameri, vendar zgolj za namen operaterjevega prepoznavanja oddaljenih objektov ter dokumentacije meritev. Kljub temu je integracija računalniškega vida smiselna predvsem v mobilni različici, ki uporablja motorizirano podnožje in s tem ponuja priložnost realizacije samodejnega usmerjanja merilnega sistema proti premikajočemu se oddaljenemu objektu. Za takšno funkcionalnost je potrebno video kamero namestiti na motorizirano podnožje, tako da je njena optična os vzporedna z optično osjo laserskega razdaljemera. Z uporabo ustreznih algoritmov računalniške obdelave slike je mogoče v vsakem trenutku določiti položaj izbranega objekta na digitalni sliki.

3.1 Izvedba vizualnega sledenja

Sledenje objektom v zaporedju slik je tema, ki je zaradi najrazličnejših možnih aplikacij predmet številnih raziskav [5, 6, 7]. Pojavljajo se mnoge nove metode, ki ponujajo vedno boljše rezultate ob sprejemljivi računski zahtevnosti za zmogljive sodobne procesorje.

Glede na zmogljivosti razvitih merilnih sistemov nas v predstavljeni študiji zanima predvsem sledenje premikajočim se objektom (vozilom, ljudem) na večjih oddaljenostih (okoli 1 km). Na podlagi testiranj na video posnetkih premikajočih se vozil in ljudi smo se odločili za izvedbo vizualnega sledenja na osnovi adaptivnih šablon [8] (angl. adaptive templates), s čimer delno zaobidemo težave pri spremembah osvetlitve, odbojev svetlobe ter majhnega barvnega kontrasta pri večjih oddaljenostih. Navkljub omenjenim prednostim se težave pojavijo pri hitri spremembi orientacije ali oblike sledenega objekta oziroma pri prepočasni adaptaciji šablone.



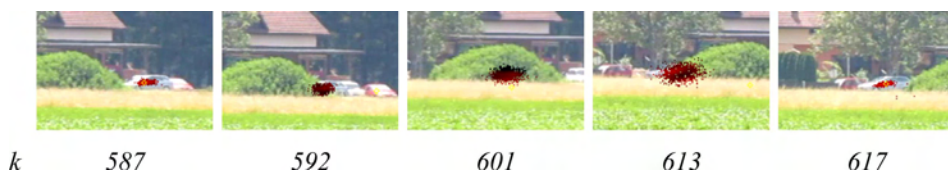
Slika 3: Adaptacija šablone pri oddaljevanju vozila.

Na začetku sledenja operater izbere objekt, ki ga želi slediti. Izgled objekta je predstavljen s šablono, ki se med potekom sledenja prilagaja spremembam videza sledenega objekta. Na osnovi šablone se nato z uporabo normalizirane križne korelacije v sliki poišče šablono najbolj podobno območje, za katerega domnevamo, da pripada sledenemu objektu. Posodobitev šablone pa nato izvedemo z uteženim seštevanjem trenutne šablone in območja z najboljšim ujemanjem. Takšen postopek adaptacije šablone v osnovi predstavlja nizkoprepustni filter, s katerim izločimo vpliv hitrih sprememb (na primer šum v ozadju, delno ali popolno zastrtje, ...), medtem ko počasnejše spremembe videza objekta filter preidejo brez pomembnih sprememb. Na sliki 3 je prikazan primer poteka adaptacije šablone (T_k) ter območja najboljšega ujemanja (N_k) med sledenjem avtomobilu. Pri postopnem delnem zastiranju objekta (na izsekih 562 in 595 avtomobil delno zakrivajo poljščine) se to sčasoma odrazi tudi v spremembi šablone.

Algoritmi sledenja večinoma vsebujejo dinamične modele sledenega objekta, ki zagotovijo konsistentnost izmerjene trajektorije z znanimi lastnostmi sledenega objekta. To pomeni, da moramo na ustrezen način združiti informacije, ki jih pridobimo na osnovi dinamičnega modela, z meritvami na osnovi ujemanja adaptivnih šablon. Eden izmed novejših, v zadnjem času vse pogosteje uporabljenih pristopov je filtriranje z delci (angl. particle filtering) [9], ki v primerjavi z alternativnim pristopom (različice

Kalmanovega filtra) ponuja dobre rezultate zlasti v primeru nelinearnih sistemov ter poljubnih porazdelitev verjetnosti. Osnova ideja filtra z delci je aproksimacija porazdelitve verjetnosti z množico točk (delcev oz. vzorcev). Za takšen opis sicer potrebujemo bistveno večje število parametrov, vendar pa lahko z njimi aproksimiramo poljubne oblike porazdelitve verjetnosti.

V primeru vizualnega sledenja uporabimo množico delcev, ki ponazarjajo možne položaje izbranega objekta na sliki ter se lahko med zaporednimi slikami premikajo skladno z dinamičnim modelom sledenega objekta. Teže delcev se določijo na osnovi normalizirane križne korelacije med šablono in območjem, ki ga posamezen delec predstavlja. Med potekom sledenja delce z majhno težo odstranjujemo, delce z veliko težo pa množimo, tako da množica delcev ostaja v območju, kjer je velika verjetnost nahajanja sledenega objekta. Za oceno trenutnega položaja objekta lahko vzamemo na primer težišče množice delcev ali pa kar koordinate najtežjega delca. Primer sledenja vozilu je prikazan na sliki 4.



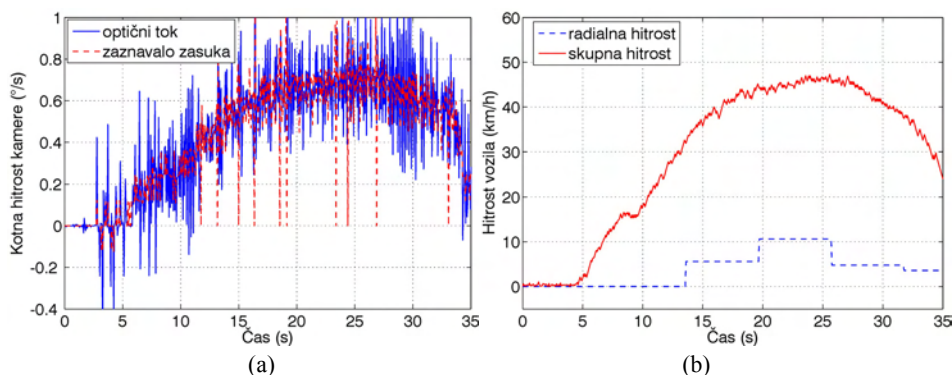
Slika 4: Sledenje s filtrom z delci pri kratkotrajnem zastrtju.

3.2 Karakterizacija gibanja

Določanje položaja objekta na sliki predstavlja prvi korak pri karakterizaciji gibanja. Če namreč želimo določiti trajektorijo gibanja objekta v prostoru, moramo poleg položaja na sliki izmeriti še oddaljenost objekta. To storimo z laserskim razdaljmerom, ki je skupaj s kamero nameščen na motoriziranem podnožju. Pri premikanju objekta je potrebno krmiljenje podnožja izvesti tako, da je optična os laserskega razdaljemera ves čas usmerjena v objekt, sicer je meritev razdalje napačna. Imamo torej primer, kjer želimo s premikajočo se kamero slediti premikajočim se objektom in na tej osnovi določati kinematske parametre opazovanih objektov. Pri tem je potrebno upoštevati, da premikanje kamere povzroči navidezne premike (tudi mirujočih) objektov na sliki. Za karakterizacijo takšnih premikov v sliki je zelo primerna analiza optičnega toka (angl. optical flow).

Z izračunom optičnega toka ozadja lahko določimo kotni hitrosti premikanja kamere, ki služita kot osnova za določanje kotnih ter kasneje tudi dejanskih hitrosti premikajočega se objekta glede na okolico. Kotne hitrosti pa lahko merimo tudi z zaznavali zasuka, ki so vgrajeni v podnožju. Primerjavo obeh meritev v primeru sledenja vozilu na razdalji 1 km vidimo na sliki 5a. Razvidno je, da se rezultati obeh meritev dobro ujemajo, kljub temu pa so opazne razlike, ki jih lahko pripišemo tresenju kamere, napakam pri izračunu

optičnega toka ter spremenljivim zakasnitvam v asinhroni komunikaciji med zaznavalom zasuka ter krmilnim računalnikom.



Slika 5: (a) Kotna hitrost kamere, izmerjena z optičnim tokom in zaznavalom zasuka; (b) hitrost sledenega vozila glede na okolico.

Med izvajanjem eksperimentov se je izkazalo, da je z razvitim merilnim sistemom na razdalji 1 km še mogoče zanesljivo izvajati karakterizacijo gibanja vozil, medtem ko je pri isti razdalji karakterizacija gibanja ljudi nekoliko manj zanesljiva. Večjo zanesljivost lahko dosežemo pri razdaljah okoli 800 m. Omejujoči dejavniki pri večjih oddaljenostih so namreč velikost sledenega objekta, divergenca laserskega žarka ter karakteristike krmilnega sistema za usmerjanje laserskega razdaljemera proti objektu.

Filter z delci v kombinaciji z adaptivnimi šablonami omogoča sledenje izbranemu objektu na sliki. Pri tem z analizo optičnega toka ozadja izvedemo stabilizacijo slike, s čimer izločimo vpliv znatnih treslajev kamere, ki so posledica skokovitih premikov koračnih motorjev vrtljivega podnožja. Z ustreznim krmiljenjem podnožja zagotovimo, da je laserski razdaljemer usmerjen proti objektu, kar omogoča izvajanje meritev oddaljenosti. S pridobljenimi podatki lahko določimo časovni potek hitrosti sledenega objekta glede na okolico, kar je prikazano na sliki 5b. Opazimo lahko, da je časovna ločljivost pri radialni komponenti hitrosti relativno nizka, kar je posledica nizke repetitivnosti laserskih meritev razdalje. Ker pa se je v prikazanem primeru avtomobil premikal pretežno v tangencialni smeri, je ta vpliv pri skupni hitrosti manj izrazit.

4. ZAKLJUČEK

Predstavljena eksperimentalna merilna sistema sta v osnovi namenjena določanju geografske lokacije oddaljenih objektov. Funkcionalnost mobilne različice smo pri tem nadgradili z integracijo strojnega vida, ki omogoča samodejno sledenje premikajočim se objektom. Za zagotovitev konsistentnosti sledenja smo uporabili filter z delci, ki združuje podatke na osnovi dinamičnega modela sledenega objekta s podatki, ki so

pridobljeni na osnovi metode adaptivnih šablon. Poleg tega je bila za stabilizacijo slike premikajoče se kamere uporabljena analiza optičnega toka.

Rezultati kažejo, da je razviti merilni sistem primeren za določanje kinematskih parametrov vozil in ljudi na oddaljenostih okoli 1 km. Opravljeni eksperimenti kažejo, da bi za sledenje vozil potrebovali višjo frekvenco in ločljivost laserskih meritev razdalje, medtem ko so rezultati meritev na osnovi računalniškega vida zadovoljivi.

LITERATURA

1. R. Madison, P. DeBitetto, A. Rocco Olean, M. Peebles (2008), Target Geolocation from a Small Unmanned Aircraft System, *2008 IEEE Aerospace Conference*, Big Sky, MT, USA str. 1-19.
2. C. Hopkinson, L. Chasmer, K. Lim, P. Treitz, I. Creed. (2006), Towards a universal lidar canopy height indicator, *Canadian Journal of Remote Sensing*, vol. 32, str. 1-14.
3. K. Han, G. N. DeSouza (2009), Instantaneous geo-location of multiple targets from monocular airborne video, *2009 IEEE Geoscience and Remote Sensing Symposium*, vol. 4, str. 1003-1006.
4. L. Kuščer, J. Diaci (2013), Measurement uncertainty assessment in remote object geolocation, *Strojniški vestnik – Journal of Mechanical Engineering*, vol. 59, str. 32-40.
5. A. G. Vicente, I. B. Munoz, P. J. Molina, J. L. L. Galilea (2009), Embedded Vision Modules for Tracking and Counting People, *IEEE Transactions on Instrumentation and Measurement*, vol. 58, str. 3004 – 3011.
6. Y. Goyat, T. Chateau, L. Trassoudaine (2010), Tracking of vehicle trajectory by combining a camera and a laser rangefinder, *Machine Vision and Applications*, vol. 21, str. 275 - 286.
7. R. Ravnik, F. Solina (2011), Paralelizacija metode AAM za sledenje obrazov v realnem času, *Zbornik konference ROSUS 2011*, Maribor: Fakulteta za elektrotehniko, računalništvo in informatiko, str. 32-37.
8. J. Ahmed, M. N. Jafri, M. Shah, M. Akbar (2008), Real-time edge-enhanced dynamic correlation and predictive open-loop car-following control for robust tracking, *Machine Vision and Applications*, vol. 19, str. 1-25.
9. B. Ristic, S. Arulampalam, N. Gordon (2004), *Beyond the Kalman Filter: Particle Filters for Tracking Applications*, London, Boston: Artech House.

PREDNOSTI IN SLABOSTI DIGITALNE OBDELAVE SLIK NA POLJSKEM ROBOTU

Peter Lepej², Peter Berk¹, Jurij Rakun¹, Miran Lakota¹

¹Fakulteta za kmetijstvo in biosistemske vede

²Fakulteta za elektrotehniko, računalništvo in informatiko

E-pošta: peter.lepej@uni-mb.si

POVZETEK: *V delu povzemamo prednosti in slabosti digitalne obdelave in analize slik, ki je prvič izvedena na ločenem, drugič pa na vgrajenem računalniko poljskega robota. Razviti algoritmi omogočajo zaznavo iskanih predmetov – rastlin, ki jih robot nato posname za nadaljno analizo ali kemično obdela (poškropi). Prednost obdelave video vsebin na ločenem računalniku se kaže predvsem v preprostosti razvoja in testiranju novih postopkov, medtem ko je prednost obdelave vsebin na vgrajenem računalniku hitrost, a za ceno robustnosti.*

1. UVOD

V kolikor želimo avtomatizirati kmetijska opravila s pomočjo mobilnih-robotiziranih konstrukcij, je nujno premagovati ovire, kot so sposobnost samodejne navigacije po polju in sposobnost zanesljive ter robustne zaznave iskanih predmetov (npr. plevela). Zato smo si za cilj zadali izgradnjo manjšega nizko cenovnega poljedelskega robota, ki bo sposoben avtonomno nadzorovati smer vožnje med vrstami koruznega polja. Uporabili smo štirikolesno mobilno platformo, avtomatizirano s pomočjo elektromotorjev, kot osnovo za naš poljedelski robot. Razvili smo algoritem, ki zazna prisotnost rastline in temelji na kombinaciji zaznavanja velikosti, oblike, ter barve zelenih elementov.

Pregled pogročja je razkril nekaj sorodnih rešitev [1, 2, 3], ki bi z nekaterimi spremembami lahko dosegli zastavljene cilje. V vseh primerih gre za robotizirane konstrukcije. Ključne slabosti sedanjih robotov za delo na polju pa so njihova velikost in teža [1], ki omejujeta mobilnost oziroma prenosljivost naprave in stroške njihovega razvoja [1].

Na univerzi v Teheranu so razvili mobilno platformo [1], ki se orientira s pomočjo računalniškega vida. Na platformi sta nameščeni dve kameri; ena na sprednji in ena na

zadnji strani ohišja. Kameri sta preko USB vodila priklopljeni na prenosni računalnik. Nadzorni program robota so razvili s pomočjo programskega paketa MATLAB in deluje na osnovi sledenja poznanih barvnih odtenkov rastlin in omogoča vožnjo med vrstami koruze. Glede na rezultate analize posameznih posnetkov, nadzorni program sprejme odločitve o kontroli krmilnega mehanizma, ki ga nadzira preko LPT vrat prenosnega računalnika. Avtorji dela so ugotovili, da njihov nadzorni program, realiziran s pomočjo programskega paketa MATLAB, v realnih razmerah deluje dokaj počasi, kar pa je posledica obdelava velike količine podatkov. Programski paket je enostaven za uporabo in primeren v fazi razvoja programa za avtonomno navigacijo platforme.

Na Univerzi za uporabne znanosti v Venlo-ju na Nizozemskem, so razvili mobilno platformo [2], ki omogoča vožnjo skozi vrste koruze. Mobilna platforma uporablja kamero, ki ima 360° kot zaznavanja v okolici mobilne platforme in je na računalnik priklopljena s pomočjo firewire vodila. Smer gibanja mobilne platforme določajo s pomočjo algoritma, ki usmerja platformo glede na zaznane vsebine. Uporabili so algoritem za zaznavo poudarjene zelene barve iz slike, ki jo posname kamera. Sliki zelenih odtenkov segmentirajo, nato pa na binariziranem posnetku uporabijo morfološke operatorje, s pomočjo katerih izločijo manjše, nepomembne objekte. Dobljeno očiščeno binarno sliko nato uporabijo, da na njej poiščejo okrogle oblike. Programsko opremo so razvili s pomočjo programskega paketa LabView in teče na prenosnem računalniku.

Na Univerzi Aalto, Fakulteta za znanstvene vede in tehnologijo so naredili mobilno platformo [3], ki se med vrstami koruze premika s pomočjo slik posnetih s pomočjo Logitech QuickCam 5000 kamere. Program za navigacijo so na vgrajenem računalniku Toradex Colibri razvili v programskem jeziku C++ in nanj naložili operacijski sistem Windows Embedded CE 6.0. Kamera zajema slike iz okolice z ločljivostjo 320×240 slikovnih elementov, kjer se algoritem navigacije zanaša na iskane barvne odtenke. Za odkrivanje rastlin koruze, so uporabili EGRBI transformacijo, ki sliko iz RGB prostora pretvori v prostor barvnih odtenkov: prekomerna zelena (EG) - rdeča-modra (RB) in intenziteta barvnih odtenkov. Cilj je odkriti zeleno barvo, kjer poudarjena zelena izključuje intenziteto spremembe tako, da je lažje ugotoviti mejne vrednosti, kot v osnovni barvni ravnini G. Rdeče-modri kanal se uporabi zato, da se ugotovi, ali je opazovani predmet bolj rdečkaste ali modrikaste barve. Problem standardne EGRBI transformacije je ta, da se osredotoča samo na zeleno-zelena barva. V primeru rastline, ki barvno ni zeleno-zelena, ampak bolj rumene barve (rdeče-zelena), kot zeleno-zelena ali zeleno-modra, ta postopek ni primeren.

V naslednjih poglavjih povzemamo delovanje naše mobilne platform – robota, ki med vrstami koruze išče znane predmete in glede na zaznavo proži aktuatorje. Zanj smo razvili svojo programsko opremo, kjer smo preizkusili kako se obnese obdelava in analiza na ločenem in nato še obdelava in analiza slik na vgrajenem računalniku. Robota opisujemo v 2. poglavju, prednosti in slabosti ene in druge obdelave in analize slik pa povzemamo v 3. poglavju.

2. ZASNOVA IN RAZVOJ POLJSKEGA ROBOTA

Da se robot lahko po polju giblje avtonomno, je zasnovan v obliki mobilne platforme s pogonom na vsa štiri kolesa, možnostjo zasuka sprednje in zadnje osi in z vgrajenim vzmetenjem za lažje premagovanje ovir. Pogon poganja trifazni brezkrtačni motor, tega pa napajajo dvo-celične litij-polimer baterije. Na robotu so nameščena tipala razdalje in digitalna kamera, ki skrbi za zajem video vsebin, vse skupaj pa je povezano na vgrajen računalnik, ki preko mikrokrmilnika krmili navigacijski del robota. Računalniški del z mikrokrmilniškimi vezjem napaja tri-celična litij-polimer baterija.

Pomembna lastnost poljskega robota je sposobnost nanosa fito-farmaceutskih snovi, kar mu omogoča vgrajen zalogovnik in par dveh elektro-magnetnih ventilov, povezanih na par nasprotno postavljenih šob. Za krmiljenje ventilov skrbi vgrajen računalnik, kar robotu omogoča nadzorovan nanos fito-farmaceutskih sredstev glede na opazovano vsebino scene.



Slika 1: Avtonomni poljedelski robot med vožnjo na polju.

2.1 Mikrokrmilniško vezje in vgrajen računalnik

Elektronika robota je razdeljena na dva dela; na vgrajen računalnik in mikrokrmilniško vezje, ki razširja funkcionalnost vgrajenega računalnika. Nalogo vgrajenega računalnika je v prvi različici prevzel računalnik VIA EPIA PX10000, ki temelji na x86 arhitekturi, ima porabo okoli 15 W moči in dosega izračun 1391 MIPS. Vgrajen računalnik zaradi nujnega aktivnega hlajenja ni bil najboljša izbira in je bil zato v izboljšani različici zamenjan z vgrajenim računalnikom BeagleBoard, zgrajenim okoli ARM Cortex A8 mikrokrmilnika, s porabo 2 W in močjo 1200 (D)MIPS. Z namenom zmanjšati nepotrebno porabo procesorskega časa, je nameščen operacijski sistem Linux, distribucija Angstrom, kjer so izbrani samo najnujnejši programski paketi.

Neglede na izbran vgrajen računalnik, robot vsebuje še razširitveno vezje z mikrokrmilnikom, ki nudi dovolj PWM izhodov za krmiljenje vseh servomotorjev, pogonskega motorja in dodatnih aktuatorjev. Zaradi pestrega in bogatega nabora perifernih razširitev je uporabljen mikrokrmilnik AVR AT MEGA 128, ki z vgrajenim računalnikom komunicira preko serijskih RS-232 vrat.

2.2 Tipala in navigacija

Razvoj poljedelskega robota je zajemal testiranje različnih tipov tipal razdalje; infra-rdeča, ultrazvočna in laserska tipala. Prva izbira so bila ultrazvočna tipala, ki v snopu izmerijo razdaljo do prve dovolj izrazite ovire. Preizkušena sta bila 2 različna tipa ultrazvočnih tipal, najprej tipala SRF04, nato pa zaradi večjega dometa, ožjega snopa in vodotesnosti še tipala Maxbotix LV-WR1. V primerjavi z infra-rdečimi tipali Sharp GP2Y0A02YK0F, saj se je izkazalo, da so ultrazvočna tipala v nekaterih primerih, kot je recimo zaznava vrst, neprimerna zaradi oblike merilnega snopa, ki je za zaznavo vrst preširok. Meritve z ultrazvokom pa se iskažejo kot neprimerne tudi v primeru zaznave manjših ovir in v primeru nepravokotnih odbojev, povzruha pa je meritve potrebno zaradi odbojev ultrazvoka opraviti zaporedno in ne istočasno na vseh tipalih hkrati. Pomankljiva se izkažejo tudi IR tipala, saj se je njihov domet v nekontrolirnih pogojih, tj. na soncu, zmanjša.

Zaradi predstavljenih pomankljivosti ultrazvočnih in IR tipal, je te zamenjal laserski merilnik razdalje SICK LMS111-10100. Ta namesto ene meritve razdalje vrača celoten nabor meritev, ki se nahajajo v izbranem območju. Merilnik zajema meritve v največ 270° območju, s 0,25° korakom in frekvenco zajema 20 Hz. Ker gre za lasersko svetlobo, njegov domet sicer na sončni svetlobi pade, a je ta še vedno večji od 10 m, kar je dovolj za večino aplikacij. Merilnik razdalje je na vgrajen računalnik priključen preko ETHERNET vmesnika.

2.3 Računalniški vid

Dogajanje v okolici robot zajema s pomočjo nameščene digitalne kamere DBK31BU03, ki omogoča zajem 30 slik v sekundi pri ločljivosti 1024 × 768 slikovnih elementov. V osnovi gre za sivinsko kamero, ki s pomočjo Bayer filtrov [4] in interpolacije [4] proizvede barvne posnetke. Kamera je na vgrajen računalnik priključena preko USB 2.0 vodila.

Digitalni posnetki iz kamere so bili uporabljeni na dva načina. V prvem poizkusu so bile ti zajeti, nato pa glede na poizvedbe odjemalca v analizo posredovani naprej na ločen računalnik. V drugem poizkusu pa so bili posnetki analizirani neposredno na vgrajenem računalniku.

2.4 Brezžična povezava

Brezžična povezava z robotom služi kot orodje za oddaljeno spremljanje dogajanja, kakor tudi kot nadzorno orodje. Osnovno povezavo z oddaljenim računalnikom predstavlja WiFi oz. IEEE 802.11g povezava, ki deluje dobro, dokler se med robotom in oddaljenim računalnikom prenašajo informacije o meritvah tipal razdalje in, povratno, informacije o krmiljenju platforme. Ko k prenosu dodamo še zajete posnetke scene, povzročimo v komunikaciji dodaten zamik. Zato smo k osnovni brezžični WiFi povezavi dodali še serijsko IEEE 802.15.4 oz. Xbee povezavo, ki je namenjena izključno krmilnemu delu komunikacije.

3. DIGITALNA OBDELAVA IN ANALIZA SLIK

Obdelavo in analizo slik ločimo glede na mesto procesiranja digitalnih slik. Prvi pristop je obdelava in analiza slik na ločenem računalniku, kjer je zaradi širokega nabora obstoječih funkcij za digitalno procesiranje slik uporabljen programski paket Matlab. Pri tem načinu je potrebno zajeto sliko najprej poslati iz vgrajenega na ločen računalnik, kar pa terja dodatno časovno zakasnitev med 0,2 in 1 s. Drugi pristop predstavlja obdelava in analiza slik neposredno na vgrajenem računalniku, kar ima sicer manjšo latenco, a je omejen iz vidika računske zahtevnosti, dodatno pa zahteva celoten razvoj funkcij za digitalno procesiranje slik. Oba pristopa imata skupen postopek zajema slik in postopek krmiljenja aktuatorjev, neglede na to ali rezultati pridejo iz vgrajenega ali ločenega računalnika.



Slika 2: Prenos slik med vgrajenim in ločenim računalnikom.

Razviti algoritmi digitalne obdelave in analize slik so bili preizkušeni na konkretnih aplikacijah v naravi, tj. v pogoji nekontrolirani, kjer se pojavijo situacije kot so močno sonce, dež in oblačno vreme. Nekontrolirani pogoji na slikah povzročajo zelo razgibano osvetlitev scene, kar lahko deloma odpravimo z adaptivnim prilagajanjem pragov [5]. V našem primeru smo uporabili Homorfni filter za odpravljanje neenakomerne osvetlitve s pomočjo nizkega sita [6]. Filter uporabimo nad vhodno sliko in šablono. Težave, kot so

popolnoma preosvetljena okolica, npr. odsev neposredne sončne svetlobe v kamero, pa smo reševali s podatki iz zabeležene zgodovine prilagajanja pragov. Na tako korigiranih posnetkih se nato opravi prepoznavna, kjer smo prepoznavali tako umetne, kot tudi realne kmetijske pridelke, kot je koruza, plevel med koruso in poljsko rožo (sivka).

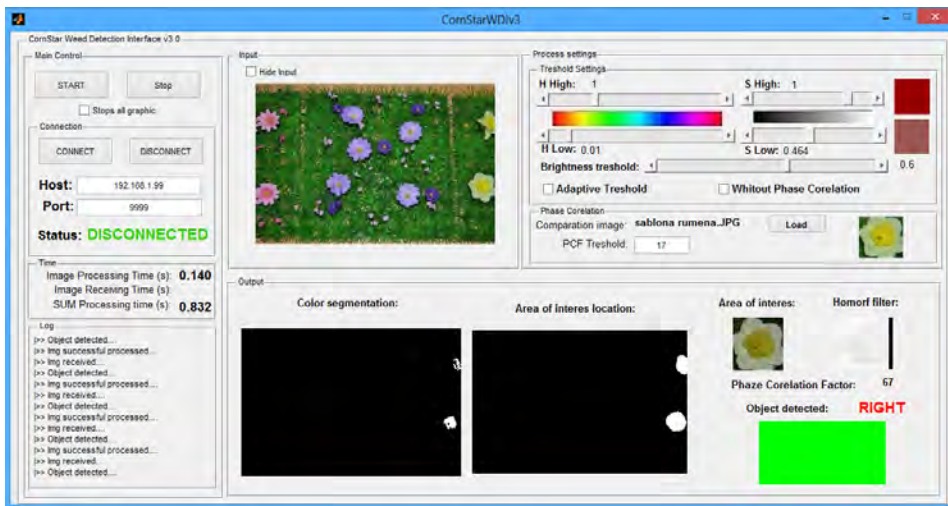
Pomemben dejavnik pri realno časovni obdelavi je hitrost prejemanja vhodnih podatkov kot je digitalna slika, čas potreben za obdelavo slike in čas potreben za izvršitev akcije. Skušali smo zmanjšati čas obdelave z izvozom Matlab kode v izvršljivo obliko, vendar s tem nismo pridobili očitne izboljšave. Vsota teh časov narekuje hitrost ter posledično učinkovitost robota. Za dober rezultat prepoznavne pridelkov pa je dodatno zelo pomembna robustnost pristopa, ki določa zanesljivost robota. V kolikor je pristop dovolj zanesljiv, je lahko uporabljen v različne namene, kot so npr.: selektivni nanos fitofarmaceutskih sredstev, štetje in analiza količine ter odkrivanje bolezni pridelka. V delu sta tako predstavljena dva pristopa k prepoznavi objektov kompleksnejših oblik, kot je npr. sivka ("*Lavandula*") ali druge vrste rož z izrazito obliko in barvo.

3.1 Na ločenem računalniku

Ozko grlo v postopku obdelave in analize slike na oddaljenem računalniku je nedvomno prenos slike preko brezžičnega omrežja. Obdelava se izvaja po uspešno vzpostavljeni komunikaciji in po prejemu in potrjevanju slike. Omogoča pa takšna obdelava preprosto uporabo že razvitih naprednejših postopkov, kakor tudi prikaz, kot je prikazan na sliki 3. Na njej so razvidni časi, ki so bili potrebni za procesiranje konkretnega primera pri iskanju rože z vijolično barvo. Najprej je potrebno opraviti TCP/IP povezavo na vgrajen računalnik, ki na poizvedbo odgovori z zajetimi digitalnimi slikami okolice. Naslednji korak je pravilna nastavitvev pragov glede na HSL barvni prostor [4,5], kar postorimo empirično z nastavitvijo drsnikov.

Naslednji postopek predstavlja iskanje podobnih oblik. V ta namen se definira šablono, ki služi kot referenca pri iskanju podobnih oblik [4]. Začnemo z določitvijo šablone, katero si predpripravimo, glede na to katere vrste rož bomo iskali in jo naložimo. Vse vhodne slike se pretvorijo v HSL barvni prostor. Prvi korak postopka iskanja podobnih oblik je barvna segmentacija [4], ki ji sledi združevanje območji [4] na podlagi katerih se določijo območja zanimanja. Nad izpostavljenim območjem se nato izvede homorfni filter za odstranitev neenakomerne osvetlitve. Zadnji korak je izvedba fazne korelacije [4] med definirano šablono in potencialnim območjem zanimanja. Glede na dobljene rezultate fazne korelacije se pragovno metodo na koncu sklepa o pristnosti iskanega predmeta oz. pridelka.

Slika 3 prikazuje zaznavo rumene rože, v pripravljenem uporabniškem vmesniku. Tako je določena šablona, v kateri je prikazan en vzorec iskane rože, ter se prilagodi drsnike glede na ustrezno barvno rože. Omenjeni primeri so bili izvedeni v kontroliranem okolju. Podobno se lahko naredi še za ostale rože, ki se nahajajo na vhodni sliki.



Slika 3: Uporabniški vmesnik za prepoznavo objektov.

Uporabniški vmesnik je zgrajen iz kontrolnih gumbov, kjer se lahko preko brezžične povezave opravi povezava na vgrajen računalnik, ki se nahaja na robotu, pod tem oknom pa se nahaja še dnevnik dogodkov. Desno je vidna vhodna slika ter nastavitve pragov. Pod nastavitvijo pragov pa se naloži željena šablona iskanega elementa. V spodnjem delu vmesnika se nahajata dve sliki, ki prikazujeta segmentirano sliko (levo) in razširjeno sliko (desno). Skrajno desno na uporabniškem vmesniku pa je prikazano trenutno opazovano območje, ter ista slika z uporabljenim homorfnim filtrom. Rezultat zaznave pa se nahaja spodaj desno. Na sliki 4 so prikazani primeri prepoznav z drugimi vzorci šablone na primeru iste vhodne slike.



Slika 4: Prepoznavna različnih objektov, vijolične levo in roza rože desno.

3.2 Na vgrajenem računalniku

Zaradi omejitve paketov operacijskega sistema na vgrajenem računalniku, ta nudi bogatega uporabniškega vmesnika, kot smo ga videli v prejšnjem podpoglavju. Do vgrajenega računalnika se namreč povežemo preko brezžične povezave, nato pa z njim komuniciramo v ukazni lupini, ki nudi vmesnik samo za spremljanje osnovnega stanja

delovanja ter vklop in izklop sistema. Vse zaznave oz. prepoznavne so tako vidne posredno preko delovanja fizičnih sistemov na robotu, kot je vklop posameznih aktuatorjev (škropilnica, LED,...). Matlabov interpreter z vgrajenim naborom funkcij v tem primeru zamenja programski jezik C, ki ga prevajalnik gcc prevaja v izvršljivo obliko.

Zajem slike se v tem primeru izvede na podoben način kot pri obdelavi na ločenem računalniku, le da se slika po zajemu obdela neposredno na vgrajenem računalniku. Obdelava zajema postopek barvne segmentacije in postopka združevanja prepoznanih območij. Zaradi hitrosti obdelave in manjše procesorske moči niso bile uporabljene kompleksnejše funkcije prepoznavne robov, Houghovega filtriranja in fazne korelacije, ampak prepoznavna temelji izrecno nabarvana segmentacija z adaptivno nastavljivimi pragovi. Funkcije za adaptacijo pragov so bile določene empirično.

Prednosti in slabosti obdelave na oddaljenem računalniku so povzete v tabeli 1.

Tab. 1: Tabela prednosti in slabosti obdelave slik na ločenem in vgrajenem računalniku.

LOČEN RAČUNALNIK		VGRAJEN RAČUNALNIK	
Prednosti	Slabosti	Prednosti	Slabosti
Kompleksnejši algoritmi	Počasen prenos slike-posledično počano delovanje	Ni potrebno brezžično omrežje	Povezava samo za spremljanje stanja
Uporabniku prijazen uporabniški vmesnik	Potrebno brezžično omrežje	Delovanje brez rizika odpovedi zaradi izpada mreže	Potreben zmogliveši procesor
Takojšne nastavljanje parametrov	Nujen dodaten, ločen računalnik	Hitrejše delovanje	
Preprosta integracija z drugimi napravami		Večja odzivnost	
Preprostejše delo in razvoj algoritmov		Manj zapletena uporaba - namestitvev	
Robusnejši (računsko zahtevnejši) postopki			

4. SKLEP

V delu sta povzeta dva načina digitalne obdelave slik, ki se med seboj ločita glede na lokacijo obdelave in uporabljene postopke. V prvem načinu je bila obdelava in analiza slik opravljena na ločenem, v drugem pa neposredno na vgrajenem računalniku. Prvi se izkaže preprostejši za razvoj, drugi pa primernejši za uporabo na terenu.

Smernice nadaljnjih raziskav kažejo v smeri vključevanja komprimiranja slik, ki jih privzeto uporabljena oprema sicer ne podpira, se pa lahko vključi programsko, torej na vgrajenem računalniku. Zaradi manjše količine podatkov, bi se tako pohitрил čas prenosa v primeru obdelave slik na ločenem računalniku, a ob zakupu, da to hkrati nekoliko povečal čas režije ob zajemu slik.

Druga izmed smernic pa kaže v smer uporabe obstoječih knjižic za digitalno procesiranje slik, ki so na voljo za operacijski sistem Linux, npr. gstreamer ali imlib2. S prvo lahko oblikujemo tok slik (pretvorba barvnih prostorov, (de)komprimiranje, ...), z drugo pa opravimo osnovne operacije (filtriranje, rotacija/translacija/skaliranje, ...) nad sliko. Prvi zahteva na izbranem opreacijskem sistemu dodatno podporo številnih drugih knjižic, drugi pa ponuja le naj osnovnejše operacije.

LITERATURA

- [1] H. H. Harms, T. Lang, J. Schattenberg (2011), *Proceedings of the 8th Field Robot Event, Institute for Agricultural Machinery and Fluid Power*, Braunschweig.
- [2] J. M. W. H. Gennip (2012), *Field Robot Event Booklet*, spletna publikacija: <http://www.fieldrobot.nl/downloads/program%20booklet%20FRE2012%20V6.pdf>, Venlo.
- [3] T. Kemppainen, T. Koski, J. Hirvelä, J. Lillhannus, T. Turunen, J. Lehto, V. Koivisto, M. Niskanen, T. Oksanen, J. Kostamo, P. Tamminen (2009), *Robot Brothers EasyWheels and ReD in Field Robot Event, Proceedings 7th Field Robot Event*, Wageningen.
- [4] R.C. Gonzalez, R. E. Woods (2002), *Digital Image Processing*, second editon, Prentice Hall.
- [5] S. Florczyk (2005), *Robot Vision*, Viley-VCH Verlag GmbH & Co. KGaA.
- [6] K. Delac, M. Grgic, T. Kos (2006), Sub-image homomorphic filtering technique for Improving facial identification under difficult illumination conditions, *International Conference in Systems, Signals and Image Processing - IWSSIP 06*, str. 95-98.

MEDICINSKE IN BIOMEDICINSKE APLIKACIJE

AUTOMATIC CALIBRATION OF VIDEO-BASED GAZE TRACKING DURING ROBOT-ASSISTED GAIT REHABILITATION

Simon Zelič, Matjaž Divjak, Aleš Holobar

System Software Laboratory

Faculty of Electrical Engineering and Computer Science

E-mail: simon.zelic@uni-mb.si

ABSTRACT: *Quantification of patient's attention to visual feedback plays an important role in assessment of patient's compliance with neurorehabilitation. One of many possible approaches builds on a remote gaze tracking with a high-speed camera mounted close to the feedback screen. Such an approach is completely unobtrusive and, thus, easy to use. However, it requires fast automatic calibration of a gaze tracker to each individual patient.*

We present a procedure for automatic calibration of a remote SVM-based gaze tracker in robot-assisted rehabilitation of stroke patients. First, automatic detection of intentional eye movements is presented. Afterwards, the information on the extracted eye movements is combined with synchronously acquired positions of gaze targets in purposely built calibration screen. As a result, the intervals in which the subject is observing the corresponding gaze targets are identified with a high confidence and used for definition of SVM learning set. When applied to video recordings of robot-assisted gait rehabilitation of three healthy controls and two stroke patients, the presented automatic learning of SVM-based gaze tracker resulted in classification of five gaze directions with accuracy of $95\% \pm 3\%$ and $89\% \pm 7\%$, respectively.

1. INTRODUCTION

EU project BETTER (FP7-2009-7.2-247935) utilizes top-down rehabilitation approach based on the visual feedback from EEG and EMG measurements to improve rehabilitation of gait-impaired patients after stroke. The rehabilitation system consists of multiple subsystems (Figure 1). The main rehabilitation utility is Hocoma's Lokomat [1] - a robot with automated gait orthosis on a treadmill. A large screen is placed in front of the Lokomat displaying visual feedback on gait rehabilitation. In the framework of BETTER project, different visual feedback modalities have been developed and/or tested, including Hocoma's graphs displaying the exerted knee and hip forces, co-activation patterns of affected leg muscles and 3D virtual environments showing the virtual avatar from the first or third person perspective (Figure 1).

In order to track the patient's attention to visual feedback, a high speed Basler Ace acA2000 CameraLink camera with frame-rate of 100 Hz is mounted on the top of the

screen displaying the visual feedback. In order to support their accurate synchronization in time, both video acquisition and graphical feedback run on the same personal computer. By using active appearance model (AAM) [2], 60 facial features are tracked in each video frame, including eyes, nose and mouth corners. A fast radial symmetry algorithm [3] is used to refine the position of pupils in real time. Afterwards, the algorithm proposed by Gee et al. [4] is applied to assess the patient's head pose. Finally, all the aforementioned extracted features are fed into SVM classifier to determine the patient's gaze direction during the robot-assisted gait rehabilitation.

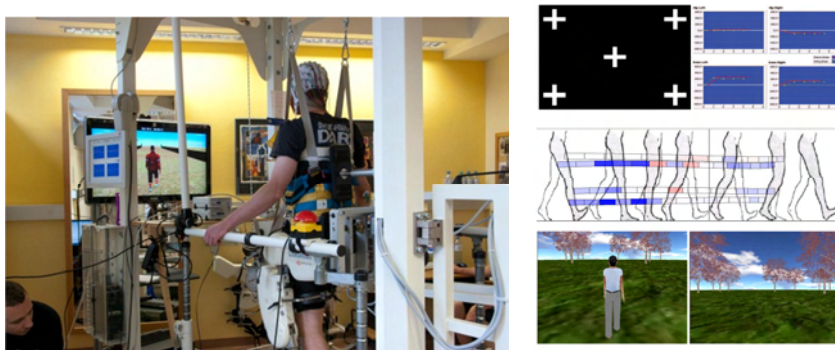


Figure 1: Rehabilitation environment; *Left panel* – Lokomat with screen for visual feedback and mounted high-speed camera above the screen; *Right panel* - different feedback modalities including the GAZE screen - video tracking calibration screen (*top left subpanel*), Hocoma's line feedback (*top right subpanel*), BETTER's top-down feedback (*central subpanel*), VR male avatar from the 3rd person perspective (*lower left subpanel*), and VE from the 1st person perspective (*lower right subpanel*).

The aforementioned gaze tracking relies significantly on the successful learning of SVM classifier, which must be fully automatic, fast and highly robust to errors in order to be performed during a short calibration to each individual stroke patient. The calibration utilizes purposely built gaze target screen (Figure 1), which sequentially displays the gaze targets in the four corners and in the centre of the screen. In order to support fast convergence and optimal classification rates of SVM, an automatic selection of the learning set must include representative and error-free samples. In this manuscript, we describe the procedure and algorithms used for automatic creation of such reliable SVM learning sets.

2. EXPERIMENTAL PROTOCOL

The experiments were carried out in the Rehabilitation Clinic Judendorf-Strassengel (Austria) and at Fondazione Santa Lucia (Rome, Italy). Three healthy volunteers, without neurological or locomotor deficits, and two stroke patients participated in this study. The experimental procedures were approved by the local ethical committees. Each participant gave written informed consent prior to the experimental session.

The experiment comprised five runs of robot-assisted walking, with five different feedback modalities as presented in the right panel of Figure 1. In all runs, the participants were instructed to walk actively, maintaining a constant speed and applying minimum force on the robot. A 42 inch screen was placed in front of the Lokomat, 1.4 m away from participant's face.

Experienced professional therapists were present during the experiments and helped to adjust the robot to the participants and to configure the avatar feedback. The parameters of the Lokomat, namely speed, body weight support, and force, were fixed before starting the experiment and kept constant during all the experimental runs. According to common clinical practice, speed was below 3 km/h and adjusted to the participant's size according to the formula: $\text{speed} = 0.54 * (\text{leg}) / 27.8$, where leg is the participant's leg length in cm and speed is given in km/h. Walking speed ranged from 1.8 to 2.2 km/h. Body weight support (BWS) was adjusted with the help of the therapists to a participant-specific minimum. This adjustment was necessary to ensure a correct activation of the foot switches (left and right heel) used to track the participant's gait cycle. The adjustment resulted in a BWS below 30% in all the participants.

3. SVM LEARNING SET

In order to automatically select the video frames with the most representative and error-free facial features, we developed an algorithm for detection of intentional eye movements and various spurious events, such as eye blinks and sudden head movements away from the feedback monitor. This is not a trivial task, as during the gait rehabilitation, the patient's head moves constantly.

The proposed algorithm design started with assessment of the average movement of the extracted facial landmarks, as depicted in Figure 2. Not all extracted AAM points follow the global average of head movement equally well. From recorded data, the eye corners (points 17, 18 for left and 22, 23 for right eye in Figure 2) have been empirically identified as the most robust indicators of average head movements. Thus, those four points have been used to approximate the movement of pupils due to the gait-induced head movement.

In order to minimize the effect of head movement, the average position of eye corners was subtracted from pupil position in each eye and each video frame. In the next step, derivatives of pupil movements in horizontal and vertical direction were calculated. Afterwards, cumulative sum of all consecutive unidirectional pupil movements is calculated. This yields a noisy representation of pupil movement (Figure 3), with the noise originating from the not completely suppressed head movements and pupil detection errors. In order to suppress this noise and detect the intentional pupil movements, we can either use a predefined threshold (i.e., a minimum distance accepted as intentional pupil movement), or detect this threshold from the extracted signal itself. Both approaches are acceptable and have their own cons and pros. However, for fully automatic detection of gaze intervals, the second approach is preferred, since it is unrelated to camera properties and changes in the patient's positions.

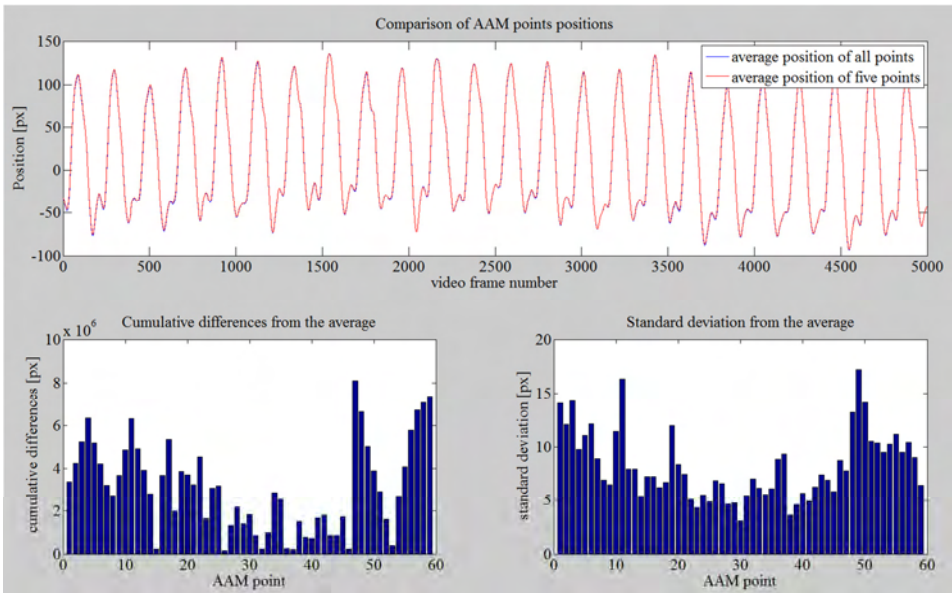


Figure 2: Result of horizontal head movement analysis for one of healthy controls. Movement of all points was brought to the same basis and averaged (*top*). Cumulative difference from mean value (in pixels) was calculated for each point separately (*bottom left*), along with standard deviation (*bottom right*).

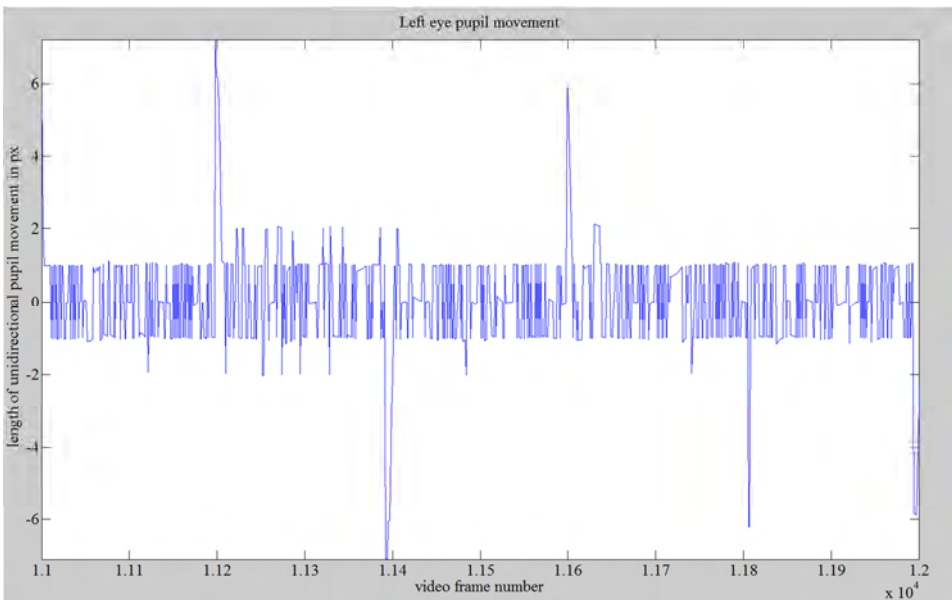


Figure 3: Length of unidirectional left eye pupil movement in left (+) or right direction (-). Large spikes result from intentional eye movements.

In the last step, the resulting signal is coupled with the synchronously acquired information on gaze target position (Figure 4). Since the exact position of gaze targets and their sequence of appearance are known, the intervals in which the patient's gaze follows the gaze targets with a high confidence can be identified by employing a set of simple rules. Two examples of such rules are:

If gaze target position changes from the centre to the corner of the screen, then movement of both right and left pupils must be detected in both x and y directions;

If gaze target changes from one to the other corner of the screen, then pupil's movement should be detected only in the corresponding direction.

Similar set of rules can be determined for detection of eye blinks and sudden spurious head movements.

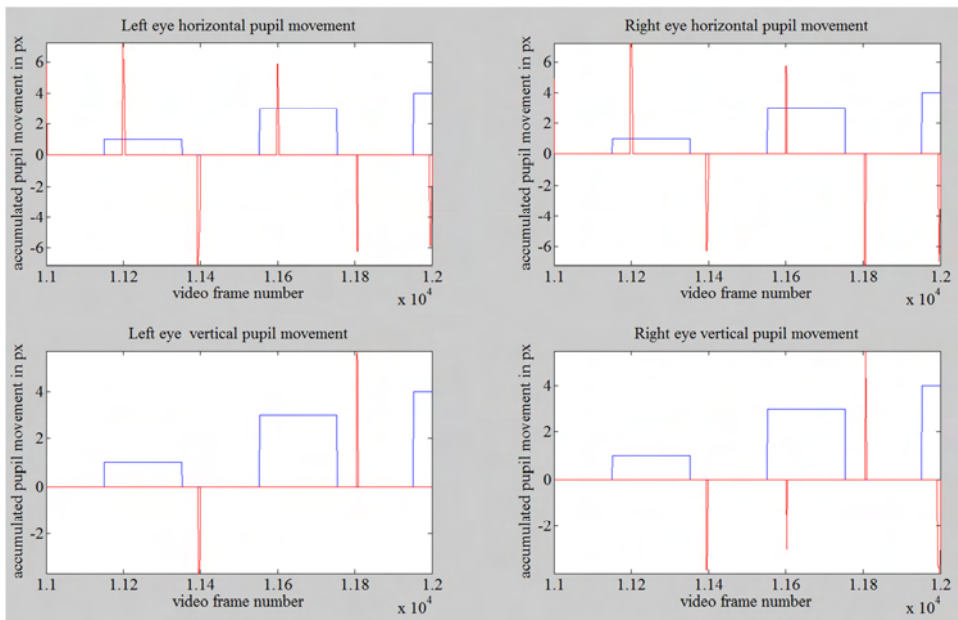


Figure 4: Final result of presented procedure for the left and right eye. For each change of position of gaze target (blue), the intentional pupil movement (red) is detected simultaneously in both eyes. Vertical eye movements are smaller and more difficult to detect than horizontal ones.

4. RESULTS

The proposed algorithm was applied to video recordings of three healthy controls (S) and two stroke patients (P) during their gait rehabilitation on Hocoma's Lokomat (Figure 1). For each subject two four minutes long videos of gaze screen observation were recorded. About half of the acquired video frames were used in aforementioned construction of SVM training set. The rest of the video was used for automatic

construction of the test set (following the same frame selection procedure as in the case of the test set). Exact distribution of the video frames for the first recording of each subject is presented in Table 1, along with the obtained SVM classification accuracy. Classification confusion matrices are presented in Figure 5. In healthy controls, the examined gaze directions were identified with average accuracy of $95\% \pm 3\%$. Most of errors originated from confusion between top-left vs. bottom-left and top-right vs. bottom-right gaze directions. In stroke patients, the examined gaze directions were classified with accuracy of $89\% \pm 7\%$.

Table 1: Number of automatically identified frames in training and test sets and accuracy of SVM-based gaze classification

Subject	recording session	No. of frames selected for training / all video frames	No. of frames selected for testing / all video frames	Classification Accuracy [%]
S6	R9	7307 / 26201	6383 / 26201	98.1
S7	R8	6662 / 25951	7313 / 25951	95.1
S8	R5	6295 / 25501	4596 / 25501	98.6
P4	Se11, R1	1704 / 4301	1815 / 4301	97.0
P7	Se4, R6	5572 / 34401	6026 / 34401	93.2

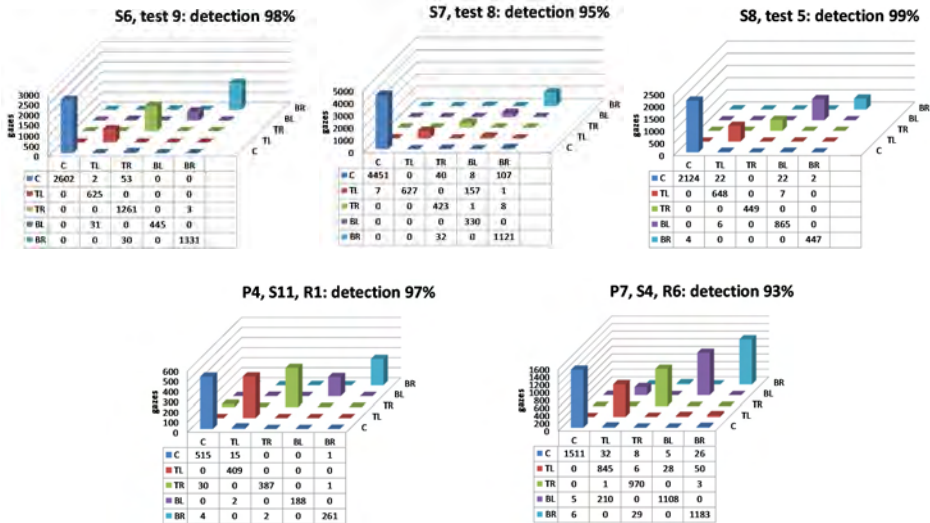


Figure 5: Confusion matrix of automatically trained SVM-based gaze tracking in three healthy controls (S6-S8) and two stroke patients (P4 and P7); Gaze directions: C – centre, TL – top-left, TR – top-right, BL – bottom-left; BR – bottom-right.

5. CONCLUSION

The proposed automatic selection of the learning set for SVM-based classification of gaze direction supports accurate gaze tracking and its fast adaptation to each individual patient or recording session without the need for expert's supervision. Moreover, by calculating the delay between the appearance of the gaze target and the corresponding intentional eye movement, the proposed procedure can also be used as an objective tool for quantification of patient's response times (results not shown). The latter should be taken into account when designing the dynamics of visual feedbacks, especially in the case of severely affected patients.

Current limitation of the proposed algorithm is its conservative identification of the representative and error-free video frames. As demonstrated in Table 1, the number of frames selected for learning and testing, respectively, is usual low. Thus, when compared to manual selection of the SVM learning set, slightly longer video recordings are required in the calibration phase of fully automatic gaze tracker. Noteworthy, this calibration has to be repeated every time the relative position of the screen with respect to the patient is changed. In the future, we plan to optimize each individual steps of the proposed procedure and increase the number of frames selected for learning or testing of proposed SVM classification. This would reduce the length of the gaze tracker calibration phase, making it more appealing to both therapists and patients.

ACKNOWLEDGMENT

This work was funded by the Commission of the European Union, within Framework 7, under Call "ICT restoring and augmenting human capabilities compensating reduced motor functions or disabilities ", Grant agreement FP7-2009-7.2-247935 "BETTER – Brain-Neural Computer Interaction for Evaluation and Testing of Physical Therapies in Stroke Rehabilitation of Gait Disorders".

REFERENCES

1. Lokomat®Pro - Enhanced Functional Locomotion Therapy with Augmented Performance Feedback, <http://www.hocoma.com/en/products/lokomat/>
2. I. Matthews, S. Baker: Active Appearance Models Revisited, *International Journal of Computer Vision* 60(2) , pp 135-164.
3. G Loy, A. Zelinsky: "A Fast Radial Symmetry Transform for Detecting Points of Interest. *IEEE PAMI*, 25 (8). pp 959-973, 2003.
4. A.H. Gee and R. Cipolla: Determining the gaze of faces in images, *Image and Vision Computing*, 12. pp. 639-647, 1994.

MODELI ŠIRJENJA SVETLOBE V OPTIČNO GOSTIH MEDIJH IN NJIHOVA UPORABA PRI DOLOČANJU OPTIČNIH LASTNOSTI

Jurij Jemec, Miran Bürmen, Boštjan Likar, Franjo Pernuš;
Fakulteta za elektrotehniko, Univerza v Ljubljani, Tržaška 25, SI-1000
Ljubljana, SLOVENIJA

E-pošta: jurij.jemec@fe.uni-lj.si

URL: <http://lit.uni-lj.si>

POVZETEK: *Spektroskopija v bližnje infrardečem območju (ang. near infrared, NIR) je neinvazivna in preprosta metoda za določanje vibracijskih lastnosti vzorca. Na podlagi izmerjenega spektra je s primernimi postopki možno določiti optične lastnosti vzorca, ki nam dajo pomembno informacijo o njegovi sestavi. V ta namen je potrebno imeti ustrezen matematični model širjenja svetlobe. Za opis širjenja svetlobe v optično gostih medijih je najpogosteje uporabljena enačba sevalnega prenosa energije (ang. radiative transfer equation, RTE) ter njeni približki. Razen v redkih primerih, ko je optične parametre mogoče določiti analitično, je standarden postopek iskanje takih parametrov modela, ki dajo najboljše ujemanje z izmerjenim spektrom. V prispevku je podrobno opisana RTE, difuzijska enačba (ang. diffusion equation, DE) kot njen najpogosteje uporabljan približek, drugi načini približkov RTE in njihovo reševanje. Podan je tudi krajši pregled postopkov za merjenje optičnih lastnosti.*

1. UVOD

NIR spektroskopija dobiva vedno več pozornosti kot postopek za kontrolo kakovosti prehranskih in farmacevtskih izdelkov, diagnostično orodje v medicini, pri forenzičnih preiskavah, v petrokemiji, kemični industriji, lesni industriji idr. Zanimiva je zaradi neinvazivnosti metode ter enostavne uporabe brez večjih zahtev za pripravo vzorca. Območje merjenja obsega elektromagnetno (EM) valovanje valovnih dolžin 750-2500 nm, posebej zanimivo pa je t.i. diagnostično okno, ki se nahaja pri valovnih dolžinah 700-900 nm. V tem območju se namreč nahajajo lokalni minimumi absorpcijskih spektrovoksi- in deoksihemoglobina, lipidov in vode [1], prav tako pa je v tem območju že zmanjšana absorpcija melanina. To omogoča fotonom teh valovnih dolžin, da prodrejo globlje v tkivo ter pridobitev informacije o tkivu globine od nekaj milimetrov do nekaj centimetrov [1].

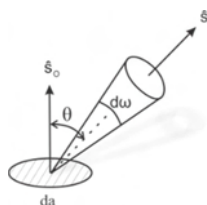
Informacijo o tkivu vsebujejo njegove optične lastnosti. V določenih konfiguracijah jih je mogoče meriti direktno, večinoma pa poteka njihova določitev prek meritev odbojnosti (ang. reflectance) in prepustnosti (ang. transmittance) tkiva. V tem primeru je

potrebno izdelati model širjenja svetlobe v mediju ter nato reševati »inverzni problem«, kjer gre za določevanje parametrov parcialne diferencialne enačbe, ki opisuje model. Problem je rešljiv analitično le pri poenostavitvah in za najenostavnejše geometrije, zato se rešitev išče numerično. Za opis širjenja svetlobe v optično gostem mediju je pogosto uporabljana RTE, čeprav ni izpeljana iz teorije širjenja EM valovanja, vendar meritve opravičujejo njeno uporabo. Zaradi kompleksnosti analitične rešitve RTE je bila izpeljana vrsta njenih približkov npr. difuzijska enačba, Bornova aproksimacija, Kubelka-Munk aproksimacija itd. V nadaljevanju sledi opis fizikalnega ozadja RTE, razlaga njenih aproksimacij, na koncu pa sledi še pregled postopkov za merjenje optičnih lastnosti tkiva.

2. MODELI ŠIRJENJA EM VALOVANJA

2.1. Opis z RTE

RTE za razliko od obravnave širjenja EM valovanja z Maxwellovimi enačbami svetlobo obravnava kot delce ter neposredno opisuje širjenje energije skozi medij. RTE ne temelji na teoretičnih osnovah porazdelitve polja, temveč je uporabljena, ker njena rešitev dobro opisuje eksperimentalne rezultate širjenja svetlobe v optično gostem mediju [2, p. 1361].



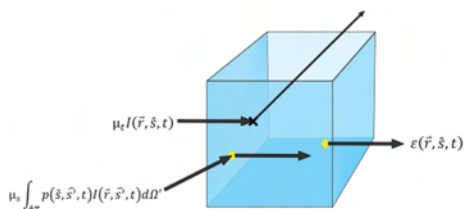
Slika 1: Specifična intenziteta v določeni točki.

Osnovna količina v RTE za opis širjenja energije je svetlost oz. specifična intenziteta (ang. radiance, specific intensity) $I_S(\vec{r}, \hat{s}, t, \nu)$ [enota: $\frac{W}{m^2 sr Hz}$] (Slika 1), ki predstavlja povprečen pretok moči v določeni točki \vec{r} in je posledica fotonov s frekvenco ν znotraj enote prostorskega kota okoli vektorja \hat{s} [3, p. 29]. Če privzamemo, da se frekvenca valovanja s širjenjem po mediju ne spreminja ter obravnavamo monokromatsko valovanje, lahko RTE zapišemo v obliki [3, p. 32]

$$\begin{aligned} & \frac{\partial}{\partial t} I(\vec{r}, \hat{s}, t) + \nabla \cdot [\hat{s} I(\vec{r}, \hat{s}, t)] \\ & = -\mu_t I(\vec{r}, \hat{s}, t) + \mu_s \int_{4\pi} p(\hat{s}, \hat{s}', t) I(\vec{r}, \hat{s}', t) d\Omega' + \varepsilon(\vec{r}, \hat{s}, t), \end{aligned} \quad (1)$$

pri čemer je ν hitrost svetlobe v mediju, $\varepsilon(\vec{r}, \hat{s}, t)$ je člen, ki predstavlja svetlobni vir, ki oddaja moč ob času t na enoto volumna in enoto prostorskega kota v smeri \hat{s} , $d\Omega'$ pa je diferencialni prostorski kot v smeri \hat{s}' . Leva stran enačbe predstavlja materialni odvod (ang. material derivative) $\frac{D}{Dt} I(\vec{r}, \hat{s}, t)$, ki opisuje spreminjanje svetlosti po času in v smeri

\hat{s} . Na desni strani pa imamo prispevke, ki so posledica sipanja iz smeri \hat{s} ter absorpcijo v točki \vec{r} , kar prispeva k zmanjšanju svetlosti, sipanje iz drugih smeri v smer \hat{s} ter viri pa prispevajo k povečanju svetlosti. Razmere ponazarja Slika 2. Numerično reševanje RTE je računsko izjemno zahtevno, analitično je bila rešena le za določene primere, splošna analitična rešitev RTE pa ne obstaja, zato so pri iskanju rešitev pogosto uporabljeni približki. Pri analizi medijev, kjer je prevladujoče večkratno sipanje je najširše in najuspešnejše uporabljan difuzijski približek [3, p. 38].



Slika 2: Možne spremembe svetlosti v točki \vec{r} .

2.1.1. Difuzijska enačba (DE)

RTE je integralno diferencialna enačba, kjer je računana veličina svetlost, medtem ko se pri DE enačbi računa svetlobni tok (ang. fluence rate), ki je definiran prispevek svetlosti iz vseh smeri in ga lahko zapišemo kot [4, p. 34]

$$\Phi(\vec{r}) = \int_{4\pi} I(\vec{r}, \hat{s}) d\omega. \quad (4)$$

Za izpeljavo DE iz RTE zato izvedemo integracijo RTE po vseh prostorskih kotih. Tako dobimo:

$$\frac{\partial}{\partial t} \Phi(\vec{r}, t) + \nabla \cdot \vec{J}(\vec{r}, t) + \mu_a \Phi(\vec{r}, t) = \int_{4\pi} \epsilon(\vec{r}, \hat{s}, t) d\Omega. \quad (5)$$

Sedaj je potrebno izraziti še vektor svetlobnega pretoka $\vec{J}(\vec{r}, t)$ ¹ (ang. net flux vector) kot funkcijo svetlobnega toka $\Phi(\vec{r}, t)$. To se izvede s ponovno integracijo RTE ter množenjem enačbe s smernim vektorjem \hat{s} . Celotna izpeljava in integracija posameznih členov RTE je izvedena v [3, pp. 246-249]. Končni rezultat je

$$\vec{J}(\vec{r}, t) = -\frac{1}{\mu_t - g\mu_s} \left[\frac{\nabla \Phi(\vec{r}, t)}{3} - \int_{4\pi} \epsilon(\vec{r}, \hat{s}, t) \hat{s} d\Omega \right]; \quad (6)$$

pri čemer lahko zapišemo $\frac{1}{3(\mu_t - g\mu_s)} = \frac{1}{3(\mu_a + \mu_s)}$ kot difuzijski koeficient. Enačba (6) je obenem tudi prvi Fickov zakon.

¹ Vektor svetlobnega pretoka $\vec{J}(\vec{r}, t)$ je po definiciji [4, p. 35] $\vec{J}(\vec{r}, t) = \int_{4\pi} I(\vec{r}, \hat{s}) \hat{s} d\omega$ in kaže v smeri neto svetlosti v točki.

2.2. Kubelka-Munk

Kubelka-Munk teorija je preprosta teorija, ki predlaga model širjenja svetlobe na podlagi dveh svetlobnih pretokov (ang. two flux theory) – enega v pozitivni smeri ter drugega v negativni smeri. Pogoja za veljavnost teorije sta plastovita oblika medija ter po celotni površini enaka osvetlitev.

3. NUMERIČNO REŠEVANJE RTE IN DE

Reševanja PDE se lahko lotimo z običajnimi metodami za numerično reševanje diferencialnih enačb. Podroben opis implementacije metod je podan v [5], za njihovo reševanje pa obstaja veliko število odprtokodnih reševalnikov [6].

3.1. Monte Carlo

Pogosto uporabljana metoda za simulacijo širjenja svetlobe v tkivih je stohastična metoda Monte Carlo, ki se z velikim številom ponovitev simulacij približuje rešitvi RTE. Fotoni so glede na vrsto vira osvetlitve temu primerno vstavljeni v tkivo. Nato prepotujejo naključno razdaljo do mesta interakcije s tkivom (porazdelitev vrednosti razdalj je eksponentna), kjer je foton nato absorbiran in sipan tako, kot to določajo optične lastnosti tkiva. Prostor simulacije je razdeljen na poljubno število delov, kamor se shranjujejo deleži absorbiranih fotonov. Iz končne vrednosti v posameznem delu prostora lahko izračunamo svetlobni tok v njem.

4. POSTOPKI ZA MERJENJE OPTIČNIH LASTNOSTI TKIVA

Merjenje optičnih lastnosti tkiva zajema merjenje koeficienta absorpcije μ_a , sipanja μ_s , atenuacije $\mu_t = \mu_a + \mu_s$, fazne funkcije sipanja $p(\cos(\theta))$, anizotropnosti sipanja g , zmanjšane koeficienta sipanja $\mu'_s = \mu_s(1 - g)$ in lomnega količnika n [4, p. 267]. Meritve lahko razdelimo na meritve *ex vivo* ter *in vivo*, kakor tudi na neposredne metode, ki ne zahtevajo modela širjenja svetlobe, ter posredne, pri katerih dobimo optične parametre iz meritev drugih količin na podlagi reševanja inverznega problema širjenja svetlobe.

4.1. Neposredno merjenje optično redkih² vzorcev

Neposredno merjenje optičnih lastnosti je pri optično redkih vzorcih možno z ustrežno postavitvijo vzorca in merilnega instrumenta. Tak način merjenja je zaradi praktičnih omejitev le delno uporaben [4, p. 270].

1. Zaradi zahteve po izločitvi večkratnega sipanja morajo biti vzorci tkiva debeline reda 10 μm , kar je možno pripraviti le z rezanjem zamrznjenega vzorca ali pa s homogenizacijo tkiva, oba postopka pa lahko spremenita optične lastnosti.

² Pri optično redkih vzorcih gre za največ enkratno sipanje svetlobe.

2. Zagotoviti je potrebno gladko površino vzorca (v praksi zelo težko), saj sicer pride do sipanja že na površini, kar vnaša motilni signal.
3. Zaradi tankosti vzorca je uporaben signal relativno majhen ter tako zelo dovzeten za motnje.

4.2. Posredno merjenje optično gostih vzorcev

Meritve lahko razdelimo na več načinov [4, p. 278]:

- glede na uporabljeno geometrijo (meritev odbojnosti ali prepustnosti),
- glede na merjeno veličino (intenziteto, svetlobni tok ali svetlost),
- glede na način, s katerim poizkusimo razlikovati med absorpcijo in sipanjem (prostorska odvisnost (ang. spatially-resolved), spektralna odvisnost (ang. spectrally-resolved), časovna odvisnost (ang. temporally-resolved)).

4.2.1. Prostorsko odvisno merjenje

Meritve odbojnosti v točki se izvajajo na različnih razdaljah od točkastega vira osvetlitve, s čimer dobimo odbojnost v odvisnosti od razdalje $R(r)$. Poizkusimo najti najboljše ujemanje izmerjene krivulje (ang. fitting) z analitično določeno krivuljo, ki jo izpeljemo iz rešitve difuzijske enačbe za polneskončen medij pri lasersko usmerjeni osvetlitvi [4, p. 192]. Nedavna uporaba prostorsko modulirane svetlobe za osvetlitev vzorca je pokazala, da je na ta način mogoče izmeriti optične lastnosti medija v vsaki točki [7] [8]. Vzorec je osvetljen s sinusno modelirano svetlobo z različno prostorsko frekvenco in fazo. Posnetki odbojnosti pri različnih fazah nam omogočajo, da izvedemo demodulacijo odbojnosti [7], različne frekvence osvetlitve pa v praksi določajo različne globine vzorca, iz katerih prihaja informacija, kar nam omogoča izvedbo tomografije.

4.2.2. Spektralno odvisno merjenje

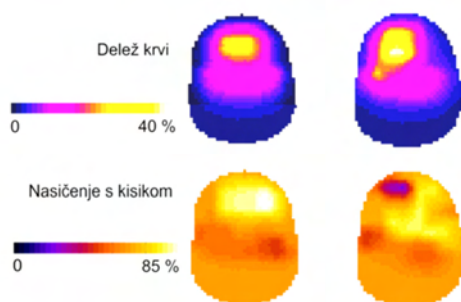
Meritev lahko izvedemo z enim točkastim virom osvetlitve in enim detektorjem. Metoda temelji na predpostavki, da so vsebovane snovi v vzorcu poznane, neznan je le njihova koncentracija. Najdemo lahko kombinacijo absorpcijskih spektrov posameznih snovi, ki jih skupaj z modelom sipanja vstavimo v izpeljano enačbo za odbojnost ter poiščemo najboljše ujemanje z izmerjenim spektrom [4, p. 293].

4.2.3. Časovno odvisno merjenje

Temelji na merjenju odziva odbojnosti v izbrani točki na impulzno osvetlitev. Meritev je ekvivalentna merjenju faznega zamika med virom in izmerjeno vrednostjo [4, p. 295]. Iz izmerjenega odziva lahko prek difuzijske enačbe za polneskončen medij izpeljemo analitični izraz za koeficient sipanja in absorpcije [9].

5. APLIKACIJE

Predstavljena teorija ima široko uporabo v medicinskih aplikacijah. Z optično tomografijo lahko ocenimo oksigenacijo možganov ter s tem sklepamo na notranje krvavitve [10]. Optična tomografija dobiva veljavo tudi pri preučevanju aktivnosti možganov. Njena prednost pred funkcijskim magnetnoresonančnim slikanjem, ki je uveljavljeno na tem področju, je večja hitrost zajemanja slik, cena opreme ter njena prenosljivost, slabost pa predvsem slabša prostorska resolucija [11]. Slabši prostorski resoluciji navkljub potekajo raziskave tudi na področju optične mamografije. Tumorje je možno zaznati, ker je njihova absorpcija od 2 do 4-krat večja od okoljskega tkiva [12].



Slika 3: Primer optične tomografije možganov dveh nedonošenčkov, ki prikazujeta delež krvi ter nasičenje s kisikom. Pri nedonošenčku na desni je vidno zmanjšanje nasičenja s kisikom ter povečanje volumna krvi, kar je posledica notranje krvavitve [11].

6. ZAKLJUČEK

Zmožnost matematičnega opisa širjenja svetlobe v tkivu je pomembna za določitev optičnih lastnosti tkiva, saj direktne metode za merjenje optičnih lastnosti v splošnem niso možne. Univerzalne metode za posredno merjenje optičnih parametrov ni. Raziskave bodo še naprej potekale v iskanju najboljše postavitve senzorjev, načina osvetlitve vzorca ter najprimernejšega postopka za izračun optičnih parametrov iz izmerjenih količin.

LITERATURA

- 1 J. Frangioni, „A Primer on Near-Infrared Imaging Systems for Image-Guided Surgery,“ 2012.
- 2 A. Ishimaru, „Wave Propagation and Scattering in Random Media and Rough Surfaces,“ v *Proceedings of the IEEE*, 1991.
- 3 F. Martelli, S. D. Bianco, A. Ismaelli in G. Zaccanti, Light propagation through biological tissue, Bellingham, Washington: SPIE, 2010.
- 4 A. J. Welch in M. J. v. Gemert, Optical-Thermal Response of Laser-Irradiated Tissue, Springer, 2011.

- 5 D. B. Davidson, *Computational Electromagnetics for RF and Microwave Engineering*, Cambridge: Cambridge University Press, 2005.
- 6 [Elektronski]. Available: http://en.wikipedia.org/wiki/List_of_finite_element_software_packages. [Poskus dostopa 4 2 2013].
- 7 D. J. Cuccia, F. Bevilacqua, A. J. Durkin in B. J. Tromberg, „Modulated imaging: quantitative analysis and tomography of turbid media in the spatial-frequency domain,“ *Optics letters*, zv. 30, št. 11, pp. 1354-1356, 2005.
- 8 T. J. Farrell, M. S. Patterson in a. B. Wilson, „A diffusion theory model of spatially resolved, steady-state diffuse reflectance for the noninvasive determination of tissue optical properties in vivo,“ *Medical Physics*, zv. 19, št. 4, pp. 879-888, 1992.
- 9 M. S. Patterson, B. Chance in B. C. Wilson, „Time resolved reflectance and transmittance for the non-invasive measurement of tissue optical properties,“ *Applied Optics*, zv. 28, št. 12, pp. 2331-2336, 1989.
- 10 J. C. Hebden, „Advances in optical imaging of the newborn infant brain,“ *Psychophysiology*, zv. 40, pp. 501-510, 2003.
- 11 A. P. Gibson, J. C. Hebden in S. R. Arridge, „Recent advances in diffuse optical imaging,“ *PHYSICS IN MEDICINE AND BIOLOGY*, zv. 50, pp. R1-R43, 2005.
- 12 D. G. e. al., „Concentration and oxygen saturation of haemoglobin of 50 breast tumours determined by time-domain optical mammography,“ *PHYSICS IN MEDICINE AND BIOLOGY*, zv. 49, pp. 1165-1181, 2004.

AVTOMATSKA ZAZNAVA ČASOVNO POGOJENIH MORFOLOŠKIH SPREMEMB V MR SLIKAH GLAVE

Žiga Lesjak*, Žiga Špiclin, Boštjan Likar, Franjo Pernuš

Laboratorija za slikovne tehnologije
Fakulteta za elektrotehniko, Univerza v Ljubljani

*E-pošta: ziga.lesjak@fe.uni-lj.si

URL: <http://lit.fe.uni-lj.si>

POVZETEK: *Morfološke spremembe v možganskem tkivu so lahko posledica procesa naravnega staranja ali procesa razvoja bolezni kot so multipla skleroza, možganski tumorji, obstruktivna spalna apneja idr. Za diagnozo in prognozo teh procesov in kot pomoč pri načrtovanju posegov se v klinični praksi najpogosteje uporabljata slikanje z magnetno resonanco (MR) in očna zaznava sprememb med MR slikami, zajetimi ob različnih časih. Očna zaznava morfoloških sprememb na 3D MR slikah je časovno zelo zahtevna, podvržena napakam in subjektivni presoji ocenjevalca. Zaradi večje objektivnosti in ponovljivosti zaznave morfoloških sprememb je vedno večja potreba po avtomatskih postopkih. V članku je podan pregled in razvrstitev avtomatskih postopkov zaznave morfoloških sprememb v MR slikah glave. Razvili smo tudi postopek zaznave morfoloških sprememb s poravnavo in primerjavo sivinske informacije MR slik in ga vrednotili na sintetičnih MR slikah, da bi ocenili pričakovano natančnost avtomatskih postopkov. Zagotavljanje ustrezne natančnosti in zanesljivosti avtomatskih postopkov ter vrednotenje le-teh na realnih MR slikah predstavljata največja raziskovalna izziva.*

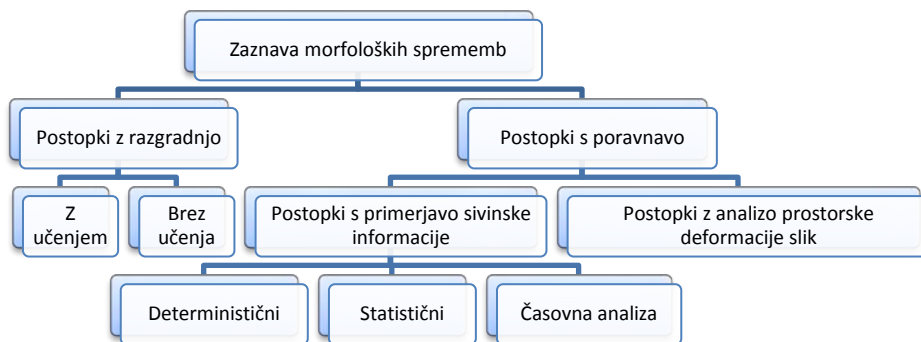
1. UVOD

Slikanje z magnetno resonanco (MR) se zaradi visokega kontrasta med mehкими tkivi najpogosteje uporablja za diagnosticiranje in prognozo pri boleznih, ki vplivajo na pojavnost možganskih tkiv, denimo multipla skleroza (MS), tumorji, obstruktivna spalna apneja (OSA), idr. Pri naštetih boleznih se simptomi med drugim odražajo v spremembah zgradbe in oblike (morfologije) možganskih tkiv, ki jih z visoko občutljivostjo zaznamo z analizo MR slik. Za paciente z MS je značilen pojav brazgotin oziroma lezij pretežno v belem možganskem tkivu in povečana atrofija možganskih tkiv [1]. Za tumorje je značilna relativno hitra rast, pri pacientih z OSA pa se v možganskem tkivu pojavljajo ishemične lezije. Ustrezna zaznava in kvantitativna analiza teh morfoloških sprememb v možganskem tkivu v času sta ključni za razumevanje nastanka in poteka bolezni, predvsem pa za vrednotenje stanja in oceno potencialno škodljivega vpliva na pacienta in kot pomoč pri odločanju ter načrtovanju posegov.

V klinični praksi prevladuje očna zaznava morfoloških sprememb s primerjavo dveh MR slik, zajetih v dveh zaporednih pregledih z ustreznim časovnim razmakom. Očna zaznava je časovno zelo zahtevna, podvržena napakam in subjektivni presoji ocenjevalca. Raziskave vpliva presoje ocenjevalca na izmerjene spremembe velikosti lezij MS kažejo na veliko variabilnost ocene posameznega ocenjevalca (ang. *intra-rater variability*) in veliko variabilnost ocene med različnimi ocenjevalci (ang. *inter-rater variability*) [2]. Zaradi potrebe po objektivnosti in večje ponovljivosti in učinkovitosti je potreba po avtomatskih postopkih zaznave morfoloških sprememb vedno večja [3].

2. PREGLED AVTOMATSKIH POSTOPKOV

Avtomske postopke zaznave morfoloških sprememb zaporedno v času zajetih MR slik glave v grobem delimo na: 1) postopke z razgradnjo MR slik glave na osnovne strukture in primerjavo morfologije teh struktur in 2) postopke s poravnavo MR slik in zaznavo sprememb med slikami. Slika 1 prikazuje podrobnejšo delitev avtomatskih postopkov za zaznavo morfoloških sprememb [4].



Slika 1: Delitev postopkov za avtomatsko zaznavo morfoloških sprememb

1.2 Postopki z razgradnjo MR slik

Zaznava morfoloških sprememb z razgradnjo MR slik temelji na primerjavi morfologije razgrajenih področij. Poleg področij zanimanja s temi postopki razgradimo tudi ostala možganska področja, s katerimi lahko ovrednotimo celostne morfološke spremembe, denimo atrofijo možganskega tkiva. Glede na število nedavnih objav [4] tovrstni postopki predstavljajo najboljše in najaktivnejše področje raziskovanja za problem zaznave morfoloških sprememb, predvsem za MR slike pacientov z MS.

Avtomske postopke razgradnje slik ločimo na postopke z učenjem in postopke brez učenja. Postopki razgradnje z učenjem potrebujejo učne (ali vzorčne) primere razgrajenih MR slik s specifično patologijo (lezije MS, tumorji, ishemične lezije, idr.). Razvrščanje možganskih struktur temelji na učnih primerih, zato mora biti v učni zbirki MR slik zajetih čim več normalnih kot tudi patoloških primerov. Med najpogosteje

uporabljenimi postopki razgradnje z učenjem so: metoda *k-najbližjih sosedov* (kNN), *umetne nevronske mreže* (ANN) in *metoda podpornih vektorjev* (SVM). Glavna slabost postopkov razgradnje z učenjem je, da natančnost zaznanih morfoloških sprememb zavisi od reprezentativnosti učnih slik.

Postopki z razgradnjo MR slik brez učenja temeljijo na specifičnih modelih možganskih tkiv. Parametri teh modelov se določijo na podlagi posamezne MR slike, običajno z različico algoritma *expectation maximization* (EM). Z modelom možganskih tkiv je dana razgradnja možganskih tkiv (normalnih in patoloških). Razni artefakti MR slik (npr. šum in sivinska nehomogenost, šum) lahko kršijo predpostavke v modelu možganskih tkiv in posledično negativno vplivajo na kvaliteto razgradnje MR slik.

S poravnavo anatomskega atlasa na dane MR slike lahko v postopek razgradnje MR slik vključimo *a priori* znanje o razporeditvi normalnih tkiv in na ta način izboljšamo natančnost razgradnje. Natančnost razgradnje lahko izboljšamo tudi z analizo večih MR sekvenc, ki v proces razgradnje vnašajo komplementarne informacije o posameznem tkivu (npr. T1-, T2- in PD-utežene in FLAIR sekvence).

2.2 Postopki s poravnavo MR slik

Morfološke spremembe lahko zaznamo preko poravnave MR slik ter s primerjavo sivinske informacije ali prostorske deformacije MR slik. Postopki s primerjavo sivinske informacije temeljijo na neposredni primerjavi sivin med istoležečimi slikovnimi elementi dveh MR slik. Ti postopki lahko zaznajo le spremembe med MR slikama, ne pa tudi ostalih struktur v MR slikah. Za grobo prostorsko poravnavo se običajno uporablja afina preslikava, za fino pa preslikava z B-zlepki [5]. Zaznavo morfoloških sprememb s primerjavo sivinske informacije lahko izvedemo 1) z deterministično analizo razlik med sivinami istoležečih slikovnih elementov dveh MR slik, 2) s statistično analizo sivinske informacije v podregijah MR slike in 3) s časovno analizo potek sivinske informacije v posameznih istoležečih slikovnih elementov v večih zaporedno zajetih MR slikah. Ključen problem pri postopkih s primerjavo sivinske informacije je variabilnost sivinske informacije MR slik, zato je potrebna ustrežna svetlostna normalizacija MR slik.

Spremembe v sivinski informaciji dveh MR slik so lahko tudi posledica spremembe molekularne strukture tkiva in ne le posledica spremembe morfologije (npr. volumna) posameznih tkiv. Postopki z analizo prostorske deformacije slik izrabljajo sivinsko informacijo le za določanje prostorske deformacije med dvema MR slikama, ki običajno temelji na nelinearni poravnavi. Prostorske deformacije so običajno določene z »gosto« neparometrično preslikavo (npr. *demons based registration* in *diffusion registration*), ki vsakemu slikovnem elementu priredi preslikavo v referenčno MR sliko. Na osnovi prostorskih deformacij med dvema MR slikama lahko hkrati izvedemo zaznavo in volumetrično analizo morfoloških sprememb. Najpogosteje se v ta namen uporablja Jakobijev operator [6], ki je relativno neobčutljiv na napake predhodne poravnave, odziv Jakobijevega operatorja pa je premosorazmeren s spremembo volumna v posameznem slikovnem elementu. Nedavno so Pieperhoff *et al.* [7] razvili postopek, ki temelji na razmerju med originalnimi in preslikanimi lokalnimi volumni slikovnih elementov (ang. *local volume ratio* - LVR). LVR je v primerjavi z Jakobijevim operatorjem manj občutljiv na šum, izognemo pa se tudi numeričnim napakam pri izračunu odvodov slike.

Zaradi visoke prostostne stopnje »goste« preslikave je ključen problem postopkov z analizo prostorske deformacije slik (ne)zanesljivost poravnave slik.

3. VREDNOTENJE POSTOPKA AVTOMATSKE ZAZNAVE

Razvili smo postopek zaznave morfoloških sprememb s poravnavo MR slik in primerjavo sivinske informacije in ga vrednotili na sintetičnih MR slikah, da bi ocenili pričakovano natančnost avtomatskih postopkov. Postopek za zaznavo morfoloških sprememb uporablja dve MR-T2 sliki zajeti ob različnih časovnih trenutkih in se izvede v treh zaporednih korakih: 1) sivinska normalizacija MR slik, 2) poravnava MR slik in 3) zaznava morfoloških sprememb na osnovi slike razlik.

Sivinska normalizacija zajema obnovo sivinskih nehomogenosti s postopkom minimizacije entropije slik [8] in prileganjem decilov histogramov MR slik [9] za izenačitev dinamičnih območij sivinskih vrednosti MR slik. V drugem koraku MR sliki poravnava z afino preslikavo, nato pa še s preslikavo z B-zlepki. V tretjem koraku poravnani MR sliki med seboj odštejemo, morfološke spremembe pa zaznamo z upragovljanjem slike razlik. Prag na sliki razlik določimo avtomatsko preko intervala zaupanja $\alpha=0.01$, pri čemer predpostavimo da so razlike normalno porazdeljene s srednjo vrednostjo 0. Slikovni elementi z razliko MR slik zunaj intervala zaupanja predstavljajo zaznane morfološke spremembe.

Predlagani postopek smo ovrednotili na treh sintetičnih MR-T2 slikah iz javno dostopne zbirke BrainWeb [10]. Da bi bili parametri simuliranih MR slik čim bolj podobni realnim MR slikam, ki so dostopne v klinični praksi smo uporabili MR-T2 slike z aksialno resolucijo 3 mm in 20% stopnjo sivinske nehomogenosti. MR-T2 slike so vsebovale lezije različnih velikosti, z eksperimentom pa smo želeli določiti natančnost in občutljivost zaznave teh lezij. Postopek zaznave morfoloških sprememb smo zagnali med vsemi tremi pari MR-T2 slik. Za vrednotenje postopka smo kot mero sposobnosti uporabili Dice-ov koeficient [11]

$$\frac{2 \times |TP|}{2 \times |TP| + |FP| + |FN|} \quad (1)$$

in občutljivost

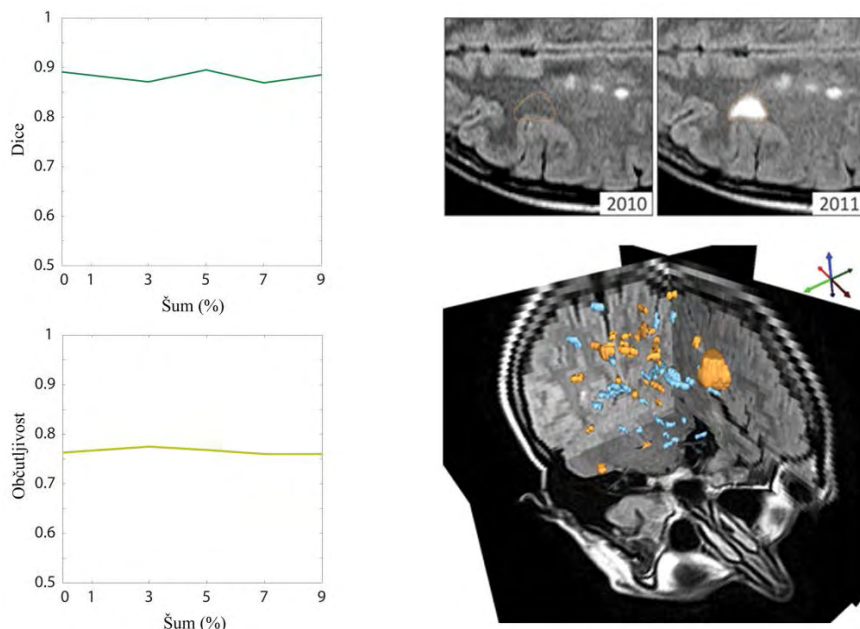
$$\frac{|TP|}{|TP| + |FN|} \quad (2)$$

kjer sta TP in FP deleža dejanskih in lažnih, FN pa delež nezaznanih sprememb.

Povprečne vrednosti mer sposobnosti in njihova odvisnost od šuma v MR-T2 slikah so prikazani v Tabeli 1 in na Sliki 2. Če primerjamo dobljene vrednosti s trenutno najboljšo razvito metodo [4], [12], ki dosega Dice-ov koeficient 0.81 in občutljivost 0.971 lahko sklenemo, da predlagani postopek dosega višjo povprečno stopnjo prekrivanja področja zaznanih in dejanskih morfoloških sprememb. Predlagani postopek ima sicer nižjo občutljivost, ki pa jo lahko povečamo z manjšanjem parametra α .

Tabela 1: Mere sposobnosti v odvisnosti od šuma

Šum (%)	0	3	5	7	9
Dice-ov koeficient	0,89	0,87	0,90	0,87	0,89
Občutljivost	0,76	0,78	0,77	0,76	0,76



Slika 2: Mere sposobnosti v odvisnosti od stopnje šuma v MR slikah (levo) in primer zaznanih morfoloških sprememb (desno)

4. ZAKLJUČEK

V članku je podan pregled in razvrstitev avtomatskih postopkov zaznave časovno pogojenih morfoloških sprememb v MR slikah glave. Razvili smo tudi postopek zaznave morfoloških sprememb s poravnavo in primerjavo sivinske informacije MR slik in ga vrednotili na sintetičnih MR slikah, da bi ocenili pričakovano natančnost avtomatskih postopkov. Ugotovili smo relativno visoko stopnjo prekrivanja (povprečen Dice-ov koeficient 0,88) med zaznanimi in dejanskimi morfološkimi spremembami.

Zagotavljanje ustrezne natančnosti in zanesljivosti avtomatskih postopkov ter vrednotenje le-teh na realnih MR slikah predstavljata največja raziskovalna izziva. Natančna in zanesljiva prostorska poravnava je ključnega pomena za robustno in natančno zaznavo morfoloških sprememb s predlaganim postopkom. Postopek je robusten na širok razpon šuma (Slika 2). Pričakujemo, da bo ključen problem pri uporabi predlaganega postopka na realnih slikah predvsem variabilnost sivinske informacije kot posledica sivinskih nehomogenosti MR slik, zato načrtujemo izboljšave svetlostne

normalizacije MR slik. Z dodatnimi MR sekvencami (MR-T1 in FLAIR) bomo v avtomatski postopek vnesli komplementarno informacijo, kar bi lahko povečalo občutljivost zaznave morfoloških sprememb. Nadaljnje potencialne izboljšave bi lahko dosegli s kombiniranjem postopkov razgradnje in poravnave slik v skupen okvir ter tako izkoristili prednosti in slabosti enih in drugih postopkov.

Objektivna in kvantitativna primerjava uveljavljenih in novih postopkov zaznave morfoloških sprememb je zelo težavna zaradi pomanjkanja reprezentativne zbirke MR slik z različnimi patologijami, z različnimi stopnjami patologije in z normalnimi, zdravimi primeri ter zaradi pomanjkanja referenčne razgradnje oziroma *zlatega standarda* morfoloških sprememb. Eden od ciljev nadaljnjih raziskav je zato izgradnja ustrežne in reprezentativne zbirke MR slik, predvidoma s pacienti z multiplo sklerozo in s pacienti z obstruktivno spalno apnejo.

5. LITERATURA

- [1] W. I. McDonald, A. Compston, G. Edan, D. Goodkin, H.-P. Hartung, F. D. Lublin, H. F. McFarland, D. W. Paty, C. H. Polman, S. C. Reingold, M. Sandberg-Wollheim, W. Sibley, A. Thompson, S. Van Den Noort, B. Y. Weinschenker, and J. S. Wolinsky, "Recommended diagnostic criteria for multiple sclerosis: Guidelines from the international panel on the diagnosis of multiple sclerosis," *Annals of Neurology*, vol. 50, no. 1, pp. 121–127, 2001.
- [2] I. L. Tan, R. A. van Schijndel, F. Fazekas, M. Filippi, P. Freitag, D. H. Miller, T. A. Yousry, P. J. W. Pouwels, H. J. Adèr, and F. Barkhof, "Image registration and subtraction to detect active T2 lesions in MS: an interobserver study," *Journal of Neurology*, vol. 249, no. 6, pp. 767–773, 2002.
- [3] M. Bosc, F. Heitz, J.-P. Armspach, I. Namer, D. Gounot, and L. Rumbach, "Automatic change detection in multimodal serial MRI: application to multiple sclerosis lesion evolution," *NeuroImage*, vol. 20, no. 2, pp. 643–656, Oct. 2003.
- [4] X. Lladó, O. Ganiler, A. Oliver, R. Martí, J. Freixenet, L. Valls, J. Vilanova, L. Ramió-Torrentà, and À. Rovira, "Automated detection of multiple sclerosis lesions in serial brain MRI," *Neuroradiology*, vol. 54, no. 8, pp. 787–807, 2012.
- [5] D. Rueckert, P. Aljabar, R. A. Heckemann, J. V. Hajnal, and A. Hammers, "Diffeomorphic Registration Using B-Splines," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2006*, R. Larsen, M. Nielsen, and J. Sporring, Eds. Springer Berlin Heidelberg, 2006, pp. 702–709.
- [6] D. Rey, G. Subsol, H. Delingette, and N. Ayache, "Automatic detection and segmentation of evolving processes in 3D medical images: Application to multiple sclerosis," *Medical Image Analysis*, vol. 6, no. 2, pp. 163–179, Jun. 2002.
- [7] P. Pieperhoff, M. Südmeyer, L. Hömke, K. Zilles, A. Schnitzler, and K. Amunts, "Detection of structural changes of the human brain in longitudinally acquired MR images by deformation field morphometry: Methodological analysis, validation and application," *NeuroImage*, vol. 43, no. 2, pp. 269–287, Nov. 2008.
- [8] U. Vovk, F. Pernuš, and B. Likar, "Intensity inhomogeneity correction of multispectral MR images," *NeuroImage*, vol. 32, no. 1, pp. 54–61, Aug. 2006.

- [9] J. K. Udupa, L. Wei, S. Samarasekera, Y. Miki, M. A. vanBuchem, and R. I. Grossman, "Multiple sclerosis lesion quantification using fuzzy-connectedness principles," *IEEE Trans. Med. Imaging*, vol. 16, no. 5, pp. 598–609, Oct. 1997.
- [10] C. A. Cocosco, V. Kollokian, R. K.-S. Kwan, G. B. Pike, and A. C. Evans, "BrainWeb: Online Interface to a 3D MRI Simulated Brain Database," *NeuroImage*, vol. 5, p. 425, 1997.
- [11] L. R. Dice, "Measures of the Amount of Ecologic Association Between Species," *Ecology*, vol. 26, no. 3, pp. 297–302, Jul. 1945.
- [12] P. Anbeek, K. Vincken, M. Vanosch, R. Bisschops, and J. Vandergrond, "Automatic segmentation of different-sized white matter lesions by voxel probability estimation," *Medical Image Analysis*, vol. 8, no. 3, pp. 205–215, Sep. 2004.

RAZGRADNJA MEDICINSKIH SLIK HRBTENICE

Robert Korez, Boštjan Likar, Franjo Pernuš, Tomaž Vrtovec

Laboratorij za slikovne tehnologije
Fakulteta za elektrotehniko, Univerza v Ljubljani
E-pošta: robert.korez@fe.uni-lj.si
URL: <http://lit.fe.uni-lj.si>

POVZETEK: *Natančna in zanesljiva razgradnja medicinskih slik hrbtenice je pomembna v klinični diagnostiki in zdravljenju poškodb oz. bolezenskih stanj hrbtenice, kot so npr. bolečina v ledvenem predelu hrbtenice, kila medvretenčne ploščice¹, spinalna stenoza², nestabilnost³, osteoporoz⁴ idr. V prispevku podamo kronološki pregled nekaterih postopkov razgradnje medicinskih slik hrbtenice, zajetih z računalniško tomografijo (CT) ali magnetno resonanco (MR).*

1. UVOD

Hrbtenica je najpoglavitejši in najbolj komplicirani del človeškega skeleta. Njen razpon sega od lobanjskega dna do medenice, povezuje glavo in ude, obdaja in varuje hrbtenjačo ter je nosilni stolp celega telesa. Zgradba hrbtenice je izredna, saj opravlja vrsto funkcij in vzdrži najrazličnejše pritiske in natege. Sestavljena je iz 33 vretenc, ki so med seboj povezana z močnimi vezmi, medvretenčnimi ploščicami in sklepi. Samo vretenca je sestavljeno iz telesa (tj. sprednji del v obliki debele koščene plošče), vretenčnega loka (ki obdaja vretenčno odprtino ter je s telesom povezan z pedikloma; vretenčne odprtine tvorijo hrbtenični kanal, v katerem se nahaja hrbtenjača), navzdol štrlečega trnastega odrastka ter dveh prečnih (ki štrlijo vstran) in štirih sklepnih (ki paroma štrlijo navzgor in navzdol) odrastkov. Funkcionalno hrbtenico delimo na zgornji (C1-C2) in spodnji (C3-C7) vratni del ter na prsni (T1-T12) in ledveni (L1-L5) del. Razvojno pa k hrbtenici spadata še križnica (S1-S5) in trtica (Co1-Co4)⁵, ki sta sestavljeni iz med seboj zraslih vretenc. Gledano s strani ima hrbtenica obliko dvojne črke S. Prsni in križnično-trtični del sta uleknjena nazaj (primarni krivini), vratni in ledveni del pa sta upognjena nekoliko naprej (vratna, ledvena krivina). Krivine se oblikujejo že pri dojenčku in dajejo hrbtenici večjo prožnost, v starosti pa postajajo bolj izrazitejše.

¹ kila medvretenčne ploščice – zdrs dela mehke sredice skozi razpoko čvrste zunanje ovojnice

² spinalna stenoza – zožitev hrbteničnega kanala, v katerem se nahaja hrbtenjača

³ nestabilnost – nezmožnost ohranjanja oblike hrbtenice pod fizičnimi obremenitvami

⁴ osteoporoz – zmanjševanje kostne mase oz. gostote kosti

⁵ C1-Co4 – oznake vretenc



Slika 1: Različni pogledi na hrbtenico (levo) in prsno vretenca (desno).

Sodobne tehnike zajema medicinskih slik omogočajo kakovosten vpogled v celotno anatomijo hrbtenice. Tehnika CT je primerna za opazovanje hrbteničnih kosti, tehnika MR pa omogoča vpogled v mehka tkiva in je tako primerna za opazovanje medvretenčnih ploščic, hrbtenjače idr. Razvoj natančnih in zanesljivih tehnik razgradnje medicinskih slik lahko torej prispeva k natančnejši diagnostiki ter načrtovanju učinkovitejših strategij zdravljenja poškodb oz. bolezenskih stanj hrbtenice.

2. POSTOPKI RAZGRADNJE

Podan je kronološki pregled nekaterih postopkov razgradnje, ki se razlikujejo predvsem po načinu uporabe predhodnega anatomskega znanja. V grobem lahko postopke razgradnje slik delimo v 2 skupini:

- i. Razgradnja, ki je povezana z deformabilnimi modeli (tj. modeli površin, ki jim spreminjamo prostorske lastnosti in teksturo, da se kar najbolj prilegajo objektom zanimanja na sliki) – Mastmeyer in sod. (2006), Ma in sod. (2010), Štern in sod. (2011).
- ii. Razgradnja, ki ni povezana z deformabilnimi modeli (upravljanje, morfološke operacije, rast področij idr.) – Hoad in Martel (2002), Carballido-Gamio in sod. (2004), Peng in sod. (2005).

2.1 Hoad in Martel (2002)

Predhodna obdelava MR slik: S kombiniranjem T1-uteženih MR slik, zajetih pri različnih časih (2.5 ms, 4.9 ms) spinskega odmeva T_E , lahko dosežemo, da bo na novih (kombiniranih) slikah intenziteta kosti nizka, intenziteta okoliškega mehkega tkiva pa visoka.

Razgradnja vretenčnih teles: Ker je samo z enostavnim upravljanjem nemogoče razgraditi vretenčne kosti, zato določimo meje, v katerih ležijo vretenčna telesa. V vsakem prečnem prerezu, v katerem je vidno vretenčno telo določimo 4 točke. Te bodo

definirale elipso, ki se bo najboljše prilegala sprednji ploskvi vretenčnega telesa. Kljub temu, da se elipsa ne bo natančno prilegala zadajšnji ploskvi, bo le-ta v celoti zajeta. Z interpolacijo nato združimo niz elips, ki bodo definirale masko M_1 . Z upravljanjem znotraj maske M_1 lahko določimo masko M_2 :

$$M_2(x, y, z) = \begin{cases} 1 \cdot M_1(x, y, z), & S(x, y, z) \leq T; \\ 0, & \text{sicer,} \end{cases} \quad (1)$$

kjer je S intenziteta slikovnega elementa. Prag T določimo tako, da zajamemo čim več kosti. Znotraj maske M_2 nato z filtriranjem zapolnimo luknje.

Razgradnja zadajšnjih struktur (odrastki): Z vpeljavo maske M_3 odstranimo vretenčna telesa iz upravljenе slike:

$$M_3(x, y, z) = \begin{cases} 1 \cdot (1 - M_1(x, y, z)), & S(x, y, z) \leq T; \\ 0, & \text{sicer.} \end{cases} \quad (2)$$

Vrednosti slikovnih elementov v maski M_3 , ki ležijo pred sprednjo stranjo vretenčne ploskve postavimo na nič. Preostanek slike pa nato obdelamo z metodo Rasti področij. Najprej s filtriranjem določimo začetne vrednosti semen (slikovnih elementov) za rast. Rezultat so semena blizu vretenčnih sklepov in semena v šumu izven telesa, katera pa avtomatsko odstranimo. Z rastjo področja, ki ga določajo semena blizu razgradimo zadajšnje strukture. Luknje ponovno zapolnimo s filtriranjem.



Slika 2: Obdelana slika (levo), enostavno upravljanje (sredina) in razgrajena slika (desno).

2.2 Carballido-Gamio in sod. (2004)

Predstavili bomo razgradnjo vretenčnih teles na izbranem stranskem prerezu T1-utežene MR slike z uporabo metode Normaliziranega reza (problem razgradnje prevedemo na problem delitve grafa). Za začetek na kratko opišimo to metodo:

- i. Na 2D sliko gledamo kot na graf $G=G(V,E)$, kjer je V množica vozlišč (slikovni elementi) in E množica obteženih povezav z utežjo $w_{ij} = e^{-\frac{d(i,j)}{\sigma}}$, kjer je $d(i,j)$ mera podobnosti (npr. intenziteta, razdalja) med vozliščema i in j .
- ii. Rešujemo posplošeni lastni problem $(D-W)x=\lambda Dx$, kjer je $W(i,j)=w_{ij}$ simetrična in $D(i,i) = \sum_j w(i,j)$ diagonalna matrika.

- iii. Lastni vektor (vsakemu njegovemu elementu pripada vozlišče grafa), ki pripada 2. najmanjši lastni vrednosti, uporabimo za delitev grafa oz. za razgradnjo slike, z delitvijo lastnega vektorja namreč določimo tudi delitev grafa.

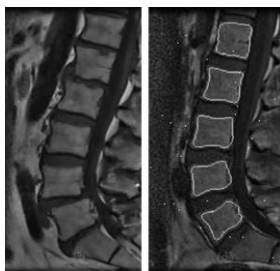
Reševanje posplošenega lastnega problema je lahko računsko precej zahtevno, saj imamo opravka z velikimi slikami (npr. 512 pikslov \times 512 pikslov, kar pomeni reševanje lastnega problema matrik velikosti $512^2 \times 512^2$). Temu se lahko izognemo tako, da ta zahteven problem prevedemo na reševanje kakšnega enostavnejšega (npr. z uporabo Nyström aproksimativne metode).

Predhodna obdelava MR slik: Izbrano T1-uteženo MR sliko, ki jo bomo razgradili, najprej obdelamo z različnimi metodami (npr. uprakovljanje, filtriranje). Nato izberemo tisti stranski prerez, ki najbolje prikazuje hrbtenični kanal.

Razgradnja vretenčnih teles: Kot smo že omenili, sliko razgradimo s pomočjo metode Normaliziranega reza (in uporabo Nyström aproksimacije). Pri tem pa je potrebno določiti dobro mero podobnosti. In sicer, za izbrani slikovni element (piksel oz. voksel) P stranskega prereza, izračunamo t. i. 3D histogram, tj. izberemo izotropično prostornino $x \times x \times z$ vokselov s centrom v P in izračunamo histogram normaliziranih intenzitet $x^2 \cdot z$ vokselov, ki je razdeljen na n prekatov. Mera podobnosti je tedaj enaka

$$d(i, j) = \frac{1}{2} \sum_{k=1}^n \frac{(h_i(k) - h_j(k))^2}{h_i(k) + h_j(k)}, \quad (3)$$

kjer je $h_i(k)$ število vokselov z intenziteto znotraj k -tega prekata 3D histograma voksla i . Sedaj lahko definiramo in rešimo posplošen lastni problem. Izbrana rešitev razgradnje bo lastni vektor, ki mu pripada 2. najmanjša lastna vrednost. Slednjega z metodo K -povprečij razdelimo na 3 dele (tj. 3 maske). Z maskiranjem sliko razdelimo na 3 dele, znotraj nje ročno izberemo 2 točki in z metodo 4-povezovalnega označevanja (potek metode v članku ni opisan) razgradimo vretenčna telesa.



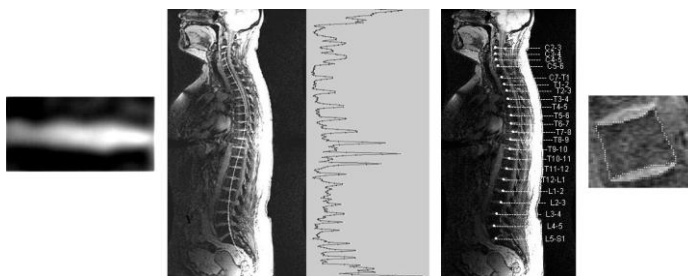
Slika 3: Obdelana (levo) in razgrajena (desno) slika.

2.3 Peng in sod. (2005)

Izbira najustreznejšega stranskega prereza MR slike: Na enem izmed stranskih prerezov vhodne MR slike celotne hrbtenice izberemo model medvretenčne ploščice velikosti 20×40 pikslov in ga sprehajamo po vsakem izmed prerezov ter pri tem iščemo najboljše

razmerje intenzitet (med modelom in območjem na katerem se model nahaja). Tako za vsak prerez izluščimo nabor koordinat (x_i, y_i) , ki predstavljajo središča modela v položaju, kjer je razmerje intenzitet nad predpisano vrednostjo. Skozi koordinate z interpolacijo določimo krivuljo, vzdolž te pa izračunamo intenzitete. Pri tem lokalni maksimumi predstavljajo središča medvretenčnih ploščic, lokalni minimumi pa vretenčna telesa. Opazimo, da razdalja med lokalnimi maksimumi približno predstavlja višino vretenčnih teles in tako je (čim manjša) varianca višin kriterij za izbiro najustrežnejšega stranskega prereza.

Številčenje medvretenčnih ploščic in razgradnja vretenc: Številčenje medvretenčnih ploščic v izbranem prerezu poteka tako, da se premikamo v obeh smereh vzdolž interpolirane krivulje ter z uporabo informacije o približni obliki in velikosti ploščic (in vretenc) natančneje določamo ter obenem številčimo njihova središča. Na koncu z Cannyjevim detektorjem določimo robove vretenc, odstrani napačne ter popravimo slabo definirane (natančen opis v članku ni podan).



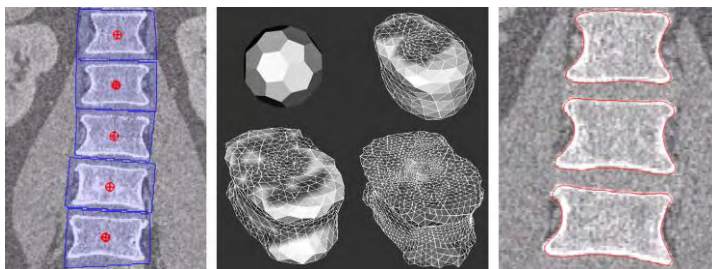
Slika 4: Od leve proti desni: Model medvretenčne ploščice, graf intenzitet vzdolž interpolirane krivulje, številčenje ploščic in razgradnja vretenc.

2.4 Mastmeyer in sod. (2006)

Ločitev vretenčnih teles: V izbranem stranskem prerezu CT slike ročno določimo točke M_i , ki predstavljajo središča vretenčnih teles. Na podlagi določenih točk zgradimo krivuljo, ki aproksimira t. i. povprečno os vretenčnih teles. V ravninah pravokotnih na krivuljo, analiziramo histogram tistega dela ravnine, ki se približno ujema s projekcijo vretenčnega telesa. Glede na analizo obdržimo tiste ravnine, ki gredo skozi središča medvretenčnih ploščic. Na koncu vzdolž krivulje definiramo eliptične cilindre (ločene z pravkar določenimi ravninami), katerih male in velike polosi bodo odvisne od velikosti posameznih vretenčnih teles. S tem je tako podana ločitev vretenčnih teles.

Razgradnja vretenčnih teles: V določeno središče vretenčnega telesa postavimo okrnjen ikozaeder (nogometna žoga) z najmanj 60 vozlišči, katerega raztezamo z minimiziranjem notranje in zunanje energije. Pri tem notranje sile določimo s povprečenjem vozlišč v njihovi okolici, zunanje sile pa so določene z intenzitetnimi profili vzdolž žarkov (z izvorom v središču ikozaedra) skozi vozlišča do roba eliptičnega cilindra. Če povzamemo, okrnjen ikozaeder s pomočjo sil iterativno raztezamo proti kostnini

vretenca. Lahko se zgodi, da razgradnja zajame tudi druge vretenčne strukture, ki pa jih odstranimo z erozijo.



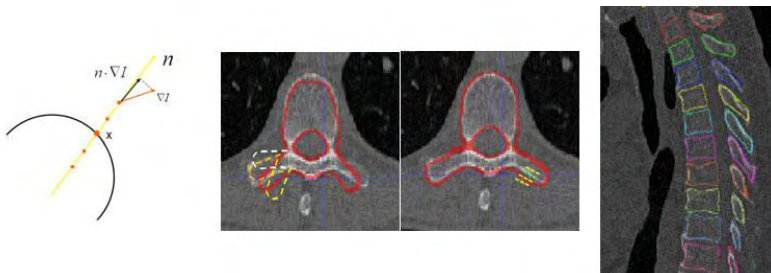
Slika 5: Ločitev vretenčnih teles (levo), raztezanje ikozaedra (sredina) in razgradnja (desno).

2.5 Ma in sod. (2010)

Učenje detektorja robov: Ročno razgradimo vsa prsna vretenca v vsaki izmed 20 CT slik (učna množica) in z metodo Sprehajajočih kock (ang. Marching cubes) konstruiramo pripadajoče triangulirane površine. Vsak model vretenca je sestavljen iz okoli 10^4 lic. Konstrukcija vektorja značilnic: Za poljubno lice izberemo 5 vzorčnih točk (vzorčna parcela) vzdolž normale v oddaljenosti 1 slikovni element. Natančneje, če je x točka na normalni premici in n enotski vektor normale na izbrano lice, potem je vzorčna parcela definirana kot $P(x) = \{x - 2n, x - n, x, x + n, x + 2n\}$. Nadalje, za vsako točko vzorčne parcele definiramo 3 značilnice: intenziteto ter projekciji gradientov (z različnima velikostma jeder) na normalo. S tem je definiran vektor značilnic vzorčne parcele, ki ga označimo z $F(x)$. Učenje detektorja robov: Za vsako težišče c poljubnega lica, definiramo mejno $P(c)$, notranjo $P(c - 3n)$ in zunanjo $P(c + 3n)$ vzorčno parcelo. Poleg tega izračunamo tudi pripadajoče vektorje značilnic: $F(c)$, $F(c - 3n)$, $F(c + 3n)$. Tako vsako lice določa 1 »pozitiven« (rob vretenca) in 2 »negativna« vektorja značilnic (notranjost, zunanost). Ker ima vsak model vretenca 10^4 lic, je s tem zagotovljena zadostna in ustrezna učna množica vektorjev značilnic. S Kvadratno diskriminantno analizo (ang. Quadratic discriminant analysis) izdelamo razvrščevalca, ki bo danemu slikovnemu elementu x in pripadajočemu vektorju $F(x)$ priredil vrednost $L(x) \in [0, 1]$, tj. verjetnost, da je slikovni element x robna točka vretenca.

Grobo-fina razgradnja: Osnovna ideja razgradnje je prilagajanje modela k robnim točkam, ki jih najdemo s pomočjo naučenega detektorja robov. Ko je model vstavljen v CT sliko (model je vstavljen tako, da se čim bolj prilega vretencu), detektor robov izračuna $L(x)$ za vsak slikovni element x vzdolž normal vseh lic. Da ohranimo topologijo in gladkost modela med deformacijo, se razgradnje lotimo postopoma. Model vretenca najprej razdelimo na 12 področij (vretenčno telo, trnasti odrastek, levi stranski odrastek, desni stranski odrastek idr.). Da med deformacijo ohranjamo njihovo obliko, vpeljemo podobnostno transformacijo tako, da je odziv detektorja robov po transformaciji maksimalen. Ko najdemo ustrezno transformacijo za vsako izmed področij, lahko s pomočjo Gaussovega filtriranja dobimo deformacijo celotnega modela. Nato pa se

lotimo še finejše deformacije. Vsak model vretenca razdelimo na približno 200 področij (krp), ki so vsaka sestavljena iz približno 50 lic. Vsako izmed krp premikamo vzdolž njenega povprečja normal, dokler ne dobimo največjega odziva detektorja. Ko določimo vse premike ponovno s pomočjo Gaussovega filtriranja določimo deformacijo celotnega modela.

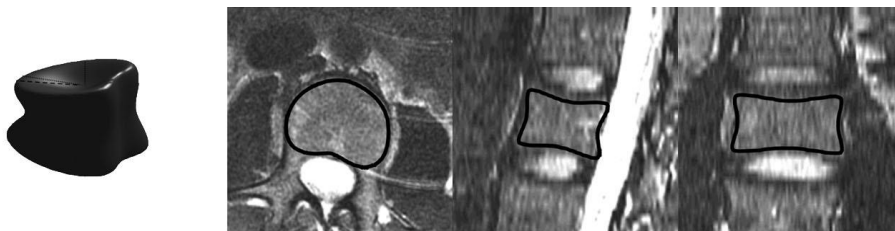


Slika 6: Vektor značilnic (levo), prilagajanje večjih in manjših področij (sredina) ter razgradnja (desno).

2.6 Štern in sod. (2011)

Modeliranje oblike in pozicije vretenčnega telesa: Osnovna oblika vretenčnega telesa je predstavljena s 3 parametri, ki določajo velikost eliptičnega cilindra (oz. superkvadrka). Prečni prerez cilindra je deformiran z 12 parametri, ki modelirajo odklon telesa od eliptične oblike v sprednjem delu ter na lokacijah levega in desnega pedikla ter vretenčne odprtine. Polmer prečnega prereza je nadalje deformiran z 2 parametroma, ki modelirata konkavnost sten telesa v sprednjem delu in na lokaciji vretenčne odprtine. Konkavnost in odklon zgornjih in spodnjih plošč telesa pa modeliramo s 6 parametri. Končno z 2 parametroma opišemo še naraščanje velikosti in torzijo. Za umestitev v sliko pa potrebujemo še 6 parametrov, ki opisujejo translacije in rotacije. Če povzamemo, obliko in pozicijo vretenčnega telesa opisuje skupaj $25+6=31$ parametrov.

Deformacija in poravnava modela: Deformacija in poravnava začetnega modela (eliptični cilindri) v izbrano vretenčno telo v CT ali MR sliki je dosežena z iskanjem maksimuma mere podobnosti (ki kombinira informaciji o intenziteti in gradientu). Maksimum torej določa optimalne vrednosti 31 parametrov, ki najbolje opisujejo obliko in pozicijo izbranega vretenčnega telesa v sliki in s tem tudi njegovo razgradnjo.



Slika 7: Parametrični model vretenčnega telesa (levo) in razgradnja (desno).

3. ZAKLJUČEK

V prispevku je predstavljenih 6 postopkov razgradnje medicinskih slik hrbtenice, po trije iz vsake skupine. Tisti, ki temeljijo na deformabilnih modelih, omogočajo poleg razgradnje tudi kvantitativno vrednotenje deformacij hrbtencičnih struktur, ki so posledice poškodb oz. bolezenskih stanj hrbtenice.

LITERATURA

1. J. Carballido-Gamio, S. J. Belongie, S. Majumdar (2004), Normalized cuts in 3D for spinal MRI segmentation, *IEEE Transactions on Medical Imaging*, vol. 23, str. 36-44.
2. J. W. M. Goethem, L. van den Hauwe, P. M. Parizel, A. L. Baert (2007), *Spinal Imaging: Diagnostic Imaging of the Spine and Spinal Cord*, Berlin: Springer.
3. C. L. Hoad, A. L. Martel (2002), Segmentation of MR images for computer-assisted surgery of the lumbar spine, *Physics in Medicine and Biology*, vol. 47, str. 3503-3517.
4. J. Ma, L. Lu, Y. Zhan, X. Zhou, M. Salganicoff, A. Krishnan (2010), Hierarchical segmentation and identification of thoracic vertebra using learning-based edge detection and coarse-to-fine deformable model, *Proceedings MICCAI 2010*, str. 19-27.
5. A. Mastmeyer, K. Engelke, C. Fuchs, W. A. Kalender (2006), A hierarchical 3D segmentation method and the definition of vertebral body coordinate systems for QCT of the lumbar spine, *Medical Image Analysis*, vol. 10, str. 560-577.
6. Z. Peng, J. Zhong, W. Wee, J.-h. Lee (2005), Automated vertebra detection and segmentation from the whole spine MR images, *Proceedings IEEE—EMBC 2005*, str. 2527–2530.
7. D. Štern, B. Likar, F. Pernuš, T. Vrtovec (2011), Parametric modelling and segmentation of vertebral bodies in 3D CT and MR spine images, *Physics in Medicine and Biology*, vol. 56, str. 7505-7522.
8. A. M. Tenllado, M. Hribernik, S. Mali, V. Smerdu, E. Cvetko, M. Števanec (2010), *Anatomski atlas: Vodnik po človeškem telesu*, Ljubljana: Tehniška založba Slovenije.

DRUGO

VICOS EYE – SPELTNA STORITEV ZA KATEGORIZACIJO VIZUALNIH OBJEKTOV

Domen Tabernik, Luka Čehovin, Matej Kristan, Marko Boben, Aleš Leonardis

Laboratorij za umetne vizualne spoznavne sisteme
Fakulteta za računalništvo in informatiko, Univerza v Ljubljani
E-pošta: domen.tabernik@fri.uni-lj.si

POVZETEK: *V članku predstavimo arhitekturo sistema za spletno storitev, ki omogoča poganjanje naprednih algoritmov računalniškega vida porazdeljenih preko večjega števila računalnikov. Arhitekturno sistem ločimo na učenje, tokovno procesiranje v realnem času in uporabniški vmesnik za spletno storitev. Učenje implementiramo v domeni MapReduce s pomočjo Hadoop poslov, medtem ko implementiramo realno-časovno procesiranje kot aplikacijo na sistemu Storm. Kot spletni vmesnik za končnega uporabnika dodatno implementiramo tudi spletno stran in Android aplikacijo. Sistem testiramo na naši gruči računalnikov in pokažemo, da se lahko slike iz podatkovne zbirke Caltech-101 naučimo v 40 minutah, medtem ko lahko tokovno procesiranje v realnem času obdela posamezno vhodno zahtevo v manj kot dveh sekundah.*

1. UVOD

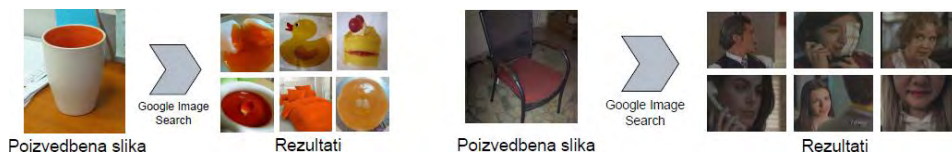
Nenehen dvig procesorske moči je v preteklih letih na področju računalniškega vida omogočil pojav spletnih storitev, ki jih v preteklosti ni bilo mogoče implementirati predvsem zaradi prevelike računske intenzivnosti algoritmov iz tega področja. Nekatere izmed novih storitev so TinEye¹, Macroglossa², Google Image Search³ in Google Goggles⁴. Vse omenjene storitve delujejo dobro predvsem na slikah, ki so jih predhodno že opazili na spletu, vendar odpovejo pri slikah, ki jih sistem še ni videl. Na primer, TinEye ne izvaja popolnoma nobenega prepoznavanja objektov, medtem ko pri vnosu slike posnete z digitalnim fotoaparatom na Google Image Search le-ta ne vrne nobenega relevantnega zadetka, če se na sliki nahaja določen splošen objekt kot sta na primer stol ali pa skodelica za kavo (slika 1).

¹ <http://www.tineye.com>

² <http://www.macroglossa.com>

³ <http://images.google.com>

⁴ <http://www.google.com/mobile/goggles>



Slika 1: Primer vrnjenih rezultatov iz Google Image Search za neznane slike.

V našem članku zato predstavimo arhitekturo sistema, ki podobno kot zgoraj omenjeni sistemi, deluje kot spletna storitev in obenem omogoča procesiranje vhodnih slik z bolj naprednimi algoritmi računalniškega vida. Pri tem je glavni cilj članka predstavitev arhitekture takšnega sistema, ki je zmožen omenjene algoritme poganjati porazdeljeno na večjem skupku računalnikov in s tem omogočiti enostavno prilagoditev storitve ob povečevanju vhodnega prometa. V splošnem je sistem zasnovan za poganjanje poljubnih algoritmov računalniškega vida, vendar se v našem primeru osredotočimo le na določen algoritem. Z našo storitvijo želimo zagotavljati kategorizacijo objektov, tako da uporabnik vnese sliko oz. izbrano regijo slike, ter jo nato naš sistem obdelja, s tem da ustvari ustrezno značilko HoC [5] in jo klasificira s pomočjo metode podpornih vektorjev (SVM) [1] v eno izmed v naprej naučenih kategorij. Informacijo o kategoriji se nato vrne uporabniku, ki lahko to informacijo uporabi za različne namene, odvisno od zelenega namena aplikacije. Na primer uporabnik informacije je lahko robotski sistem, ki želi identifikacijo objekta, ki ga vidi pred sabo, ali pa je to lahko uporabnik, ki želi prepoznati določeno vrsto živali, rastline in podobno.

2. ZAHTEVE SISTEMA

Pri implementaciji sistema moramo upoštevati sledeče zahteve:

- nudenje spletne storitve za bolj napredne algoritme računalniškega vida
- zmožnost poganjanja algoritmov preko večjega skupka računalnikov v porazdeljenem načinu
- zmožnost hitre obdelave po več sto zahtevkov na sekundo

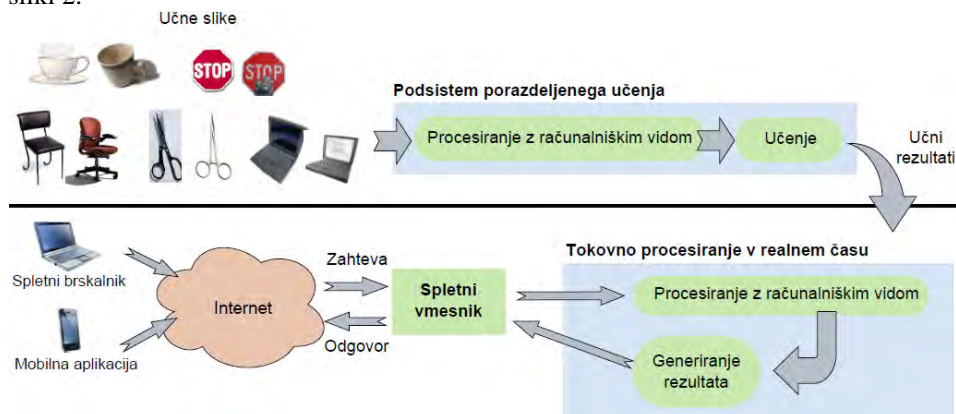
Pri prvi zahtevi moramo zagotoviti, da bo naš sistem služil kot podpora za spletno storitev, pri kateri se lahko promet konstantno spreminja. Zaradi tega moramo zagotoviti učinkovito fleksibilnost sistema, da bo imel zmožnost nenehnega prilagajanja prometu, ki se s časom lahko močno poveča. To zagotovimo z implementacijo algoritma v načinu, ki omogoča porazdelitev procesiranja preko večjega števila računalnikov. Porazdeljena implementacija mora biti dovolj fleksibilna, da lahko z enostavnim dodajanjem nove procesorske moči zagotovimo učinkovito rabo celotne računske zmogljivosti in s tem omogočimo večjo prepustnost sistema. To dosežemo z implementacijo algoritma v sistemih Hadoop [4] in Storm [6], ki sta specifično namenjena za porazdeljeno procesiranje in zelo enostavno omogočata povečanje prepustnosti s preprostim dodajanjem nove procesorske moči.

Za učinkovito zagotavljanje spletne storitve moramo prav tako poskrbeti za procesiranje večjega števila zahtevkov sočasno, pri čemer morajo biti vsi zahtevki obdelani v zelo kratkem času. V splošnem se za spletne strani priporoča čas TWT (ang. *tolerable waiting*

time) v rangu dveh sekund [1], vendar pa lahko le-ta varira glede na specifično aplikacijo. V našem primeru bi z zagotovljenim odzivnim časom do 5 sekund lahko omogočili uporabo sistema v različnih aplikacijah.

3. ARHITEKTURA

Za zagotavljanje zgornjih zahtev sistem arhitekturno razdelimo v dva glavna podsistema: *porazdeljeno učenje* in *tokovno procesiranje v realnem času*. Pri prvem podsistemu implementiramo učinkovito procesiranje večjega števila slik v paketnem načinu in ustrezno učenje na podlagi obdelanih slik. Medtem ko drugi podsistem priklopimo direktno na spletno storitev in lahko tako zagotavlja hitro procesiranje in streženje vseh vhodnih zahtev iz spleta. Kot tretji podsistem velja omeniti še uporabniški vmesnik, ki uporabnikom zagotavlja dostop do storitev sistema, in tako deluje kot posrednik med uporabnikom in sistemom za procesiranje zahtev. Spletni vmesnik smo implementirali kot preprosto spletno stran ter kot Android aplikacijo. Celotna arhitektura je prikazana na sliki 2.



Slika 2: Dvostopenjska arhitektura sistema, ki zagotavlja storitev računalniškega vida.

4. PODSISTEM ZA PORAZDELJENO UČENJE

V tem poglavju opisani podsistem je zadolžen za učenje knjižnic ali modelov, ki so nato potrebni pri izvajanju podsistema za tokovno procesiranje v realnem času. V našem primeru kategorizacije vizualnih objektov s pomočjo značilke HoC in metode podpornih vektorjev (SVM) to pomeni, da moramo vse učne slike v prvem koraku obdelati z metodo LHOP [7] in ustvariti ustrezne značilke HoC. V naslednjem koraku pa te značilke nato vstopijo kot učne vrednosti metode podpornih vektorjev. S to metodo je potrebno naučiti ločen model za vsako od učenih kategorij.

Zgoraj opisan postopek realiziramo v domeni MapReduce [2], kar nam omogoči, da lahko sistem zelo enostavno prilagodimo za povečano množico vhodnih podatkov, oz. ogromno količino slik v našem primeru. Za implementacijo MapReduce smo vzeli odprto-kodni sistem Hadoop [4], ki je spisan v Javi. Glavne značilnosti tega sistema so

zmožnost procesiranja ogromne količine podatkov (več deset Terabyte-ov podatkov) in relativno enostavna skaliranost na gruče računalnikov z 1000 ali več vozlišči.

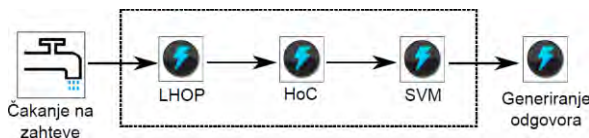
Za učinkovito implementacijo v MapReduce domeni je potrebno identificirati vhodne elemente, ki se lahko procesirajo neodvisno od ostalih elementov, in jih lahko Hadoop sistem samodejno razporedi po razpoložljivih procesorskih enotah. V primeru prvega koraka kategorizacije vizualnih objektov so ti vhodni elementi kar slike, saj lahko za vsako sliko ustvarimo značilko HoC povsem neodvisno od ostalih slik. Paralelizacija na tem nivoju je smiselna tudi zaradi ogromnega števila vhodnih slik (od 1000 do 100000 slik), ki jih potrebujemo za učenje kategorij. Ta korak implementiramo kot en Hadoop posel, ki se tako izvede za vse učne slike naenkrat, nato pa se jih shrani na porazdeljen datotečni sistem (HDFS). Izhod tega posla služi kot vhod naslednjem poslu, ki implementira učenje vseh kategorij z metodo podpornih vektorjev. V koraku strojnega učenja za vhodni element paralelizacije ne izberemo posamezne slike, ampak izberemo posamično kategorijo, saj je potrebno za učenje ene kategorije uporabiti vse učne slike (bodisi za pozitivne primere bodisi za negativne primere). V našem primeru smo za implementacijo SVM uporabili LIBSVM [1], vendar je Hadoop posel napisan tako splošno, da bi bilo mogoče metodo zamenjati s poljubno metodo strojnega učenja.

Pri obstoječi implementaciji Hadoop posla za učenje z metodo podpornih vektorjev smo lahko večjo gručo računalnikov izkoristili tudi za optimizacijo parametrov SVM metode (ang. *parameter grid optimization*). To smo storili enostavno s podvajanjem učnih kategorij, pri čemer smo za vsak element paralelizacije izbrali poleg kategorije še ustrezno vrednost parametra, ki ga optimiziramo. Metoda Reduce v postopku MapReduce je nato služila za izbiro najbolj optimalnega naučenega modela za vsako kategorijo posebej.

5. PODSISTEM ZA TOKOVNO PROCESIRANJE V REALNEM ČASU

V tem podsistemu je potrebno vsako zahtevo, ki pride iz interneta obdelati z metodo LHOP, ustvariti ustrezno značilko HoC ter značilko klasificirati v eno od predhodno naučenih kategorij. Ta postopek je zaradi zahtev po spletni storitvi potrebno narediti v manj kot 5 sekundah, kar izključuje uporabo Hadoop sistema za porazdeljeno procesiranje tega problema. Hadoop je namenjen paketnemu procesiranju ogromne količine podatkov naenkrat, kar pomeni, da je potrebno vse vhodne podatke pripraviti v naprej. Zaradi nepredvidljivega prihajanja zahtev v sistem pa to v tem podsistemu ni mogoče zagotoviti. V tem primeru lahko učinkovito paralelizacijo izvedemo z implementacijo aplikacije v sistemu Storm [6]. Slednji je namenjen za porazdeljeno procesiranje v realnem času in omogoča izredno hiter odziv na posamezno zahtevo.

Aplikacijo v Storm sistemu definiramo s t.i. topologijo (usmerjen graf), ki določa posamezne procesne elemente in tok podatkov, ki se pretaka med njimi. Primer topologije za našo aplikacijo je prikazan na sliki 3.



Slika 3: Storm topologija za kategorizacijo vizualnih objektov z uporabo značilke HoC in metode podpornih vektorjev.

Naša topologija je sestavljena iz enega izvora (ang. *spout*), ki čaka na vhodne zahteve ter jih nato posreduje naslednjim elementom v procesiranje. Naslednji procesni element (ang. *bolt*) obdela sliko z LHOP knjižnico, in podatke pošlje v naslednji element. Ta ustvari značilko HoC in jo pošlje naprej v procesni element, ki z metodo podpornih vektorjev za vsako od naučenih kategorij preveri ali značilka pripada tej kategoriji. Zadnji procesni element obdelane rezultate zapakira v ustrezen odgovor originalnemu uporabniku.

Zgornja topologija je povsem zaporedne narave in je zato odzivni čas sistema vsota vseh procesnih elementov:

$$t_{odziv} = t_{LHOP} + t_{HoC} + t_{SVM} + \bar{t} \quad (1)$$

Vendar vsak od elementov lahko zaradi narave Storm-a teče v svojem procesu in tako omogoči delovanje sistema po principu cevovoda. Prav tako lahko za vsak procesni element skrbi več procesov, zaradi česar se lahko istočasno obdeluje več zahtev naenkrat.

6. ZMOGLJIVOST SISTEMA

Zmogljivost podsistema za porazdeljeno učenje smo preverili s postavitvijo Hadoop sistema na treh računalnikih, kjer ima vsak 35 vozlišč (40 CPU jeder) in 80 GB spomina. Za učno množico smo izbrali podatkovno zbirko slik Caltech-101 in naučili nekaj več kot 100 različnih kategorij. Za učenje 18248 slik (polovica od teh slik je prezrcaljenih) na 105 vozliščih smo potrebovali 37 minut, od tega je bilo 10 minut porabljenih za sam zagon, saj le-ta ni bil spisan povsem optimalno.

Tabela 1: Povprečen odzivni čas procesiranja zahteve po procesnih elementih.

t_{LHOP}	t_{HoC}	t_{SVM}	\bar{t}	t_{odziv}
1547 ms	85 ms	462 ms	48 ms	2144 ms

Podsistem tokovnega procesiranja v realnem času smo ovrednotili na enem računalniku z 8 procesnimi jedri in 16 GB spomina. Za topologijo smo namenili 4 različne procese, tako da je vsak procesni element potekal v svojem procesu. Rezultati testiranj so podani v tabeli 1. Podane vrednosti predstavljajo povprečje 20 ponovitev. V primeru implementacije na podobni konfiguraciji kot je bila uporabljena za sistem Hadoop (105 vozlišč), lahko po teoriji strežnega procesa [8] izračunamo, da je ob povprečnem času procesiranja 2144 ms maksimalno število zahtevkov, ki bi jih sistem še lahko sprejel in obdelal v tem času, 48 zahtev na sekundo.

7. ZAKLJUČEK

V članku smo predstavili arhitekturo, ki omogoča porazdeljeno procesiranje kategorizacije vizualnih objektov in služi za procesiranje zahtevkov iz spletne storitve. Implementirali smo kategorizacijo objektov z uporabo značilke HoC [5] in metode podpornih vektorjev [1]. Arhitekturo smo razdelili na dva dela: *porazdeljeno učenje in tokovno procesiranje v realnem času*. Prvi del smo implementirali v MapReduce [2] domeni s pomočjo implementacije Hadoop [4] in pokazali, da lahko celotno podatkovno zbirko Caltech-101 naučimo v manj kot 40 minutah na konfiguraciji računalnikov s 105 vozlišči. Za tokovno procesiranje v realnem času smo uporabili sistem Storm [6] in implementirali obdelavo spletnih zahtev kot Storm aplikacijo. Pokazali smo, da z našim sistemom dosežemo povprečen odzivni čas v rangu dveh sekund ter, da bi ob konfiguraciji računalnikov s 105 vozlišči lahko obdelali do 48 zahtev na sekundo. V prihodnje želimo nadgraditi sistem z dodatnimi algoritmi računalniškega vida, specifično z iskanjem slik po vsebini, tako da bi upoštevali tudi informacijo o prepoznanih objektih na slikah.

LITERATURA

1. C.-C. Chang and C.-J. Lin. *Libsvm: A library for support vector machines*. ACM Transactions on Intelligent Systems and Technology, 2:27:1–27:27, 2011.
Programska oprema na voljo na <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
2. J. Dean, S. Ghemawat, and G. Inc. *Mapreduce: simplified data processing on large clusters*. V OSDI'04: Proceedings of the 6th conference on Symposium on Operating Systems Design & Implementation. USENIX Association, 2004.
3. F. F.-H. Nah. *A study on tolerable waiting time: how long are web users willing to wait?* Behaviour & IT, 23(3):153–163, 2004.
4. T. White. *Hadoop: The Definitive Guide*. O'Reilly Media, Inc., 1st edition, 2009.
5. D. Tabernik, M. Kristan, M. Boben, and A. Leonardis. *Learning statistically relevant edge structure improves low-level visual descriptors*. V International Conference on Pattern Recognition, 2012.
6. <http://storm-project.net>
Informacije o sistemu Storm za porazdeljeno procesiranje v realnem času-
7. S. Fidler and A. Leonardis. *Towards scalable representations of object categories: Learning a hierarchy of parts*. V CVPR. IEEE Computer Society, 2007.
8. A. O. Allen. *Probability, statistics, and queueing theory with computer science applications*. Academic Press Professional, Inc., San Diego, CA, USA, 1990.

UVAJANJE 3D TEHNOLOGIJ PRI VARSTVU KULTURNE DEDIŠČINE

Žiga Stopinšek ^a, Gregor Berginc ^b, Miran Eric ^c, Franc Solina ^a

^a Laboratorij za računalniški vid
Fakulteta za računalništvo in informatiko, Univerza v Ljubljani
E-pošta: ziga@stopinsek.eu
URL: <http://lrv.fri.uni-lj.si>

^b Tretja dimenzija d.o.o, XLAB d.o.o.
E-pošta: gregor.berginc@3dimenzija.si
URL: <http://www.3dimenzija.si>

^c Zavod za varstvo kulturne dediščine Slovenije
E-pošta: miran.eric@guest.arnes.si
URL: <http://www.zvkds.si>

POVZETEK: *Pocenitev tridimenzionalnih merilnikov in napredek tehnik snemanja in zajemanja podatkov omogočata njihovo uporabo na različnih področjih. Nova strojna in programska orodja, ki omogočajo široko računalniško podprto tridimenzionalno dokumentiranje, so še posebej pomembna tudi pri varstvu kulturne dediščine. Različne metode zajemanja podatkov, bodisi z merilniki, bodisi z fotogrametričnimi ovrednotenji množic slikovnega gradiva, je sedaj mogoče 3D dokumentirati tako posamezne predmete kot tudi prostor, v katerem živimo. V zadnjih letih postaja tridimenzionalna dokumentacija zaradi dostopnosti do strojnih orodij nepogrešljivo orodje tudi v sodobnem varstvu kulturne dediščine. Žal pa je za namensko uporabo pri analizi, standardizaciji, dokumentiranju, arhiviranju in interpretaciji tridimenzionalne dokumentirane kulturne dediščine na razpolago zelo malo programskih orodij.*

1. UVOD

Zavest o pomenu raznovrstne in bogate kulturne dediščine v svetu in še zlasti v Evropi, skokovit tehnološki napredek v informacijskih tehnologijah, ki nam omogoča digitalizacijo te dediščine, ter potreba po javni dostopnosti do dediščine, je skrbnike dediščine prepričalo o nujnosti ustvarjanja digitalnega arhiva in vzpostavitve dostopnosti do dediščine prek spletnega omrežja. Del te skupne dediščine je tudi arheološka dediščina, katere najpomembnejši del procesa razumevanja je celovito, natančno in čim bolj popolno dokumentiranje [1, 3, 4]. Tovrstna dediščina je zaradi naravnih erozijskih

procesov kot tudi namernega ali nenamernega uničevanja minljiva in izginja. Javna dostopnost do arheološke kulturne dediščine, še zlasti tiste v prostoru (pokrajine, območja bivanja, naselbine, kompleksnejše arhitekturne in druge strukture), je zelo omejena ali pa celo nedostopna (podvodna dediščina). Zato je naloga arheologov in strokovnjakov na področju varstva kulturne dediščine sistematično pripraviti dobro dokumentarno gradivo.

Analogne oblike dokumentiranja najdišč in predmetov (risana gradiva) imajo vrsto nezanemarljivih pomanjkljivosti, kakršne so lahko človeška površnost in subjektivne interpretacije v postopku sistematičnega ustvarjanja dokumentarnega gradiva. Zaradi zmogljivosti in dostopnosti do informacijskih tehnologij se je v sodobni arheologiji povečala in razširila izraba digitalnih naprav in programskih orodij za ustvarjanje dobrega dokumentarnega gradiva, hkrati pa so se povečale možnosti za pripravo dobrega interpretacijskega gradiva.

V ta nabor sodijo tudi aktivni 3D merilniki in najsodobnejša programska orodja za 3D analizo in 3D interpretacijo nizov fotografskega gradiva.

Večplastno 3D gradivo nadomešča pomanjkljivosti analognega gradiva in postaja pomemben del arheološkega dokumentarnega gradiva [7]. Dobro pripravljene dokumentarne 3D modeli postajajo pomembna podlaga za interpretativno raziskovalno delo. Sodoben razvoj programskih orodij že omogoča računalniško segmentiranje 3D modelov in analitično ter sintetično razumevanje elementov arheološke dediščine. Vid je za človeka najpomembnejši čut za zaznavanje svetlobnih dražljajev in s tem razumevanje prostora v katerem se giba. Slika, ki jo ustvarijo možgani, sekundarno pa jo človek ustvarja na različne podlage, je najlažje razumljiva oblika sporazumevanja, vsled tega pa je tudi razumevanje in predstava o arheološkem dokumentarnem gradivu najlažja z ogledom 3D modela.

Kljub temu da so v arheologiji že v uporabi metodologije zajemanja 3D podatkov, tudi aktivni 3D merilniki ter posodobljene stereofotogrametrične analize za potrebe ustvarjanja 3D modelov, pa postopki ustvarjanja 3D dokumentarnega gradiva, naprave, formati modelov ter programska orodja še niso standardizirani.

V nadaljevanju bomo predstavili nekatere postopke in oblike globinskega zajemanja podatkov na kopnem ter ocenili prednosti in slabosti teh oblik. Teste smo opravili s pripravo dokumentarnega gradiva notranjskega plovila *drevaka*, ki smo ga posneli z aktivnim 3D merilnikom. V nadaljevanju smo preverili tudi uporabnost razpoložljivih oblik dokumentiranja v podvodni arheologiji. V ta namen smo sodelovali pri dokumentiranju rimskega tovornega plovila v Sinji Gorici. V zadnjem delu bomo preverili še pomen 3D modelov v arheologiji in opisali nekatere možne rešitve pri standardizaciji in programski opreми.

2. 3D DOKUMENTACIJA NA KOPNEM

Na podlagi testnih snemanj smo ugotovili, da je uporaba aktivnih 3D merilnikov primerna predvsem za predmete in območja, ki jih želimo natančno dokumentirati.

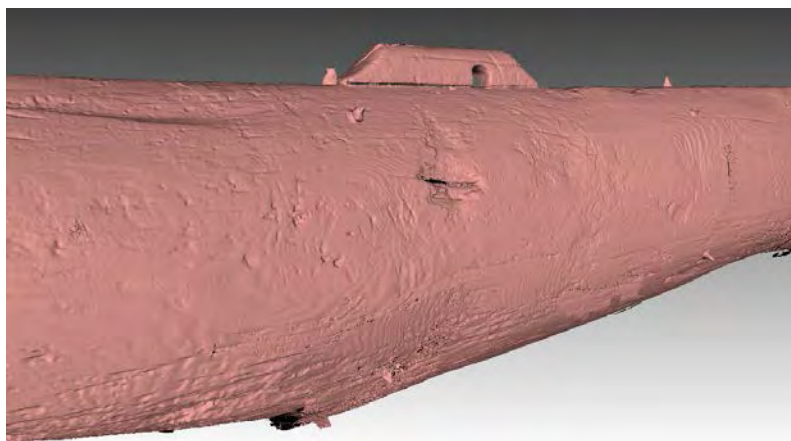
Danes je na razpolago množica različnih vrst merilnikov. Za uporabo v arheologiji so zanimivi predvsem laserski merilniki, ki delujejo na podlagi triangulacije (LMT) in merilniki s strukturirano svetlobo (MSS) [6]. Merilniki TOF (Time-Of-Flight) so primernejši za snemanje prostora in obsežnejših najdišč.

Za dokumentiranje arheoloških najdišč je pomembna mobilnost, ustrezna natančnost, terenska prilagodljivost (močna svetloba, oblačno vreme, deževje), prostorski pogoji (gradbišče, obala, jame), hitrost snemanja, sprotni nadzor nad zajetimi podatki, zajem tekstur ter cenovna ugodnost.

LMT in MSS zagotavljajo visoko natančnost in hitrost zajemanja podatkov. Slednje je predpogoj za sočasen nadzor nad zajetimi podatki. Oboji imajo težave z gladkimi, ravnimi in bleščečimi površinami ter s površinami, ki svetlobo močno vpijajo. Za večja arheološka območja in arhitekturo so primernejši LMT, MSS, ki imajo goriščno razdaljo običajno le okoli enega metra, pa so primernejši za manjša območja, dele najdišč in predmete. MSS za zajem podatka uporabljajo belo svetlobo, vsled česar so občutljivi na močno svetlobo iz okolice. Pri LMT je potrebno poskrbeti za varnost, saj so laserski žarki nevarni za oči. Na tržišču je danes dostopna množica cenovno dostopnih merilnikov (ArtecGroup, DAVID, NextEngine, Microsoft in drugi).

3D modele je mogoče ustvarjati tudi s pasivnimi metodami, predvsem fotogrametrijo. V arheologiji je zelo uporabna zaradi natančnosti, hitrosti zajemanja 3D podatkovnih slojev za večja območja in cenovne dostopnosti, vendar zahteva veliko procesorske moči pri obdelavi fotografij. Zaradi naštetega je fotogrametrija dobra zamenjava za drage merilnike TOF.

2.1 Primer: snemanje in modeliranje drevaka



Slika 1: Del 3D modela drevaka. Na sliki se vidijo napake, ki so rezultat težav pri snemanju dolgih enakomernih površin. Podrobnosti na plovilu so zajete z zadovoljivo natančnostjo.

Drevak, v svoji zbirki ga hrani tehniški muzej Slovenije v Bistri, je tipična notranjska jezerska ladja [7] dolga 5,5 m in široka do 1 m. Predstavlja primer večjega in zahtevnega predmeta za snemanje. Snemanje je trajalo 6 ur z 48 zajemi podatkov ter težavami zaradi velikosti in močnega opoldanskega sonca. Snemanje je bilo izvedeno z MSS Artec MHT, ki zajema tudi teksturo površin snemanih predmetov. Podatki so bili procesirani v aplikaciji Artec Studio 8, ki omogoča združevanje, urejanje, registracijo 3D modelov, poenostavljanje, polnjenje lukenj in odstranjevanje šuma. Zaradi velikega števila in skupne velikosti zajemov (okoli 30 GB) v glavnem pomnilniku je registracija in pretvorba dna *drevaka* v model trajala 12 dni. Rezultat je viden na Sliki 1. Pridobljena triangulirana mreža je velika okoli 2 GB.

3. 3D DOKUMENTACIJA V PODVODNI ARHEOLOGIJI

Sodobno varstvo kulturne dediščine, tako tudi v podvodni arheologiji, priporoča ohranjanje dediščine *insitu*. Vodno okolje je za arheološko raziskovalno delo v primerjavi s suhozemnimi najdišči izjemno težko [4]. Sodelavci pri projektih morajo biti vrhunsko usposobljeni potapljači, delo pod vodo pa je iz različnih razlogov (omejeno gibanje, tok, slaba vidljivost, dostopnost, globina, razpoložljivi čas) lahko tudi življenjsko nevarno. Zato je pomembno oblikovati dobro razmerje med vrhunsko izvedenim dokumentiranjem najdišča (omogoča poznejše meritve, analizo in interpretacijo) in porabljenim časom. K izboljševanju tega razmerja danes močno prispevajo metodologije 3D dokumentiranja. Pogoji za snemanje pod vodo so zahtevni. Na kvaliteto podvodnega zajema podatkov vpliva razpršitev in absorpcija svetlobe ter motnost medija, kar zmanjšuje kontrast in svetilnost, poveča izgubo podrobnosti in spremembo barve.

Aktivni 3D merilniki so prirejani za delo na kopnem. Laboratorijsko so sicer raziskovalci s primernimi ohišji in drugimi rešitvami LMT prilagodili tudi za delo pod vodo [8]. Žal pa ni cenovno dostopnih naprav za operativno terensko podvodno delo. Poseben problem predstavlja tudi uporaba bele svetlobe, zato je priporočena uporaba monokromatske ali koherentne svetlobe. Prilagoditev aktivnih merilnikov za delo pod vodo odpira vrsto težko rešljivih problemov. Za 3D modeliranje so tako primernejši in cenovno dostopnejši pasivni sistemi, npr. eno ali večsmerni stereo sistemi [9] ter fotogrametrija. Težava pri pasivnih sistemih je iskanje ujemajoče se točke na slikah (korespondenca) zaradi enostavnih in monotonih tekstur pod vodo. Dobre rezultate trenutno prispevajo kombinacije različnih metod. Na Univerzi v Calabriji so uporabili več MSS in jih povezali v stereo sistem [2]. Na Univerzi v Shizouki pa so s projekcijo binarnega vzorca med fotografiranjem predmeta rešili problem korespondence [5].

Operativnega podvodnega arheološkega dokumentiranja pa laboratorijske rešitve še vedno ne morejo povsem zadovoljiti.

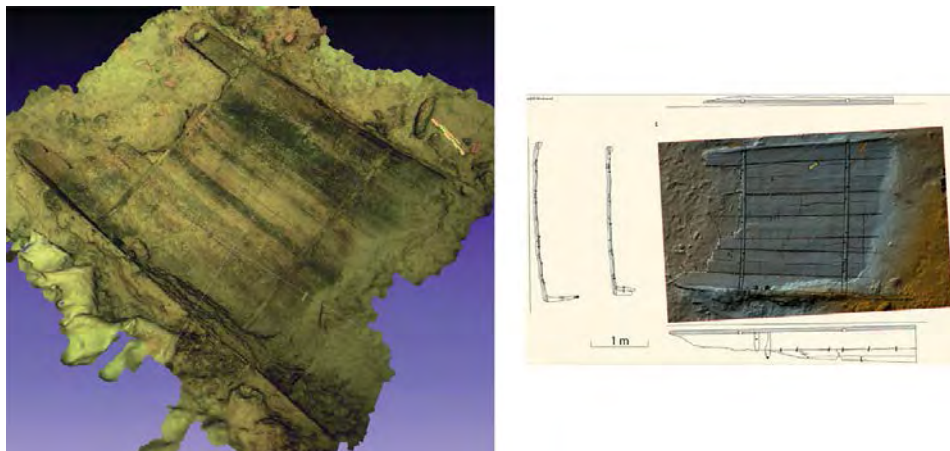
3.1 Primer: dokumentiranje rimske tovorne ladje v Sinji Gorici

Oktobra 2012 je v Sinji Gorici v Ljubljani potekalo dokumentiranje rimskega tovornega plovila iz prve polovice prvega stoletja [3]. To je avgustejsko obdobje, ko je

bilo delovanje rimskega mesta Nauporta, ki je ležal na območju današnje Vrhnike, na višku. Zaradi zahtevnosti posega in zelo kratkega razpoložljivega časa je bila v Sloveniji za 3D dokumentiranje prvič pod vodo uporabljena metoda fotogrametrije. Kljub želji po primerjalni dokumentaciji analogno dokumentiranje z risanjem tlorisa in presekov v mreži kvadrantov zaradi pomanjkanja časa ni bilo izvedeno. Vendar pa so se s fotogrametrijo pridobljeni 3D modeli izkazali za zelo natančne, reprezentativne ter izjemno analitično uporabne (Slika 2).

Na podlagi izkušenj¹ iz analognega dokumentiranja je, kljub temu, da primerjalna analiza ni bila izvedena, mogoče oceniti, da bi bila kvaliteta in natančnost analognega dokumenta zaradi omejenosti in odstopanj pri meritvah z metri in risanju na tablo vsaj 20% slabša, 3D modela pa na podlagi teh podatkov sploh ni mogoče izdelati. Izjemno pomembno je dejstvo, da bi za analogno dokumentiranje 12 m² velike površine plovila potrebovali najmanj 15 do 20 dragih in potencialno nevarnih potapljaških ur.

Za namenska fotografiranja² plovila v 4 različnih stanjih raziskav so bile potrebne 3 potapljaške ure. Iz setov fotografij (ločljivosti 4288x2848 dpi) so bili sestavljeni trije 3D modeli. V Laboratoriju za računalniški vid FRI UL za potrebe preverjanja natančnosti fotogrametričnih 3D modelov nastalih iz dveh skupin fotografij, posnetih v dveh različnih potopih, poteka primerjalna analiza, ki bo potrdila ali zavrgla pričakovanja o absolutni natančnosti dokumentirane najdbe.



Slika 2: 3D model rimske tovarne ladje (levo), in predelava modela za nadaljno analizo (desno).

¹ Soavtor pričujočega članka M. Erič na področju raziskovanj v podvodni arheologiji deluje od leta 1988, s problematiko 3D dokumentiranja podvodnih arheoloških najdišč pa se poglobljeno ukvarja od leta 1999.

² Opravil jih je Rok Kovačič (svetovni prvak v podvodni fotografiji 2011) s fotoaparatom Nikon D300 ter 18 mm objektivom. Za vsak 3D model je bilo narejenih okoli 900 fotografij.

4. UPORABA 3D PODATKOV V ARHEOLOGIJI

Potencial dobljenih 3D podatkov je, mimo atraktivnega pogleda na virtualni 3D model, širše uporaben. V nasprotju z 2D fotografijo, ki nespremenljivo določa smer in kot pogleda ter interpretirano 2D tlorisno dokumentacijo, omogoča 3D model simuliran virtualen pogled v dokumentirane površine ali predmete, ki jih je mogoče proučevati skoraj povsem tako kot *insitu*. Pomembneje pa je, da so zaradi morfoloških značilnosti 3D oblakov prostorsko umeščenih točk, absolutni posnetek trenutnega stanja, povsem odprte možnosti nadaljnjih analiz 3D modela, ki jih ne bi mogli opraviti niti realno na originalnih površinah arheoloških najdišč. Sistematično in načrtno je mogoče proučevati, segmentirati ali klasificirati izbrane površine na 3D modelu. Z avtomatiziranim iskanjem in analiziranjem površin je mogoče na modelu poiskati značilne elemente, ki bi sicer bili zaradi omejenega časa v naravnem okolju spregledani. Pomembna pa je tudi verjetna³ arhivska trajnost digitalnega arheološkega dokumentarnega gradiva kot podlaga za proučevanje, promocijo, interpretacijo in turistično trajnostno rabo.

4.1 Standardizacija

V arheologiji je v uporabi množica merilnih naprav za prostorske meritve naravnega in kulturno spremenjenega prostora.⁴ Dokaj podhranjeno je 3D dokumentiranje v večji ločljivosti, saj so tehnike, ki so v uporabi pri daljinskem zaznavanju, premalo natančne. Tudi uporaba TOF in LMT za širša 3D dokumentiranja so v proučevanju podrobnosti premalo natančna.

Zato se pojavlja potreba po standardiziranju 1) primernih tehnik in merilnikov za razumevanje razlik v natančnosti, obsegu, formatih in kompresijskih algoritmov visoko ločljivih 3D modelov, 2) dostopne in primerne programske opreme za procesiranje modelov ponuja različne algoritme in filtre za poenostavitve in druge posege v 3D model, 3) potek 3D snemanja, 4) kombinirana 3D snemanja s fotografskim arhivom in analognimi načrti, 5) možnosti analize 3D podatkov, segmentiranje, fragmentiranje ter uporaba podatkov v svetovnih spletih in mobilnih aplikacijah

4.2 Programska oprema

Le nekaj računalniških aplikacij rešuje izrazito arheološka vprašanja. Mednje sodi predvsem program za Harrisovo matriko časovnih soodvisnosti stratigrafskih enot pri izkopavanjih ter nanj vezan StratiGraf za arhiviranje arheoloških izkopavanj, ter izjemno dober Site Recorder za kartiranje dediščine. Deluje na podlagi množice medsebojnih

³ Vprašanje trajnosti digitalnih podatkov je široka in zelo zahtevna tema, predmet številnih znanstvenih raziskav, ki doslej še niso povsem zanesljivo ponudile končnega odgovora.

⁴ Posebej intenzivno so se različne merilne tehnike v arheologijo implementirale po spremenjenih načelih arheoloških znanosti, ki govorijo o čim boljšem ohranjanju arheoloških najdišč. Razvijati so se začele metode daljinskega zaznavanja (Remote Sensing), z uporabo laserskih, sonarskih, geoelektričnih, geomagnetnih, geotermalnih, ultrafrekvenčnih, spektrometričnih, seizmičnih, geokemičnih, fotogrametričnih in drugih tehnik proučevanja morfologije in notranje strukture zemljine.

meritev, ki jih aplikacija triangulira v absolutno natančnost. Žal pa je zelo zamudno zajemanje podatkov pod vodo. Nekaj programov je razvitih tudi v posameznih univerzitetnih okoljih, vendar so slabo ali celo nedostopni. Tu je še nekaj specializiranih programov za dendrokronologijo (Past) in datacije radikarbonskih datacij (Ox-Cal)

Za procesiranje 3D podatkovnih slojev, ki izvirajo iz naslova arheoloških raziskav na področju daljinskega zaznavanja, je na razpolago množica programskih orodij, v večini razvitih kot podpora strojni merilni opremi. Nekaj je tudi univerzalnih, zelo močnih, vendar tudi zelo dragih programskih rešitev. Mednje sodijo programske rešitve AutoCad, Surfer, ArcGis, PhotoModeler, Maya, Catia, Erdas, MapInfo, Las, Intergraf in mnogi drugi. S temi programi se tudi rešuje večina potreb po analitičnem dokumentiranju arheoloških podatkovnih slojev.

Za poglobljeno procesiranje 3D arheoloških podatkovnih slojev še ni razvitega učinkovitega namenskega programskega orodja. Obstajajo pa programske rešitve za izdelavo 3D arheoloških podatkovnih slojev iz digitalnih fotografij v obliki prostodostopnih in plačljivih storitev v oblaku (PHOV/Mementify⁵, 123D Catch) Omenjene rešitve temeljijo na tehnologiji »oblika iz gibanja« (ang. *structure from motion*), pri kateri se na podlagi ujemanj slikovnih značilnic sosednjih slik izvede samodejna kalibracija sistema kamer ter izgradnja gostega oblaka točk. Za prikazovanje in urejanje 3D modelov je zelo primerno prostodostopno programsko orodje MeshLab, ki omogoča tudi delo z obsežnimi 3D modeli..

Uporabniki pričakujejo naslednje rešitve: 1) interaktivno ogledovanje in urejanje 3D modelov, 2) ogledovanje in izvoz presekov modela v linearnih 3D oblikah v druga programska orodja, 3) pretvorbe 3D modelov v realno velikost in integracijo v absolutne koordinatne sisteme CAD in GIS okolja, 4) določanje in odpiranje točk na 3D modelu, za povezavo z drugimi deli arhivskega gradiva, 5) segmentacijo, poenostavljanje, dvosmerno avtomatizirano interpretacijo delno prepoznanih elementov 3D modela ter izločanje prepoznanih elementov za kasnejšo analizo, 6) klasifikacija elementov z uporabo prosto dostopnih podatkovnih zbirk ter 7) možnost fotogrametričnega 3D modeliranja brez posrednikov.

Izpostavljam nekaj brezplačnih in odprtokodnih platform, ki bi lahko bile primerne kot izhodišče razvoja rešitve. SketchUp [12] ni odprtokoden, vendar ponuja vmesnik v programskem jeziku Ruby za manipulacijo scene. Poleg tega omogoča komunikacijo z vgrajenim brskalniškim oknom. Ima pa hude pomanjkljivosti povezane s hitrostjo in odpiranjem večjih 3D modelov, ki morajo biti v COLLADA formatu.

Velik potencial ima lahko uporaba WebGL-a [13] preko knjižnic three.JS, SceneJS ali OSG.JS, vendar se lahko pojavijo prostorske in časovne težave z velikimi modeli.

MeshLab [10] je odprtokoden program za prikazovanje in urejanje 3D modelov. Je hiter, omogoča pa uvažanje in izvažanje modelov v veliko formatov (OBJ, PTY, STL, itd).

⁵ Ena pomembnejših svetovnih programskih rešitev za izdelavo 3D modelov s pomočjo fotogrametrije je nastala v podjetjih Xlab in Tretja Dimenzija iz Slovenije [11].

Ponuja vmesnik v programskem jeziku C++ za pripravo različnih razširitev: uvažanje oz. izvažanje informacij, filtre in orodja za urejanje modela.

5. ZAKLJUČEK

Varstvo kulturne dediščine je zapleten ciklični proces, ki ga vzdržujejo različni dejavniki. Razvoj 3D tehnologij je ponudil možnosti, ki bi lahko podkrepile učinkovitost njene zaščite. Omogočil je nove oblike zajema informacij ter njihove predstavitve, analize in obdelave. V sodobni arheologiji se že uporabljajo tehnike globinskega zaznavanja, vendar pa pomanjkanje standardizacije postopkov ogroža enoten pristop k arhiviranju ter analizi in obdelavi pridobljenih 3D modelov.

V članku smo predstavili nekaj metod globinskega zaznavanja, ki so primerne za 3D dokumentacijo na kopnem in pod vodo. Metode smo tudi ovrednotili na plovilu *drevaku* in potopljeni rimski ladji. Opisali smo rešitve, ki jih obetajo 3D informacije v arheologiji. Izpostavili smo nerešena vprašanja pri standardizaciji ter programske rešitve, ki bi lahko predstavljale iztočnico za nadaljnje delo.

V nadaljevanju bi bilo potrebno izvesti temeljite študije o natančnosti in uporabnosti 3D modelov ter preveriti možnost delne ali popolne zamenjave analognih metod dokumentiranja s 3D dokumentacijo. Pripraviti bo potrebno programsko opremo, ki bo sposobna takšne informacije obdelovati in učinkovito pomagati arheologom pri analizi in vrednotenju teh informacij.

LITERATURA

1. W. Böhler, G. Heinz (1999), Documentation, surveying, photogrametry, *XVII Cypa Symposium*, Olinda, Brazil.
2. F. Bruno, G. Bianco, M. Muzzupappa, S. Barone, A. V. RZIONALE (2011), Experimentation of structured light and stereo vision for underwater reconstruction, *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, str. 508-518.
3. M. Erič (2012), Documenting a Roman Vessel at Sinja Gorica, *Yearbook of the International Centre for Underwater Archaeology in Zadar*, no. 2, str. 58-65.
4. A. Gaspari, M. Erič (2010), *Minimalni standardi podvodnih arheoloških raziskav: izhodišča in smernice*, projektna študija, Ministrstvo za kulturo RS. (http://www.arhiv.mk.gov.si/fileadmin/mk.gov.si/pageuploads/Ministrstvo/raziskave-analize/dediscina/STANDARDI_PODVODNIH_RAZISKAV_20100301.pdf)
5. R. Kawai, A. Yamashita, T. Kaneko (2009), Three-dimensional measurement of objects in water by using space encoding method, *Proceedings of the 2009 IEEE international conference on Robotics and Automation*, str. 2890-2835.

6. G. Sansoni, M. Trebeschi, F. Docchio (2009), State-of-the-art and applications of 3d imaging sensors in industry, cultural heritage, medicine and criminal investigation, *Sensors*, vol. 9, str. 586-601.
7. Ž. Stopinšek (2012), *Uporabnost 3D merilnikov v kulturni dediščini*, diplomsko delo, Fakulteta za računalništvo in informatiko, Univerza v Ljubljani.
8. A. Yamashita, S. Ikeda, T. Kaneko (2005), 3-d measurement of objects in unknown aquatic environments with a laser range finder, *Proceedings of the 2005 IEEE international conference on Robotics and Automation*, str. 3912-3917.
9. A. Yamashita, R. Kawanishi, T. Koketsu, T. Kaneko, H. Asama (2011), Underwater sensing with omni-directional stereo camera, *Proceedings of the 11th Workshop on Omnidirectional Vision, Camera Networks and Non-classical Cameras (OMNIVIS2011)*, str. 304-311.
10. <http://meshlab.sourceforge.net/>
MeshLab.
11. <http://phov.eu>
PHOV.
12. <http://www.sketchup.com>
SketchUp.
13. <http://en.wikipedia.org/wiki/WebGL>
WebGL.

ŠTUDENTSKE APLIKACIJE

UJEMANJE PRSTNIH ODTISOV NA PODLAGI GREBENOV

Jaka Pohar, Jernej Bule, Peter Peer

Laboratorij za računalniški vid, Fakulteta za računalništvo in informatiko,
Univerza v Ljubljani, Tržaška cesta 25, 1000 Ljubljana
E-pošta: jakapohar@gmail.com, {jernej.bule, peter.peer}@fri.uni.lj.si

POVZETEK: *V tem delu je predstavljena nadgradnja sistema za verifikacijo oseb na podlagi prstnega odtisa FingerIdent, ki v obstoječi obliki uporablja primerjanje na podlagi značilk. Z namenom izboljšanja varnosti sistema smo implementirali še algoritem za primerjanje na podlagi grebenov in ga integrirali v obstoječ sistem. Nadgrajen sistem smo preizkusili na štirih testnih množicah s tekmovanja FVC 2002. Rezultati kažejo, da je nadgrajen sistem, v katerem se končni rezultat ujemanja oblikuje kot utežena vsota stopnje ujemanja obstoječega primerjanja značilk in implementiranega primerjanja grebenov, zanesljivejši od sistema pred nadgradnjo.*

1 UVOD

Sistemi za verifikacijo oseb na podlagi prstnega odtisa običajno temeljijo na treh glavnih korakih: zajem prstnega odtisa, iskanje značilk in odločanje [4]. Kot zadnji korak v sistemu se izvede primerjanje prstnih odtisov. Algoritem primerja dva prstna odtisa in vrne rezultat, ki je v obliki stopnje ujemanja (angl. matching score). Obstajajo tri različne metode primerjanja: korelacijska metoda, primerjanje na podlagi značilk in metoda primerjanja grebenov [6].

Korelacijska metoda (angl. correlation-based matching) deluje tako, da sliki dveh prstnih odtisov položi eno na drugo, nato pa se za različne poravnave (npr. translacije in rotacije) izračuna korelacija med ustreznimi slikovnimi elementi.

Primerjanje na podlagi značilk (angl. minutiae based matching) deluje tako, da se najprej poišče značilke dveh prstnih odtisov in se jih shrani kot množici točk v dvodimenzionalnem prostoru. Nato se množici točk poravnava tako, da je število ujemanj značilk največje.

Metoda primerjanja grebenov (angl. ridge feature-based matching) deluje tako, da primerja prstna odtisa na podlagi značilnosti, pridobljenih iz vzorca grebenov (npr. lokalna orientacija in frekvenca, oblika grebenov). Tovrstne značilnosti so v splošnem manj razločujoče, vendar pa jih je možno iz slik prstnih odtisov, ki so slabe kvalitete, lažje in bolj zanesljivo izluščiti kot značilke.

Z namenom doseganja večje zanesljivosti sistemov je smiselna tudi hkratna uporaba več različnih načinov primerjanja [1].

Sistem FingerIdent je sistem za verifikacijo oseb na podlagi prstnega odtisa, ki je bil razvit v Laboratoriju za računalniški vid na Fakulteti za računalništvo in informatiko Univerze v Ljubljani [4]. Sistem smo nadgradili tako, da smo vanj integrirali algoritem za primerjanje na podlagi grebenov, ki se skupaj z obstoječim algoritmom za primerjanje na podlagi značilk uporablja pri ugotavljanju ujemanja dveh prstnih odtisov.

2 ALGORITEM ZA PRIMERJANJE GREBENOV

Predstavljen algoritem primerjanja je modificirana verzija algoritma opisanega v [3]. Največ podobnosti ima z metodami primerjanja grebenov, saj temelji na uporabi grebenov, vendar hkrati skuša združiti prednosti vseh treh različnih metod primerjanja. Tako kot korelacijske metode ohrani večino informacije v prstnih odtisih, je razločujoč in fleksibilen kot primerjanje na podlagi značilk in odporen protu šumu, kot metode primerjanja grebenov.

Algoritem za primerjanje grebenov ugotavlja kakšna je podobnost med grebeni dveh prstnih odtisov. Greben pomeni strukturo, ki povezuje dve značilki. Algoritem najprej poišče par grebenov, ki si je najbolj podoben. Ta par uporabi kot bazni par in primerja grebene, ki so sosednji baznemu paru. Vsakega izmed parov ujemajočih se grebenov uporabi kot nov bazni par in primerja grebene, ki so mu sosednji. Postopek se rekurzivno nadaljuje dokler ne uporabimo vseh parov, ki se ujemajo. Na koncu se izračuna stopnja ujemanja, ki odraža dolžino grebenov, ki se ujemajo. Možno je, da prvi bazni par ni pravi, zato se postopek ponovi večkrat, vsakič z drugim baznim parom. Za končno stopnjo ujemanja se uporabi maksimalno ujemanje.

2.1 Predstavitev grebenov in strukturiranje podatkov

Algoritem kot vhod sprejme dva prstna odtisa, pri čemer je vsak predstavljen kot:

$$\begin{aligned} F &= \{R_1, R_2, \dots, R_n\} \\ R_i &= \{P_{i1}, P_{i2}, \dots, P_{im}\} \\ P_{ij} &= (x_{ij}, y_{ij}) \end{aligned} \tag{1}$$

F pomeni prstni odtis oz. seznam vseh grebenov tega prstnega odtisa, R_i posamezen greben oz. seznam vseh vzorčenih točk na tem grebenu, P_{ij} pa posamezno točko s koordinatama x_{ij} in y_{ij} .

Koordinate točk na grebenih pridobimo iz skeleta prstnega odtisa, ki je vmesni rezultat obdelave prstnega odtisa v sistemu FingerIdent. V fazi primerjanja bi bilo nepotrebno in časovno potratno primerjati vse točke na grebenih, zato jih pred zapisom v seznam enakomerno vzorčimo.

Za nekatere operacije, ki jih algoritem izvaja, seznamski predstavitev prstnega odtisa ni primerna. Zato povsem na začetku za vsak prstni odtis ustvarimo še matrično predstavitev prstnega odtisa. Matrično predstavitev si lahko predstavljamo kot sliko prstnega odtisa, na kateri je ozadje predstavljeno z vrednostjo 0, grebeni pa z vrednostmi večjimi od 0. Točke, ki pripadajo istemu grebenu, so označene z isto vrednostjo. Vrednost je določena z indeksom grebena v seznamski predstavitvi prstnega odtisa.

2.2 Primerjanje para poravnanih grebenov

Algoritem primerja poravnana grebena $R_1 = \{A_i\}_{i=1}^m$ in $R_2 = \{B_j\}_{j=1}^n$. Najprej primerjamo vsako točko A_i z vsako točko B_j . Poljubni A_i in B_j se ujemata, če je evklidska razdalja med njima manjša ali enaka od praga *ed_th* (*ed_th* smo empirično nastavili na 5). Ujemanje točk predstavimo z dvodimenzionalno matriko $T(i, j)$, pri čemer $i = 1, \dots, m$ in $j = 1, \dots, n$. $T(i, j) = 1$ pomeni, da se točki A_i in B_j ujemata, $T(i, j) = 0$ pomeni, da ujemanja ni.

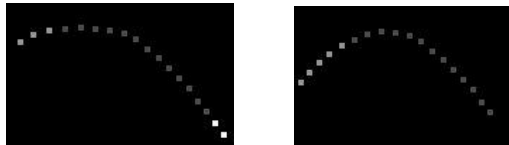
V T nato poiščemo najdaljši povezan seznam ujemajočih se točk (angl. longest point string), ki ga predstavljajo pari $\{(p_l, q_l)\}_{l=1}^k$, pri čemer sta p_l in q_l indeksa v R_1 oz. R_2 . Pari (p_l, q_l) morajo zadostiti naslednjim pogojem:

- $T(p_l, q_l) = 1$
- p_l in q_l monotono naraščata oz. padata
- $|p_l - p_{l-1}| \leq 2, |q_l - q_{l-1}| \leq 2$, pri čemer $l = 2, \dots, k$

Za iskanje seznama smo uporabili podoben postopek opisanemu v [2]. Če najdaljši povezan seznam ujemajočih se točk vsebuje vsaj k_th parov točk, se ujemanje teh točk zabeleži (k_th smo empirično nastavili na 3).

Rezultat primerjanja dveh poravnanih grebenov je lahko neujemanje, ujemanje ali delno ujemanje. V primeru, da se dva grebena le delno ujemata, njune dele, ki se ne ujemajo in vsebujejo vsaj k_th točk, na sprednjem in/ali zadnjem koncu grebenov odcepimo. Tako omogočimo, da kot novi grebeni ostanejo na voljo za kasnejšo primerjavo.

Primerjanje poravnanih grebenov R_1 in R_2 je predstavljeno na sliki 1. Temno sive točke se ujemajo, bele točke se ne ujemajo, skupini svetlo sivih točk pa se prav tako ne ujemata in jih odcepimo ter predstavimo kot nova grebena.



Slika 1: Primerjanje poravnanih grebenov

2.3 Iskanje začetnega baznega para grebenov

V tem koraku pripravimo seznam potencialnih začetnih baznih parov grebenov $I = \{(R_{1i}, R_{2i}) | R_{1i} \in F_1, R_{2i} \in F_2\}$ za prstna odtisa F_1 in F_2 . Začetne bazne pare naj bi sestavljali grebeni, ki se najbolje ujemajo, saj z njimi začnemo postopek primerjave prstnih odtisov. Pare poiščemo tako, da vsak R_{1i} primerjamo z vsakim R_{2i} . Najprej izračunamo transformacijo med R_{1i} in R_{2i} , ju poravnamo in izračunamo stopnjo ujemanja s po naslednji formuli:

$$s = \frac{m^2}{size} \quad (2)$$

m pomeni število ujemajočih se točk in *size* dolžino grebenov R_{1i} in R_{2j} . S kvadriranjem m dosežemo, da se pri uvrščanju (R_{1i}, R_{2i}) v I ne upošteva samo proporcionalno ujemanje grebenov, ampak tudi dejansko število ujemajočih se točk. V primeru, da transformacija med grebenoma presega 45%, para ne uvrstimo v I . Stopnji s za par (R_{1i}, R_{2i}) prištejemo še stopnji ujemanja dveh sosednjih parov grebenov, ki jih poravnamo glede na transformacijo med R_{1i} in R_{2i} . Z omenjenimi pravili poskušamo zagotoviti, da se na začetek seznama I uvrstijo pari grebenov, ki se dejansko ujemajo.

Na sliki 2 je prikazano računanje ujemanja za potencialni začetni bazni par. Svetlo sive točke predstavljajo ujemanje grebenov R_{1i} in R_{2i} , temno sive pa ujemanje sosednjih grebenov.



Slika 2: Računanje ujemanja za potencialni začetni bazni par

2.4 Primerjanje vseh grebenov

Ob podanem začetnem baznem paru grebenov, lahko izračunamo transformacijo med grebenoma tega para, poiščemo sosednje grebene, jih poravnamo in primerjamo. Vsak par ujemajočih se grebenov uporabimo kot nov bazni par in ponovimo računanje transformacije, iskanje, poravnavanje in primerjanje sosednjih grebenov za ta par. Postopek smo implementirali kot rekurzivno metodo, ki kot vhod sprejme ujemajoča se grebena R_1 in R_2 in ju uporabi kot bazni par. Ob prvem klicu metode kot bazni par uporabimo enega od potencialnih začetnih baznih parov grebenov.

2.4.1 Računanje transformacije

Pri računanju transformacije (translacija in rotacija) med dvema grebenoma $R_1 = \{(x_{11}, y_{11}), (x_{12}, y_{12}), \dots, (x_{1n}, y_{1n})\}$ in $R_2 = \{(x_{21}, y_{21}), (x_{22}, y_{22}), \dots, (x_{2n}, y_{2n})\}$ smo uporabili podoben postopek kot je opisan v [7]. Za grebena R_1 in R_2 , ki sta enake dolžine, v idealni situaciji velja:

$$\begin{bmatrix} \cos(\Delta\alpha) & -\sin(\Delta\alpha) \\ \sin(\Delta\alpha) & \cos(\Delta\alpha) \end{bmatrix} \cdot \begin{bmatrix} x_{2i} \\ y_{2i} \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} x_{1i} \\ y_{1i} \end{bmatrix} \quad (3)$$

α je kot rotacije in $[\Delta x \ \Delta y]^T$ translacija. Translacijo izračunamo po naslednji formuli:

$$[\Delta x \ \Delta y]^T = [x_{11} - x_{21} \ y_{11} - y_{21}]^T \quad (4)$$

Nadaljujemo z računanjem rotacije, pri čemer želimo translacijo izločiti. Od vsakega (x_{1n}, y_{1n}) odštejemo (x_{11}, y_{11}) , podobno naredimo za (x_{2n}, y_{2n}) . Spremenljivkama a_1 in a_2 pripišemo vrednosti $a_1 = \cos(\Delta\alpha)$ in $a_2 = \sin(\Delta\alpha)$. Enačba (3) se poenostavi v:

$$\begin{bmatrix} x'_{2i} & -y'_{2i} \\ y'_{2i} & x'_{2i} \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} x'_{1i} \\ y'_{1i} \end{bmatrix} \quad (5)$$

Za izračun kota orientacije α izvedemo (5) nad vsemi vzorčenimi točkami. Rešitev metode najmanjših kvadratov za parameter $a = [a_1 \ a_2]^T$ lahko določimo z reševanjem naslednje enačbe:

$$a = [A^T A]^{-1} A^T b \quad (6)$$

kjer
$$A^T = \begin{bmatrix} x'_{21} & y'_{21} & \cdots & x'_{2n} & y'_{2n} \\ -y'_{21} & x'_{21} & \cdots & -y'_{2n} & x'_{2n} \end{bmatrix}$$

in
$$b^T = [x'_{11} \ y'_{11} \ \cdots \ x'_{1n} \ y'_{1n}]$$

Kot rotacije α ocenimo iz (6) po naslednji enačbi:

$$\alpha = \arctan\left(\frac{a_2}{a_1}\right) \quad (7)$$

2.4.2 Poravnavanje grebenov

Grebena R_1 in R_2 poravnamo tako, da premaknemo R_2 glede na dano translacijo $[\Delta x \ \Delta y]^T$, kot rotacije α in bazno točko ($base_x, base_y$). Bazna točka predstavlja točko okoli katere zavrtimo greben. Določimo jo kot prvo točko grebena trenutnega baznega para.

Za premikanje grebena R_2 izvedemo naslednje enačbe za vsako točko R_2 :

$$x_i = x_{2i} - base_x \quad (8)$$

$$y_i = y_{2i} - base_y \quad (9)$$

$$\begin{bmatrix} x'_{2i} \\ y'_{2i} \end{bmatrix} = \begin{bmatrix} x_i & -y_i \\ y_i & x_i \end{bmatrix} \cdot \begin{bmatrix} \cos(\Delta\alpha) \\ \sin(\Delta\alpha) \end{bmatrix} + \begin{bmatrix} \Delta x + base_x \\ \Delta y + base_y \end{bmatrix} \quad (10)$$

2.4.3 Iskanje sosednjega grebena

Pri iskanju sosednjega grebena želimo najti greben, ki se nahaja poleg grebena R . Za to potrebujemo dve poljubni sosednji točki $P_1 = (x_1, y_1)$ in $P_2 = (x_2, y_2)$, ki se nahajata na grebenu R , in matrično predstavitev prstnega odtisa M . Iskanje sosednjega grebena poteka tako, da pregledujemo celice matrike M . Če $|x_2 - x_1| \geq |y_2 - y_1|$ pregledamo celice v vertikalni liniji med P_1 in P_2 , sicer celice v horizontalni liniji. Če naletimo na celico, ki vsebuje vrednost večjo od 0, pomeni, da smo našli sosednji greben. Vrednost celice predstavlja indeks grebena v seznamski predstavitvi prstnega odtisa.

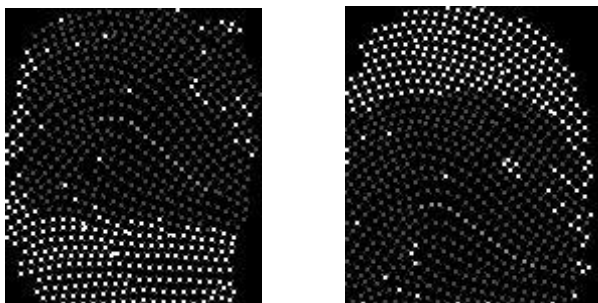
2.5 Računanje stopnje ujemanja

Pri računanju stopnje ujemanja dveh prstnih odtisov upoštevamo le območje, kjer se prstna odtisa prekrivata. Prekrivno območje (angl. overlapped region) obeh prstnih odtisov izračunamo kot presek območij posameznih prstnih odtisov, ki sta poravnana glede na transformacijo izračunano za začetni bazni par grebenov. Stopnjo ujemanja izračunamo po naslednji formuli:

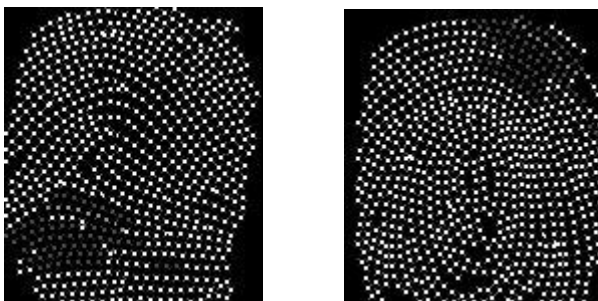
$$score_i = \frac{N_{im}^2}{N_{i1} \cdot N_{i2}} \quad (11)$$

N_{i1} in N_{i2} sta števili vzorčenih točk na grebenih dveh prstnih odtisov na prekrivnem območju, N_{im} pa je število vseh vzorčenih točk, ki se ujemanjo. Postopek računanja stopnje ujemanja in s tem tudi postopek primerjanja vseh grebenov, ponovimo za n najboljših potencialnih začetnih baznih parov grebenov, pri čemer $score_i$ predstavlja stopnjo ujemanja za i -to iteracijo (n smo empirično nastavili na 3). Za končno stopnjo ujemanja dveh prstnih odtisov uporabimo maksimalno vrednost $score_i$.

Na sliki 3 je prikazano ujemanje prstnih odtisov istega prsta, pri čemer stopnja ujemanja znaša 0,87. Na sliki 4 je prikazano ujemanje prstnih odtisov dveh različnih prstov, za katera je stopnja ujemanja 0,19.



Slika 3: Ujemanje prstnih odtisov istega prsta



Slika 4: (Ne)ujemanje prstnih odtisov različnih prstov

2.6 Konsistentne omejitve

Razcepljanje grebenov ima lahko za posledico, da se grebeni prstnih odtisov različnih prstov dobro ujemanjo. Z uporabo konsistentnih omejitev poskušamo to preprečiti.

Za vsako iteracijo primerjanja dveh prstnih odtisov z množicama grebenov R_1 in R_2 ustvarimo množico konsistentnih omejitev C , ki je na začetku prazna. Vsakič, kadar pride do razcepljanja grebena iz R_1 ali R_2 , dodamo v množico C novo omejitev, ki pove s katerim grebenom se novonastali greben sme primerjati.

Če novonastali greben nastane z odcepom prednjega oz. zadnjega dela grebena r_{1i} iz R_1 , se sme primerjati samo z grebenom, ki je nastal z odcepom prednjega oz. zadnjega dela grebena r_{2j} iz R_2 , pri čemer se morata r_{1i} in r_{2j} ujemanjati. V primeru, da do odcepa

prednjega oz. zadnjega dela r_{2j} ne pride, se novonastali greben ne sme primerjati z nobenim grebenom. V primeru, da novonastali greben nastane z odcepom prednjega ali zadnjega grebena iz R_2 , zgornje pravilo ustrezno prilagodimo.

Vsakič ob primerjanju poljubnih grebenov r_{1i} iz R_1 in r_{2j} iz R_2 za vsakega od njiju preverimo tudi ustreznost konsistentnim omejitvam C . S spremenljivko c štejemo kolikokrat je v iteraciji primerjanja dveh prstnih odtisov prišlo do neupoštevanja omejitev. Pričakujemo lahko, da bo vrednost c ob primerjanju prstnih odtisov istega prsta ostala majhna oz. enaka 0, če primerjamo dva identična prstna odtisa. Z namenom pohitritve delovanja algoritma vrednost c preverjamo sproti. Ko c preseže vrednost c_{th} , prekinemo izvajanje iteracije in stopnjo ujemanja dveh prstnih odtisov nastavimo na 0 (c_{th} smo empirično nastavili na 90).

3 REZULTATI

Sistem FingerIdent smo pred in po integraciji algoritma za primerjanje grebenov testirali na štirih množicah slik prstnih odtisov s tekmovanja FVC2002 [5]. Vsaka množica je vsebovala 800 slik. Slike so bile pridobljene z optičnim (DB1, DB2) in kapacitivnim (DB3) čitalcem, zbirka DB4 pa je vsebovala sintetično generirane prstne odtise. Z uporabo protokola iz FVC2002 smo merili naslednje parametre:

- EER (angl. equal error rate): vrednost FMR in FNMR, kadar velja $FMR = FNMR$,
- ZeroFMR: najnižja vrednost FNMR za $FMR = 0\%$,
- ZeroFNMR: najnižja vrednost FMR za $FNMR = 0\%$,
- FMR100: najnižja vrednost FNMR za $FMR \leq 1\%$,
- FMR1000: najnižja vrednost FNMR za $FMR \leq 0,1\%$.

Stopnja FNMR (angl. false non-match rate) označuje neujemanje dveh vzorcev istega osebk, stopnja FMR (angl. false match rate) pa ujemanje dveh vzorcev različnih osebkov [6].

V tabeli 1 so prikazani rezultati testiranja sistema FingerIdent pred nadgradnjo.

Tabela 1: Parametri sistema FingerIdent pred nadgradnjo

	EER	ZeroFMR	ZeroFNMR	FMR100	FMR1000
DB1	1,89%	5,18%	99,09%	2,29%	3,57%
DB2	1,65%	3,29%	93,80%	1,82%	2,75%
DB3	7,33%	29,96%	99,90%	12,21%	18,32%
DB4	3,72%	14,54%	96,91%	5,36%	9,00%

Algoritem za primerjanje grebenov smo integrirali tako, da se končni rezultat ujemanja dveh prstnih odtisov oblikuje kot utežena vsota stopnje ujemanja obstoječega primerjanja značilk in implementiranega primerjanja grebenov:

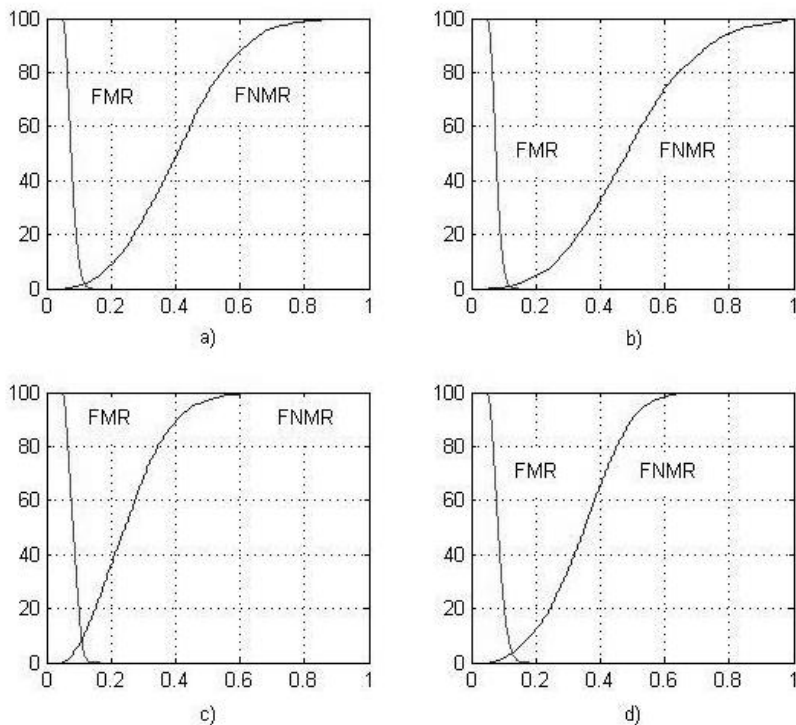
$$s = (1 - w) \cdot s_m + w \cdot s_r \quad (12)$$

Stopnja ujemanja je na intervalu [0,1], večja kot je, bolj se prstna odtisa ujemata.

Največje izboljšanje večine parametrov smo v nadgrajenem sistemu FingerIdent dosegli, ko smo parameter w nastavili na 0,1. Rezultati testiranja za omenjeno nastavitve so prikazani v tabeli 2. Vrednosti v oklepajih povejo za koliko odstotkov se je zmanjšala vrednost posameznega parametra glede na parameter sistema pred nadgradnjo. Na sliki 5 so prikazane še ROC krivulje nadgrajenega sistema.

Tabela 2: Parametri sistema FingerIdent po nadgradnji

	EER	ZeroFMR	ZeroFNMR	FMR100	FMR1000
DB1	1,85% (2,02%)	3,04% (41,38%)	94,22% (4,91%)	1,93% (15,62%)	2,89% (19,00%)
DB2	0,93% (43,69%)	1,82% (44,57%)	98,38% /	0,93% (49,02%)	1,54% (44,16%)
DB3	8,44% /	25,39% (15,25%)	98,79% (1,11%)	13,07% /	20,36% /
DB4	3,38% (8,95%)	9,61% (33,91%)	95,07% (1,90%)	4,89% (8,67%)	8,86% (1,59%)



Slika 5: ROC krivulje nadgrajenega sistema za testne množice a)DB1, b)DB2, c)DB3 in d)DB4

Izboljšanje sistema je najbolj izrazito pri DB1 in DB2, ki vsebujeta kvalitetnejše slike kot DB3 in DB4. Sklepamo lahko, da je delovanje algoritma za primerjanje grebenov na slikah dobre kvalitete zanesljivejše.

Glavna pomanjkljivost algoritma za primerjanje grebenov se nahaja v delu za iskanje potencialnih začetnih baznih parov grebenov, saj se je večkrat zgodilo, da so se na začetek seznama potencialnih začetnih baznih parov grebenov uvrstili pari, ki so se kasneje izkazali za neustrezne. Pomanjkljivost je tudi čas izvajanja algoritma (407 ms), ki je precej daljši od časa algoritma za primerjanje značilk (18 ms). Relativna počasnost algoritma za primerjanje grebenov je delno posledica tega, da je implementiran v programskem jeziku Matlab, medtem ko je algoritem za primerjanje značilk implementiran v C#. Pomembno je tudi dejstvo, da prvi algoritem v postopek primerjanja zajame približno 10-krat več točk prstnega odtisa kot drugi. Čas primerjanja dveh prstnih odtisov v nadgrajenem sistemu FingerIdent je enak vsoti obeh časov (425 ms).

4 ZAKLJUČEK

Implementiran algoritem primerja prstna odtisa na podlagi grebenov. Glavni posebnosti algoritma sta razcepljanje grebenov in sprotno izračunavanje transformacije. S tem smo omogočili, da se tolerirajo tudi morebitne nelinearne deformacije prstnih odtisov. Zanesljivost sistema FingerIdent smo z integracijo algoritma za primerjanje grebenov izboljšali. Izboljšanje parametrov je najbolj opazno pri testiranju nad množico DB2, kjer smo stopnjo EER zmanjšali za 43,69%, ZeroFMR za 44,57%, FMR100 za 49,02% in FMR1000 za 44,16%. Težava je daljši čas, ki ga sistem zdaj porabi za primerjavo dveh prstnih odtisov. V nadaljnjem delu bi zato veljalo pohitriti celoten algoritem za primerjanje grebenov in odpraviti težave pri iskanju potencialnih začetnih baznih parov grebenov.

LITERATURA

1. J. Fierrez-Aguilar, L. Nanni, J. Ortega-Garcia, R. Cappelli, D. Maltoni (2005), Combining multiple matchers for fingerprint verification: a case study in FVC2004, *Proceedings ICIAP 2005*, Cagliari: str. 1035-1042.
2. T.H. Cormen, E.C. Leiserson, L.R. Rivest, C. Stein (2009), *Introduction to Algorithms*, Cambridge (Massachusetts); London: The MIT Press.
3. J. Feng, Z. Ouyang, F. Su, A. Cai (2005): An Exact Ridge Matching Algorithm for Fingerprint Verification, *Proceedings IWBRIS 2005*, Beijing, str. 103-110.
4. U. Klopčič, P. Peer (2010), Fingerprint-based human verification system, *Proceedings IWSSIP 2010*, Rio de Janeiro, str. 150-153.
5. D. Maio, D. Maltoni, R. Cappelli, J.L. Wayman, A.K. Jain (2002), FVC2002: Second Fingerprint Verification Competition, *Proceedings 16th International Conference on Pattern Recognition*, Québec City: str. 811-814.
6. D. Maltoni, D. Maio, A.K. Jain, S. Prabhakar (2005), *Handbook of Fingerprint Recognition*, New York: Springer.
7. X. Zheng, Y. Wang (2008), Fingerprint matching based on ridge similarity, *Proceedings ICASSP 2008*, Las Vegas: str. 1701-1704.

RAZPOZNAVANJE ČLOVEŠKIH EMOCIJ NA OSNOVI DIGITALNIH POSNETKOV

Uroš Mlakar, Božidar Potočnik

**Center za obdelavo biomedicinskih signalov in slik
Fakulteta za elektrotehniko, računalništvo in informatiko
E-pošta: uros.mlakar@uni-mb.si
URL: <http://storm.uni-mb.si>**

POVZETEK: *V tem prispevku se ukvarjamo s prepoznavanjem šestih osnovnih človeških emocij, in sicer gnusa, jeze, presenečenja, straha, veselja in žalosti. Razvili smo tri metode za prepoznavanje emocij, pri čemer vse temeljijo na natančnem izrezovanju obraza v obliki elipse in sliki razlik. Pri prvi metodi izrežemo samo elipso obraza, pri drugi metodi pomembnejše dele obraza obtežimo, pri tretji metodi pa izrezano elipso z obrazom pretvorimo v polarni koordinatni sistem. Sliko razlik izračunamo med sliko z nevtralnim obrazom in sliko obraza iste osebe, ki vsebuje emocijo. Dobljeno sliko razlik popišemo z značilkami LPQ , razpoznavalnik za emocijo pa realiziramo s svojim strojem SVM. Predlagane metode smo testirali na javno dostopni bazi MMI Facial Expression Database. Rezultati so pokazali, da je naujspešnejša metoda tista, pri kateri obtežimo pomembnejše dele obraza.*

1. UVOD

Na področju računalniškega vida je v razvoju vedno več sistemov, ki se ukvarjajo z obdelavo slik ali video posnetkov z namenom avtomatizacije določenih procesov iz vsakdanjega življenja. Med bolj iskanimi objekti na slikah so človeški obrazi in njihova obrazna mimika, s pomočjo katere lahko določimo čustvo, ki ga človek doživlja. Na osnovi del avtorja v [1] ugotovimo, da so obrazno mimiko preučevali predvsem psihologi, zdravniki in umetniki, in sicer že v 17. stoletju. So se pa v zadnji četrtini 20. stoletja z napredkom na področju robotike, računalniške grafike in računalniškega vida, animatorji in računalniški inženirji vse bolj zanimali za študije obraznih mimik. Avtomatski sistemi za prepoznavanje obrazne mimike se danes uporabljajo že na številnih zanimivih področjih. Raziskovalci, ki se ukvarjajo s študijem interakcije med človekom in računalnikom se osredotočajo predvsem na oblikovanje smiselnih in odzivnih uporabniških vmesnikov. Nista pa področji robotike in komunikacije človek-računalnik edini področji uporabe avtomatske prepoznavne mimike. Uporabo teh algoritmov zasledimo tudi na področjih telekomunikacij, video iger, animacije, psihiatrije, avtomobilske varnosti, v izobraževalni opremi ipd. [1]. Kot primer podajmo aplikacijo EmotiChat [1], tj. klepetalica, ki na podlagi uporabnikove mimike vstavlja čustvene simbole v tekst.

Računalniški sistem za prepoznavo obrazne mimike pogosto razdelimo v tri podsklope oz. module, in sicer v: 1) modul za detekcijo in sledenje obrazu, 2) modul za izločanje značilk obraza, ter 3) modul za razvrščanje oz. klasifikcijo obrazne mimike. V tem prispevku se bomo osredotočili zgolj na modul za razvrščanje in klasifikacijo in pa deloma tudi na izločanje značilk obraza.

Po detekciji obraza je naslednji korak poiskati obrazne značilke, s pomočjo katerih klasificiramo obrazno mimiko in določimo, katero čustvo izkazuje opazovana oseba. V našem delu smo se osredotočili na značilke, ki jih tvorimo na osnovi informacij iz celotne slike in pa na metode, ki določijo značilke na osnovi primerjave (slike razlik) slike z nevtralnim obrazom in slike z emocijo [2]. Za popis obraznih značilk veliko avtorjev uporablja deskriptorja tekstur LBP (*angl. local binary patterns*) [6] in LPQ (*angl. local phase quantization*) [4,7].

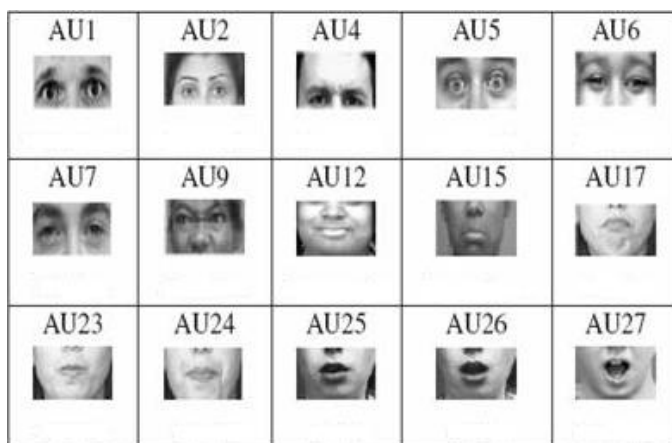
V našem delu smo najprej reproducirali rezultate metode iz [4], ki temelji na obeh zgoraj omenjenih deskriptorjih in sistemu FACS (*angl. facial action coding system*). Ta algoritem smo poimenovali algoritem LAUD. Preučili in razvili pa smo tudi algoritem za prepoznavanje emocij, ki temelji na razliki slik ter ga nadgradili z inteligentnejšim izrezom obraza iz digitalnega posnetka. Obraz smo izrezali na 3 različne načine: 1) obraz smo izrezali tako, da smo izračunali velikost elipse s pomočjo koordinat, ki jih dobimo iz detektorja obraza; 2) obraz smo izrezali na enak način kot pri točki 1, le da smo še natančneje izrezali oklepajočo škatlo elipse. Piksele, ki so padli izven meja elipse, smo nato napolnili s povprečno sivino obraza oz. z ničlami; 3) obraz smo iz digitalnega posnetka ponovno določili s pomočjo izreza elipse, piksele iz te elipse pa smo nato transformirali v polarni koordinatni sistem. Tako dobljeno območje obraza smo nato popisali z značilkami. Dobljene vektorje značilk smo zatem uporabili za učenje razpoznavalnika na osnovi strojev SVM (*angl. support vector machines*), uspešnost posameznega pristopa pa smo detajlno validirali na javno dostopni bazi MMI Facial Expression Database, ki vsebuje digitalne (video) posnetke oseb z emocijami. Vsako izmed metod smo zatem še nadgradili tako, da smo obtežili pomembnejše dele obraza (tj., področja okoli oči in ust).

Na koncu podajmo še strukturo tega prispevka. Po tem uvodu, bomo na kratko v 2. poglavju predstavili sistem FACS, ki ga uporabljamo za kodiranje obrazne mimike, nato pa bomo v 3. poglavju povzeli korake algoritma LAUD, na katerem temelji naš algoritem. V 4. poglavju bomo opisali našo idejo, ki temelji na odštevanju slik in natančnem izrezovanju obraza iz digitalnega posnetka. Zatem bomo v 5. poglavju predstavili rezultate naših eksperimentov, delo pa bomo zaključili s kratko diskusijo in smernicami za nadaljni razvoj.

2. SISTEM FACS

Sistem FACS (*angl. Facial Action Coding System*) je sistem za taksonomizacijo obrazne mimike, ki ga je razvil Paul Ekman s kolegi leta 1978. Premiki posameznih obraznih mišic se kodirajo s sistemom FACS na podlagi različnih hitrih sprememb obraza. Te spremembe in mišice, ki so odgovorne za te spremembe, se imenujejo aktivne enote –

AU (*angl. action units*). Sistem FACS je sestavljen iz več aktivnih enot. Aktivne enote delimo na aditivne in neaditivne. O aditivnih govorimo takrat, ko je pojav vsake aktivne enote neodvisen od drugih, neaditivne so pa tiste, katere posredno vplivajo na druge aktivne enote. S to definicijo postane klasifikacija obrazne mimike zelo preprosta, saj jo lahko klasificiramo s kombinacijo aditivnih in neaditivnih aktivnih enot. Vsaka aktivna enota ima lahko dodano tudi intenziteto, ki jih označujemo s črkami od A do E, kjer A pomeni najšibkejše in E najmočnejše. Na sliki 1 lahko vidimo nekaj glavnih aktivnih enot, v tabeli 1 pa, katere aktivne enote definirajo posamezno emocijo.



Slika 1: Osnovne aktivne enote v sistemu FACS.

Tabela 1: Definicije šestih osnovnih čustev z enotami AU v sistemu FACS.

EMOCIJA	AKTIVNE ENOTE
Veselje	AU6 + AU12
Žalost	AU1 + AU4 + AU15
Presenečenje	AU1 + AU2 + AU5B + AU26
Strah	AU1 + AU2 + AU4 + AU5 + AU20 + AU26
Jeza	AU4 + AU5 + AU7 + AU23
Gnus	AU9 + AU15 + AU16

3. ALGORITEM LAUD

Med raziskovalci je za detekcijo emocij in obrazne mimike zelo priljubljen sistem FACS (glej poglavje 2). V [4] so predstavili zanimivo idejo o detekciji aktivnih enot iz statičnih slik s pomočjo lokalnih deskriptorjev LBP in LPQ.

Operator LBP [6] segmentira vsak slikovni element slike glede na centralni element v maski, pri čemer sta velikost maske in oddaljenost od centralnega elementa nastavljava parametra. Vsak slikovni element tako dobi vrednost. Če upoštevamo 8-bitni rezultat operatorja LBP z osnovno masko velikosti 3×3 nad vsakim slikovnim elementom, potem vsako opazovano območje slike opišemo z histogramom z 256 elementi. Operator LPQ [7] pa uporablja fazne informacije slikovnih elementov, ki jih dobimo s pomočjo kratko časovne Fourierove transformacije, izračunane v oknu poljubne velikosti pri štirih izbranih frekvencah. Za vsak slikovni element zato dobimo 4 koeficiente, s pomočjo katerih izračunamo 8 bitno vrednost. Fazno informacijo določimo na osnovi predznaka realnih in imaginarnih delov koeficientov. Za vsako opazovano območje v sliki dobimo kot rezultat histogram z 256 elementi.

Algoritem LAUD, s katerim detektiramo aktivne enote, sestoji iz naslednjih štirih korakov. Na sliki najprej poiščemo obraz in izrežemo oklepajočo škatlo najdenega obraza. V naslednjem koraku sliko razdelimo na poljubno število regij, vsako regijo pa popišemo z izbranim deskriptorjem (bodisi LPB ali LPQ). Za vsako regijo slike torej dobimo histogram z n elementi (v delu je $n=256$). V zadnjem koraku algoritma vse dobljene histograme posameznih regij konkateneramo, s čimer dobimo histogram, ki predstavlja obraz. V algoritmu LAUD je razpoznavanje aktivnih enot realizirano s pomočjo strojev SVM.

4. ALGORITEM NA OSNOVI RAZLIKE SLIK

Izražanje emocije se odraža kot deformacija oz. sprememba na obrazu, kar lahko seveda s pridom izkoristimo. Če imamo na razpolago sliko, na kateri oseba ne izraža nobene emocije, in sliko, na kateri oseba izraža emocijo, potem lahko sliki preprosto odštejemo, s čimer dobimo slikovit opis sprememb obraza za določeno emocijo. To idejo smo vgradili v algoritem LAUD. Ker vhodne slike običajno vsebujejo preveč šuma (npr. moteče ozadje, oblačila opazovane osebe ipd.), ki lahko ima pri učenju klasifikatorja velik vpliv, smo naš algoritem, poimenovan RAZ, nadgradili še z metodo, ki iz digitalnega posnetka natančno izreže obraz osebe. Razvili smo tri variante izrezovanja obraza, in sicer: 1) natančen izrez obraza v kartezičnem koordinatnem sistemu; 2) natančen izrez obraza v kartezičnem koordinatnem sistemu in obežitev pomembnejših delov obraza; in 3) natančen izrez obraza in transformacija v polarni koordinatni sistem. V nadaljevanju opisujemo vse tri variante izrezovanja obraza.

4.1 Natančen izrez obraza v kartezičnem koordinatnem sistemu

Človeški obraz je po svoji obliki silno podoben elipsi. Če hočemo obraz aproksimirati z elipso, potem moramo natančno določiti parametre elipse, tj. položaj središča in velikost polosi. Treba je omeniti, da detektor obraza vrača obraz, ki ni rotiran, zato kota rotacije pri elipsi ne upoštevamo.

V nadaljevanju podajamo postopek določanja parametrov elipse. Koordinate središča izračunamo s pomočjo koordinat, ki jih vrača Haarov detektor obrazov. Za absciso središča vzamemo kar absciso koordinate nosu, ordinato pa izračunamo kot aritmetično sredino med ordinatami oči in nosu. Nato določimo še polosi. Za daljšo polos izberemo kar polovico višine slike, za določitev krajše polos pa analiziramo vrednosti slikovnih elementov v vrstici, ki jo določa ordinata središča elipse. Dejansko opazujemo položaj dveh največjih gradientov v tej vrstici, in sicer največjega na levo in desno stran od središča elipse. Na osnovi vseh izračunanih parametrov elipse ustvarimo masko "elipso", s pomočjo katere izrežemo obraz iz digitalnega posnetka. Z masko smo torej izrezali obraz, izračunali oklepajočo škatlo obraza, dobljeno sliko pa smo skalirali na vnaprej predpisano velikost (saj se velikost elipse lahko seveda od osebe do osebe spreminja).

V našem algoritmu smo po zgoraj opisanem postopku izrezali obraz tako iz slike z nevtralnim obrazom, kakor tudi iz slike obraza, ki vsebuje vrh emocije (tj., slike, kjer je emocija najbolj izrazita). Oba izreza smo nato odšteli, razliko slik pa popisali z značilkami LPQ. Dobljene histograme smo podali učnemu algoritmu. Primer izrezovanja obraza in odštevanje izrezov lahko vidimo na sliki 2.



Slika 2: Primer izrezovanja obraza: izrez nevtralnega obraza (levo) in izrez obraza z emocijo (sredina). Na desni je prikazan rezultat odštevanja izrezov.

4.2 Natančen izrez obraza v kartezičnem koordinatnem sistemu in obtežitev pomembnejših delov obraza

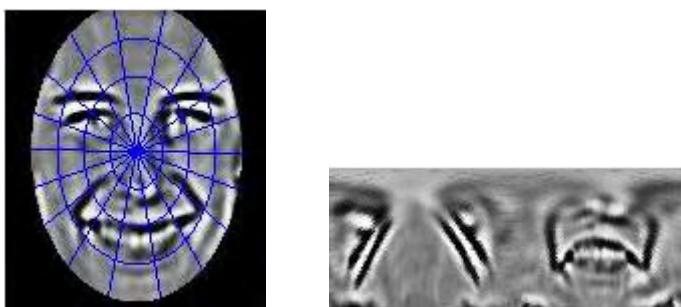
Pri izražanju emocij sodelujejo le določeni deli obraza (glej poglavje 2), zato smo dobili idejo, da bi posameznim delom obraza dali večjo pomembnost kot drugim (tj., da bi pomembnejše dele obraza obtežili). Na enak način kot v podpoglavju 4.1 smo izrezali obraz, določili oklepajočo škatlo elipse, nato pa glede na koordinate, ki jih je vrnil Haarov detektor obrazov, obtežili oči, nos in usta. Za pomembne dele obraza smo uteži nastavili na 1, ostala področja v sliki pa smo postavili na 0. Nadaljnji koraki so identični zgoraj opisanemu postopku (tj., določitev slike razlik in njen popis z značilkami LPQ). Primer rezultata z obtežitvijo pomembnejših delov obraza lahko vidimo na sliki 3.



Slika 3: Primer izrezovanja z obtežitvijo pomembnejših delov obraza : izrez nevtralnega obraza (levo) in izrez obraza z emocijo (sredina). Na desni je prikazan rezultat odštevanja izrezov.

4.3 Natančen izrez obraza in transformacija v polarni koordinatni sistem

V tretji različici izrezovanja obraza pa smo dobljeno elipso pretvorili v polarni koordinatni sistem. S pretvorbo v polarni koordinatni sistem dejansko razrešimo problem skaliranja obrazov na isto velikost. Pretvorba je potekala po naslednjem postopku. Za izhodišče uporabimo središče elipse. Iz središča pošiljamo žarke po kotu med 0 in 360 stopinj. Pri vsakem kotu izračunamo, kateri slikovni elementi pripadajo izbranemu kotu, ter tudi razdalje vseh slikovnih elementov pri izbranem kotu. Nato oblikujemo akumulator, kjer na abscisni osi teče kot, na ordinatni osi pa razdalje od središča. V eno celico akumulatorja lahko pade več slikovnih elementov, zato vrednost celice v akumulatorju določimo kot povprečje vrednosti slikovnih elementov v tem področju. Primer pretvorbe izrezanega obraza v polarni koordinatni sistem je prikazan na sliki 4.



Slika 4: Primer pretvorbe izrezanega obraza v polarni koordinatni sistem: izrez obraza v kartezičnem koordinatnem sistemu z označenimi polji (levo) in akumulator v polarnem koordinatnem sistemu, kjer je izbran korak po kotu 2 stopinji in število razdelkov je 70 (desno).

Nadaljnji koraki so identični zgoraj opisanim postopkoma (tj., določitev slike razlik in njen popis z značilkami LPQ). Rezultat lahko vidimo na sliki 5.



Slika 5: Primer rezultata pretvorbe v polarni koordinatni sistem: izrez nevtralnega obraza (levo) in izrez obraza z emocijo (sredina). Na desni je prikazan rezultat odštevanja izrezov.

5. REZULTATI

V tem poglavju bomo opisali uporabljeno eksperimentalno okolje, zatem bomo predstavili uporabljeno metriko za ocenjevanje uspešnosti algoritmov, na koncu pa bomo podali še dobljene rezultate.

5.1 Eksperimentalno okolje

Za namen testiranja smo izbrali podatkovno bazo MMI Face database [5], v kateri je zbranih več kot 1500 slik oz. videoposnetkov oseb različnih etničnih skupin, starosti in spola, ki izražajo različne emocije, ali pa so aktivne bodisi posamezne aktivne enote ali pa kombinacije le-teh. Velik del te podatkovne baze je anotiran, kar pomeni, da so poznani trenutki aktivacije posameznih aktivnih enot. Podatkovno bazo sta anotirala dva strokovnjaka za sistem FACS. V naši raziskavi smo se osredotočili na prepoznavanje šestih osnovnih emocij, in sicer: gnusa, jeze, presenečenja, straha, veselja in žalosti.

Za potrebe testiranja smo množico vseh slik razdelili na dva dela, in sicer na fiksno učno in testno množico. Množici sta bili razdeljeni tako, da se ista oseba ni pojavljala hkrati v učni in testni množici. V bazi imamo 27 oseb, od tega smo jih za učenje izbrali 21, za testiranje pa 6 oseb. V tabeli 2 je seznam vseh oseb in njihovih slik pri izbrani emociji. V tabeli smo tudi označili 6 oseb, ki smo jih uporabili za testiranje. Opazimo, da smo kot testne osebe izbrali tiste, ki imajo zabeležene slike pri vseh emocijah.

Tabela 2: Seznam oseb in število njihovih slik pri posamezni emociji. Skrajni levi stolpec vsebuje identifikator osebe iz podatkovne baze.

	GNUS	JEZA	PRESENEČENJE	STRAH	VESELJE	ŽALOST	SLIK OSEBA
1	50	75	70	64	77	63	399
2	27	114	55	22	183	35	436
3	-	-	-	-	17	-	17
6	-	23	-	-	-	-	23
15	16	-	-	-	-	-	16
16	16	-	-	-	-	-	16
28	-	-	59	29	-	-	88
30	4	-	-	-	139	-	143

31	-	22	39	39	-	21	121
32	22	48	26	26	29	-	151
33	62	20	50	17	67	59	275
34	42	22	15	12	65	-	156
35	77	49	40	-	-	17	183
36	-	45	21	-	23	-	89
37	19	51	111	13	55	116	365
39	41	29	44	56	31	14	215
40	45	10	-	6	8	-	69
41	-	108	55	-	-	-	163
42	-	47	-	-	5	3	55
43	-	-	-	-	33	-	33
44	-	-	3	33	50	-	86
45	29	57	42	31	4	-	163
46	-	135	110	17	33	15	310
47	40	68	55	17	56	34	270
48	30	20	33	-	16	-	99
49	64	76	125	11	59	-	335
50	61	-	28	60	121	19	289
Skupaj slik	645	1019	981	453	1071	396	4565
Učna množica	406	662	596	264	602	75	2605
Testna množica	239	357	385	189	469	321	1960

5.2 Metrike za merjenje uspešnosti

Uspešnost algoritmov smo želeli ovrednotiti tudi kvantitativno. V ta namen smo uporabili metriko CRE, tj. uspešnost klasifikacije emocije, ki je definirana kot

$$CRE = \frac{t}{N}$$

kjer t število pravilno razpoznanih slik za izbrano emocijo, N pa število vseh slik izbranega razreda oz. emocije.

5.3 Rezultati

V našem algoritmu smo predobdelane slike razdelili na 64 regij in jih popisali z značilkami LPQ. Treba je omeniti, da so vse slike, ki jih uporabljamo za prepoznavo emocij, predobdelane z algoritmom, opisanim v [8]. Deskriptor LPQ pa smo izbrali zato, ker je v [4] dajal dobre rezultate pri razpoznavi aktivnih enot.

Razpoznavalnik smo realizirali s pomočjo strojev SVM. Vhod v postopek učenja so bili torej obrazi, popisani z značilkami LPQ. Uporabljena funkcija za ločitev učnih vzorcev v prostoru značilk je RBF (*angl. radial basis function*). Ker smo iskali optimalne parametre za nastavitve ločitvene funkcije smo uporabili 5-kratno navzkrižno validacijo (*angl. 5-fold cross validation*).

V nadaljevanju prikazujemo rezultate naših metod. Rezultati metode, kjer obraz izrežemo s pomočjo postopka, opisanega v podpoglavju 4.1, so zbrani v tabeli 3. Postopek prepoznavanja emocij smo zagnali deset krat, pri čemer se je pri vsakem zagonu naključno izbrala slika z nevtralnim obrazom osebe. V zadnjih dveh vrsticah sta izračunana povprečna uspešnost in standardni odklon.

Tabela 3: Rezultati prepoznavanja osnovnih emocij z algoritmom, kjer je bil obraz izrezan s postopkom iz podpoglavja 4.1. V vrsticah AVG in STD sta izračunana povprečna uspešnost in standardni odklon preko vseh zagonov.

zagon	GNUS	JEZA	PRESENEČENJE	STRAH	VESELJE	ŽALOST
1	36	25	24	21	29	0,6
2	38	47	10	3	43	0,3
3	38	47	10	3	42	0,3
4	42	24	17	9	35	5
5	33	34	39	10	32	2
6	31	50	35	13	23	2,4
7	16	25	29	22	42	1,2
8	39	51	34	36	38	3
9	31	34	39	16	39	0,9
10	26	32	42	3	35	0,6
AVG	33	36,9	27,9	13,6	35,8	2,72
STD	7,6	10,8	12,1	10,5	6,4	1,4

Rezultate naše metode z natančno izrezanim obrazom in obteženimi očmi ter usti (glej podpoglavje 4.2 so prikazani v tabeli 4.

Tabela 4: Rezultati prepoznavanja osnovnih emocij z algoritmom, kjer je bil obraz izrezan s postopkom iz podpoglavja 4.2. V vrsticah AVG in STD sta izračunana povprečna uspešnost in standardni odklon preko vseh zagonov.

zagon	GNUS	JEZA	PRESENEČENJE	STRAH	VESELJE	ŽALOST
1	39	50	42	10	64	0,3

2	36	46	29	8	52	3
3	47	45	39	7	57	3
4	42	42	57	5	46	4
5	41	45	32	20	61	0,3
6	43	39	44	2	63	2
7	32	48	41	11	71	2
8	42	25	59	28	69	2
9	41	42	41	16	62	3
10	54	43	49	17	57	0
AVG	41,7	42,5	43,3	12,4	60,2	2,3
STD	5,9	6,9	9,6	7,8	7,5	1,1

Rezultate naše metode z natančno izrezanim obrazom in transformiranim v polarni koordinatni sistem (glej podpoglavje 4.3) vidimo v tabeli 5.

Tabela 5: Rezultati prepoznavanja osnovnih emocij z algoritmom, kjer je bil obraz izrezan s postopkom iz podpoglavja 4.3. V vrsticah AVG in STD sta izračunana povprečna uspešnost in standardni odklon preko vseh zagonov.

zagon	GNUS	JEZA	PRESENEČENJE	STRAH	VESELJE	ŽALOST
1	29	52	44	11	31	0,3
2	16	42	36	16	55	0,3
3	25	46	45	16	32	1,5
4	25	40	34	9	39	0,9
5	24	59	41	7	59	1,5
6	26	49	35	14	47	0,6
7	21	60	40	11	47	4
8	32	52	44	9	34	0,9
9	25	47	39	14	28	0,6
10	19	57	34	10	58	0,3
AVG	24,2	50,4	39,2	11,7	43	1,52
STD	4,6	6,8	4,2	3,1	11,7	1,5

6. DISKUSIJA IN SKLEP

V tem poglavju bomo naredili kratek pregled dobljenih rezultatov, identificirali bomo težave, s katerimi smo se srečevali, na koncu pa bomo podali še smernice za nadaljnji razvoj.

Proti pričakovanju so naše metode z natančnim izrezom obraza dajale precej slabše rezultate. Ne glede na izbrano metodo smo najuspešneje razpoznavali emocije veselja, presenečenja in pa jeze, medtem ko smo bili pri prepoznavi gnusa, strahu in žalosti precej slabši. Po našem mnenju so bile slednje tri emocije slabše razpoznane zaradi

velike medsebojne podobnosti. Pri žalosti se pa je pojavil tudi problem premajhne učne množice.

Analizirali smo tudi anotirano podatkovno bazo. Na osnovi naše subjektivne ocene smo ugotovili, da so nekatere slike anotirane kot določena emocija, čeprav z njo nimajo prav nič skupnega (tj., slika ne predstavlja emocije za katero je bila označena). Zaradi tega smo celotno testno množico anotirali še enkrat, in sicer z našimi lastnimi anotacijami. Slike smo označili le takrat, kadar smo menili, da je emocija res prisotna. Primerov "slabih" anotacij iz podatkovne baze ne moremo pokazati, saj nobena od teh oseb, pri kateri smo zaznali takšne anotacije, ni podala soglasja za objavo. Na koncu smo naš najboljši algoritem še enkrat preizkusili na naših anotacijah. Po pričakovanju se je uspešnost prepoznavanja bistveno izboljšala. Rezultate lahko vidimo v tabeli 6.

Tabela 6: Rezultati razpoznavanja emocij z najboljšo metodo (glej podpoglavje 4.2), pri čemer smo uporabili naše anotacije.

zagon	GNUS	JEZA	PRESENEČENJE	STRAH	VESELJE	ŽALOST
1	45	58	55	39	63	0,6
2	48	53	56	39	67	5
3	47	59	54	39	77	5
4	56	62	60	42	56	3
5	53	43	59	33	61	1
6	60	56	62	34	63	3
7	64	66	64	36	71	4
8	47	64	66	38	79	5
9	56	46	67	39	72	5
10	54	49	66	29	77	1
AVG	53	55,6	60,9	36,8	68,6	3,5
STD	6,2	7,7	4,8	3,8	7,8	1,6

Delo zaključimo še s kratkimi smernicami za nadaljnji razvoj. V nadaljevanju bomo skušali izboljšati naš algoritem predvsem z izbiro učinkovitejših deskriptorjev za opis obraza.

LITERATURA

1. V. Bettadapura (2012), Face Expression Recognition and Analysis: on Recognition and Analysis: The State of the Art, Tech Report.
2. M. Pantic, L. J.M. Rothkrantz (2000), Automatic Analysis of Facial Expressions: The State of the Art", IEEE Transactions Pattern Analysis and Machine Intelligence, vol. 22, št. 12, str. 1424 – 1445.

3. R. C. Gur, R. Sara, M.Hagendoorn, O. Marom, P. Hughett, L. Macy, T.Turner, R. Bajcsy, A. Posner, R. E. Gur (2002), A method for obtaining 3-dimensional facial expressions and its standardization for use in neurocognitive studies, *Journal of Neuroscience Methods*, vol. 115, str. 137-143.
4. B. Jiang, M. F. Valstar, M. Pantic (2011), Action unit detection using sparse appearance descriptors in space-time video volumes: *Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition (FG'11)*. Santa Barbara, CA, USA, str. 314 – 321.
5. <http://www.mmifacedb.com/>
Javno dostopna podatovna baza MMI facial expression database
6. T. Ojala, M. Pietikäinen, T.Mäenpää (2002), Multiresolution Gray-scale and Rotation Invariant Texture Classification with Local Binary Patterns, *IEEE Transactions Pattern Analysis and Machine Intelligence*, vol. 24, št. 7, str. 971-987.
7. V. Ojansivu, J. Heikkila (2008), Blur insensitive texture classification using local phase quantization, *Proceedings on Image and Signal Processing*, vol. 5099, str. 236 - 243
8. M. Šavc, B. Potočnik (2012), Steerable Pyramids Coded by Local Binary Patterns and Fisher Discriminant Analysis, for Facial Recognition, zbornik enaindvajsete mednarodne Elektrotehniške in računalniške konference ERK 2012, str. 173 - 176

DOLOČANJE CENTRA OČESA S POMOČJO KRIVULJ ENAKE INTENZITETE

Janez Cesar, Borut Batagelj

Fakulteta za računalništvo in informatiko, Univerza v Ljubljani

E-pošta: janez.cesar@gmail.com

POVZETEK: *V prispevku predstavljamo iskanje centra očesa na sivinskih slikah s pomočjo krivulj enakih intenzitet. Implementirali smo postopek za določanje centra očesa na osnovi krivulj, ki jih tvorijo črte oziroma krivulje enakih intenzitet na sliki. Lokalizacija centra očesa ter sledenje premikom oči se danes uporablja za številne raziskovalne in komercialne namene, v aplikacijah kot so prepoznavanje obrazov in pri interakciji človek-računalnik.*

1. UVOD

Določanje lokacije centra očesa na sliki je osnovi korak pri bolj kompleksni nalogi sledenja očem. Sledenje premikom oči je postopek, s katerim določimo smer pogleda in točko v prostoru, na katero je oseba usmerila svojo vizualno pozornost. Začetki segajo daleč nazaj v prejšnje stoletje. Na področju psihologije se sledenje očesu uporablja za preučevanje delovanja možganov, na primer zakaj ljudje usmerimo svojo pozornost na en objekt izmed množice, zakaj neko stvar gledamo dlje kot drugo ipd. Eden bolj znanih poskusov na tem področju je Yarbusov poskus [1], ki je pokazal, da so trajektorije gibanja oči odvisne od naloge, ki jo poskuša oseba opraviti. To je pokazal s sledenjem premikom oči med opazovanjem umetniške slike.

Do danes se je postopek za sledenje očem zelo razvil in se ne uporablja samo v psihologiji, temveč tudi na drugih področjih. Postopek je postal pomemben v medicini (nevrologija, oftalmologija, diagnosticiranje shizofrenije), v oglaševanju, pri preučevanju kvalitete oziroma učinkovitosti določene spletne strani ali televizijske oddaje. V oblikovalski industriji se poslužujejo sledenju pogledov za primerjavo učinkovitosti različnih uporabniških vmesnikov, na primer za razvoj kontrolnih plošč v pilotskih kabinah in armaturah avtomobilov, za primerjavo različnih zasnov menijev pri uporabniških vmesnikih. Pri učenju nekega opravila se spremenijo oziroma prilagodijo premiki oči, zato je možno s preučevanjem premikov pred in po treningu oceniti uspešnost učenja. Sledenje očem se uporablja tudi za ugotavljanje osredotočenosti uporabnika pri določeni nalogi, zaznavanju utrujenosti pri uporabi naprav in vožnji ter nenazadnje pri računalniških vmesnikih za hendikepirane osebe.

2. PREGLED TEHNIK ZA SLEDENJE PREMİKOM OČI

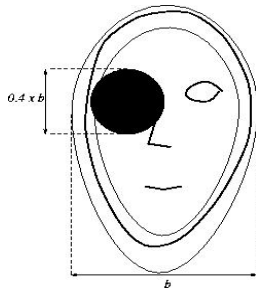
Postopke za merjenje premikov oči in določanje smeri pogleda lahko razdelimo v tri skupine: 1. kontaktne leče s tuljavo, 2. postopki, ki temeljijo na video posnetku očesa (IR ali vidna svetloba) in 3. elektrookulografija.

Pri elektrookulografiji merimo spremembe električnega potenciala na koži, do katerih pride predvsem zaradi premikov očesa. Uporaba kontaktne leče z integrirano tuljavo, je eden izmed najbolj invazivnih, a zelo natančnih postopkov. V oko se vstavi leča z vgrajeno tuljavo, premiki očesa pa se zaznavajo kot premiki te tuljave v elektromagnetnem polju. Najbolj razširjen in danes tudi najbolj raziskovan je postopek, kjer se obdeluje slika, pridobljena s pomočjo video kamere. Ta pristop je najmanj invaziven, a še vedno dovolj natančen za praktično uporabo.

Za določanje smeri pogleda moramo poznati položaj glave ter usmerjenost oči. S kombinacijo obojega lahko izračunamo absolutno smer pogleda. Nekateri pristopi se osredotočijo samo na položaj glave. Seveda taki pristopi ne morejo biti zelo natančni, saj nam da položaj glave samo neko globalno smer, medtem ko lahko oči gledajo v drugo stran. Drugi pristopi se osredotočijo samo na položaj oči. Ti temeljijo na tem, da sta zenica in šarenica okrogli, in če jih na sliki gledamo od strani, vidimo elipse [2]. Na podlagi oblike in orientacije teh elips se lahko izračuna smer pogleda. Ker imata oba pristopa svoje prednosti in pomanjklivosti, je logično, da se lahko pri iskanju položaja oči uporabita oba, kar ima za posledico hitrejši in bolj robusten sistem. Pristop, ki uporablja stereo vid ter vnaprej pripravljeno šablono zazna položaj obraza na vhodni sliki, nato na tem poišče koordinate značilnih točk, ki predstavljajo model obraza (kotički ust, oči itd.) [3]. S pomočjo teh značilk se določi orientacijo obraza. Nato se za vsako oko izračuna relativno smer pogleda. Iz koordinat kotičkov očesa se določi središče očesa in njegov premer. Med kotičkoma očesa se s Houghovo transformacijo poišče položaj šarenice. Smer pogleda je potem definirana kot razlika med središčem očesa in središčem šarenice. Eden od možnih načinov za odkrivanje šarenice znotraj očesa je, da se premikamo s primerjalno masko ustrezne oblike in velikosti po sliki očesa in iščemo korelacije med očesom in masko [4]. Pomankljivost tega postopka je ta, da je precej računsko zahteven, potrebuje pa tudi začetno inicializacijo za posameznega uporabnika.

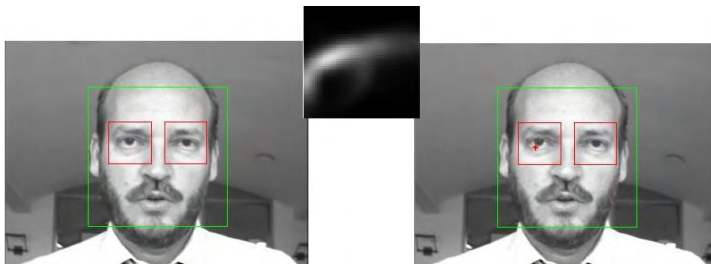
3. DOLOČANJE CENTRA OČESA S POMOČJO KRIVULJ Z ENAKO INTENZITETO

Na vhodni sliki je potrebno najprej določiti pozicijo obraza. Za ta postopek smo uporabili algoritem avtorjev Viola in Jones [5]. Nato iščemo območje kjer naj bi se nahajalo oko. Regija, ki zajema oko, je tukaj mišljena širše; ne le zenica in šarenica, temveč celotna regija, ki obsega tudi veke, trepalnice in obrvi. V našem primeru smo za omejitev regije uporabili znanje o geometrijski zgradbi obraza (Slika 1).



Slika 1: Natančnejše določanje regije iskanja očesa.

Po uspešnim lociranjem regije očesa se zdaj osredotočimo na iskanje centra očesa. Ideja katero smo uporabili temelji na tem, da so oči sestavljene iz radialno simetričnih vzorcev s podobno svetlostjo (intenziteto). Vendar pa v nasprotju z drugimi metodami, ki se osredotočajo in opirajo na simetričnost, tukaj uporabimo izophote (krivulje, katere povezujejo območja z enako intenziteto) [7]. Izophote na sliki so krivulje, ki povezujejo točke enake intenzitete (podobno kot izohipse na zemljevidu povezujejo kraje z isto nadmorsko višino). Ker se izophote ne sekajo ali stikajo, lahko sliko vedno opišemo z izophotami. Imajo pa še dodatno pozitivno lastnost, in sicer da so neodvisne od rotacij in linearnih sprememb osvetlitvenih pogojev. Za boljšo predstavitev vpeljemo lokalni koordinatni sistem, kjer ima vsaka točka svoj lokalni koordinatni sistem, ki je nastavljen tako, da kaže v smer maksimalne spremembe intenzitete, ki je določena z gradientom. Ker je ukrivljenost krivulje obratno sorazmerna z radijem, lahko izrazimo radij kroga, katerega (lahko tudi smo del) predstavlja krivulja. Radij sam nam ne pomaga veliko, če ne vemo orientacije in smeri kamor kaže. Orientacija se lahko izračuna iz gradienta, usmerjen pa je vedno proti največji razliki v svetlosti. Predznak krivulj je odvisen od intenzitete na drugi strani (če je zunanja stran svetlejša je predznak pozitiven). Z množenjem gradienta z obratno vrednostjo krivulje, dualnost krivulje poskrbi, da lahko ločimo smer centra. Te centre nato preslikamo v tabelo centrov krivulj tako imenovani akumulator IC (Slika 2).



Slika 2: Iz natančno locirane slike očesa se izračuna matrika centrov krivulj iz katere se določi središče očesa.

Ker vsak vektor določa približno pozicijo centra, akumulator zgladimo z Gaussovimi filtrom, tako da na koncu vsake skupine centrov, ki so relativno skupaj tvorimo en skupen center. Nadalje lahko doprinos vsakega vektorja pri iskanju centra še utežimo.

V realnem svetu nimamo nobenega zagotovila, da so meje objekta enake intenzitete, torej da rob oziroma mejo objekta opisuje samo ena izophota. Če bi v takem primeru dovolili glasovanje vsaki izophoti, bi dobili nesmiselne rezultate, saj oblika izophot ne ustreza obliki objekta. Da se izognemo temu, moramo zagotoviti, da glasujejo samo relevantne izophote, to so tiste, ki sledijo robu objekta.

S pomočjo operatorja ukrivljenost, ki določa ukrivljenost krivulje dobimo več informacije o povezanosti krivulj in meje objekta. Operator ukrivljenost ima nizek odziv na ravnih površinah in robovih, velik pa tam, kjer je gostota izophot maksimalna. To je okoli robov objekta, kar pomeni, da tam izophote zelo verjetno sledijo robovom objekta in tudi glasujejo za isti lokalni center. Prednost uporabe ukrivljenosti v primerjavi s pristopom samo na podlagi iskanja robov je v tem, da vsak slikovni element lahko prispeva k odločitvi, s seštevanjem glasov pa dobimo visok odziv na izophotah z istim centrom (z upoštevanjem da se te krivulje nahajajo na robu). Te visoke odzive imenujemo izocentri (IC).

Pri iskanju očesa lahko izkoristimo tudi barvo zenice in šarenice. Ker je predznak izophote odvisen od svetlosti na zunanji strani krivulje, lahko predvidevamo da negativni predznak pomeni spremembo gradienta iz svetle v temno območje. S pomočjo tega razlikujemo temne in svetle centre. Ker je pri očesu beločnica svetlejša kot roženica in šarenica lahko ignoriramo glasove, ki jih prispevajo izophote, kjer je ukrivljenost pozitivna. Metoda se dobro obnese tudi pri zelo močni osvetlitvi, saj so svetli glasovi ne upoštevajo ter pri zabrisanih oziroma nejasnih slikah.

V praksi pa se srečamo še z različnimi svetlobnimi pogoji ter napakami zaradi napol zaprtih oči, kjer veke zakrivajo del očesa. V teh primerih dobimo napačen maksimalni IC (MIC), zato za odpravo teh napak uporabimo algoritem Mean Shift in strojno učenje. Mean Shift se uporabi direktno na akumulatorju IC, z razmislekom, da imajo najrelevantnejši MIC najbolj gosto okoliško porazdeljene glasove in da napačni MIC nikoli niso prav daleč do pravih. Iskalno okno algoritma začne na najdenem maksimalnem IC in potem konvergira k regiji z maksimalno distribucijo. Po nekaj iteracijah pa je izocenter najbližji centru iskalnega okna določen za nov izračunan center očesa. Pri rešitvi strojnega učenja pa uporabimo drugo idejo. Namesto osredotočanja na en MIC, se raje osredotočimo na n najbolj relevantnih in med njimi razlikujemo s pomočjo klasifikacije. Z algoritmom SIFT se izračuna značilnica na centru okna okoli določenega MIC. Tu se tudi predvideva, da je center očesa. Izračunano značilnico potem primerjamo s knjižnico že izračunanih značilnic, ki smo jih dobili na testni množici in z uporabo klasifikatorja KNN, ki je eksperimentalno dosegel tudi najboljše rezultate, izberemo najustreznejši MIC.

4. REZULTATI

Poleg predstavljene metode na osnovi krivulj enake intenzitete v kombinaciji z algoritmom avtorjev Viola in Jones [5] za iskanje obrazov, smo metodo za primerjavo preizkusili še na regiji, katero smo pridobili iz informacije o poziciji oči, ki je priložena vsakemu obrazu v posebni datoteki. Metoda avtorjev Viola-Jones pa ni omejena samo na iskanje obrazov, temveč je s primernimi xml datotekami možno iskati tudi druge vzorce. Za primerjavo smo uporabili algoritem avtorjev Viola in in za detekcijo oči uporabili datoteko haarcascade_eye.xml [6]. Postopek iskanja centra očesa po metodi krivulj enake intenzite poteka tako, da na sivinski sliki najprej lociramo obraz. Predvidene pozicije oči nato izluščimo na podlagi atropoloških značilnostih obraza. Na tako pridobljeni manjši sliki, ki predstavlja oko in njegovo ožjo okolico nato poženemo algoritem. Algoritem vrne matriko, kamor so preslikani glasovi posameznih slikovnih elementov. Z iskanjem območij z največjo gostoto glasov nato poiščemo najverjetnejše središče očesa.

Rezultate metod smo primerjali z rezultati v članku kjer je bila metoda na osnovi krivulj tudi prvič predstavljena [7]. Za ugotavljanje napake smo vpeljali tako imenovano maksimalno normalizirano napako e , kot je razvidno iz enačbe 1, ki temelji na izračunu najslabšega zadetka očesa in se uporabi kot mera za določanje natančnosti lokacije očesa. Mero je vpeljal avtor Jesorsky [8] in je definirana kot:

$$e = \frac{\max(d_{levo}, d_{desno})}{w}, \quad (1)$$

kjer sta d_{levo} in d_{desno} evklidska razdalja med lociranim (izračunanim) očesom in pravo pozicijo očesa v priloženi datoteki. Razdalja med očmi v priloženi datoteki pa je w . Za primerjavo metod smo uporabili podatkovno bazo slik BioID [9] v kateri se nahaja 1520 slik obrazov.

Rezultati pridobljeni s kombinacijo detekcije obraza in naše metode so občutno slabši v primerjavi s temi, ko je regija že znana (Tabela 1). Ta razlika se pojavi zaradi nekoliko večje začetne regije, saj pri detekciji obraza ne moremo zelo natančno locirati regije očesa in tako lahko pride iskani center očesa v skrajni rob regije. To rezultate znatno poslabša, saj v iskanje niso bili vključeni vsi slikovni elementi, ki bi lahko glasovali za pravi center.

Tabela 1: Rezultati različnih pristopov

Metoda	($e < 0.05$)	($e < 0.10$)	($e < 0.25$)
Metoda v kombinaciji z algoritmom Viola-Jones	12.43%	52.11%	86.84%
Metoda na znani regiji	39.93%	72.70%	92.37%

Algoritem Viola-Jones z uporabo datoteke za iskanje direktno na očeh	28.95%	64.47%	71.78%
Rezultati iz članka [8]	77.15%	82.11%	96.35%

Pomanjkljivost metode v kombinaciji z algoritmom avtorjev Viola-Jones, je tudi da algoritem za prepoznavanje obrazov na sliki ne najde obraza (45 primerov) ali pa prepozna obraz na napačnem področju slike (89 primerov), kar vse pripomore k slabšemu rezultatu. Rezultati na znani regiji objektivneje kažejo, s kolikšno natančnostjo deluje iskanje oči s pomočjo krivulj z enako intenziteto. Rezultat je delno odvisen tudi od velikosti regije, na kateri poganjamo algoritem. Pri večjih regijah k rezultatu prispevajo tudi očesni lok, obrvi, pri manjših regijah pa se lahko zgodi da za center ne glasujejo vsi relevantni slikovni elementi. Pri uporabi `haarcascade_eye.xml` je algoritem za detekcijo oči na bazi BioID prepoznal še 799 oči katerih na sliki ni bilo. Na 102 slikah pa ni našel nobenega očesa, kar pri celotnem naboru slik pomeni 6,7%. Uporaba Viola-Jones algoritma za iskanje oči relativno natančno določi center očesa, vendar pa ga pesti veliko število napačno ugotovljenih oči. Kar 26% je lažno ugotovljenih oči. Iz tega sledi, da zanašanje samo algoritem Viola-Jones ni dovolj zanesljivo.

Primerjava naše metode in metode iz članka pokaže precejšnje razlike. Razlike nastanejo zaradi boljše interpretacije mapiranih rezultatov, saj so v članku v postopek vključili še metodo Mean Shift ter SIFT z uporabo KNN klasifikatorja, to pa je tudi motivacija za nadaljnje delo.

5. ZAKLJUČEK

V prispevku smo predstavili problem iskanja centra oči in pregledali nekaj najbolj razširjenih pristopov. Metoda iskanja centra očesa s pomočjo krivulj enake intenzitete je časovno nezahtevna in tako primerna za uporabo v realnočasnih sistemih. Je robustna in odporna na linearne spremembe svetlosti in rotacije. Za detekcijo očesa potrebuje metoda približno stotinko sekunde. Metoda je primerna za obdelavo slik v nizki resoluciji. Zaradi nizke časovne zahtevnosti, je metoda primerna za uporabo tudi v kombinaciji z drugimi pristopi. Problem predstavljajo svetlejšje oči, saj tam šarenica in roženica manj prispevata k izračunu centra. Tudi pri slikah na katerih so osebe z očali, kjer se lahko pojavi odblesk svetlobe ali pa senčenje okoli oči je zanesljivost metode manjša. Pričakujemo, da bo v prihodnosti z uporabo te metode možno hitro in natančno slediti uporabnikovemu pogledu z uporabo relativno preproste kamere.

LITERATURA

1. A. Yarbus (1967), *Eye Movements and Vision*, Plenum, New York.

2. J. G. Wang, E. Sung, R. Venkateswarlu (2005), Estimating the eye gaze from one eye, *Computer vision and image understanding*, vol. 98, str. 83-103.
3. Y. Matsumoto, A. Zelinsky (2000), An algorithm for real-time stereo vision implementation of head pose and gaze direction measurement, *Proceedings AFGR 2000*, str. 499-504.
4. N. Bhoi, M. N. Mohanty (2010). Article:Template Matching based Eye Detection in Facial Image. *International Journal of Computer Applications*, vol.12, št. 5, str. 15-18.
5. P. A. Viola, M. J. Jones (2001), Rapid Object Detection using a Boosted Cascade of Simple Features, *Proceedings CVPR (2001)*, Kauai, HI, USA, str. 511-518.
6. http://www-personal.umich.edu/~shameem/haarcascade_eye.html
Datoteka opisov za algoritem Viola-Jones.
7. R. Valenti, T. Gevers (2008), Accurate Eye Center Location and Tracking using Isophote Curvature, *Proceedings CVPR 2008*, Anchorage, Alaska, USA, str. 1-8.
8. O. Jesorsky, K. J. Kirchberg, R. Frischholz (2001). Robust face detection using the Hausdorff distance., *Proceedings AVBPA 2001*, str.90-95.
9. <http://www.bioid.com/index.php?q=downloads/software/bioid-face-database.html>
Baza slikovnih datotek obrazov.

RAZVIJANJE 360° PANORAMSKE SLIKE IZ LEČE S SFERIČNIM ZRCALOM

Jernej Kranjec, Božidar Potočnik
Fakulteta za elektrotehniko, računalništvo in informatiko
E-pošta: jernej.kranjec@gmail.com

POVZETEK: V prispevku predstavimo postopek za razvijanje 360° panoramske slike iz nastavka s sferičnim zrcalom Kogeto dot, namenjenim za mobilni telefon iPhone. Demonstriramo uporabo tega nastavka na drugem mobilnem telefonu ter razvoj postopka kalibracije in razvijanje panoramske slike. Preizkusimo tudi izboljšanje kvalitete razvite slike s pomočjo interpolacije, opravljeno s kubičnimi b-zlepki. Dobljeni rezultati so obetavni.

1. UVOD

Panoramske slike so slike z razširjenim vidnim poljem, ki omogoča pregled nad večjo, nepretrgano sceno, kot pa nam jo omogoča človeško oko. Ta efekt lahko dosežemo bodisi s programsko obdelavo več zaporednih slik ali pa z uporabo namenskih leč.

Danes že najdemo na tržišču amaterske potrošniške komplete za zajem in izdelavo 360° panoramskih fotografij, kot je npr. Kogeto dot [1] za mobilni telefon iPhone. So pa računalniška orodja za obdelavo takšnih posnetkov močno vezana na izbrano sistemsko rešitev. Z omenjeno strojno opremo lahko sliko razvijemo le na napravah z operacijskim sistemom iOS ali izbranim predvajalnikom na njihovi spletni strani.

V tem prispevku se ukvarjamo z razvijanjem 360° panoramske slike iz nastavka s sferičnim ogledalom. Za potrebe te raziskave smo priredili nastavek Kogeto dot na takšen način, da ga lahko uporabljamo na mobilnem telefonu Samsung Galaxy S2. V tem delu bomo tudi predstavili metodo za razvijanje panoramske slike, ki smo jo razvili v prenosljivem programskem jeziku Python. Poleg same transformacije iz logaritemsko-polarnega koordinatnega sistema v kartezični prostor moramo opraviti tudi kalibracijo, ki popravi razmerja na razviti sliki. Kvaliteto razvite slike skušamo izboljšati z uporabo interpolacije, ki jo izvedemo s pomočjo b-zlepkov.

Na koncu podajmo še strukturo tega prispevka. V 2. poglavju bomo predstavili uporabljeno strojno opremo za zajem panoramskih slik. V poglavju 3 predstavimo metode za razvijanje in interpoliranje zajete slike. V 4. poglavju prikažemo in analiziramo dobljene rezultate, delo pa zaključimo s kratkim sklepom v poglavju 5.

2. UPORABLJENA STROJNA OPREMA

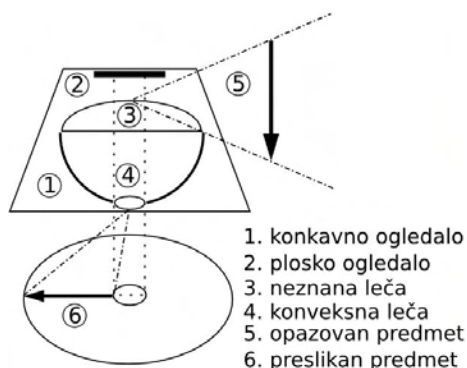
V naši raziskavi smo uporabili nastavek s sferičnim zrcalom Kogeto dot, ki omogoča zajem 360° panoramske slike s pomočjo mobilnega telefona iPhone. Ta nastavek smo za namen našega raziskovanja ustrezno prilagodili.

Najprej smo mehansko odstranili nastavek s sferičnim zrcalom ter ga pritrdili na plastični zaščitni ovitek za mobilni telefon Samsung Galaxy S2. Nastavek je bilo potrebno še mehansko obdelati, da se je prilegal s kamero mobilnega telefona. Končni rezultat, skupaj z nastavkom, lahko vidimo na sliki 1.



Slika 1: Nastavek Kogeto dot pritrjen na zaščitni ovitek.

S pomočjo vizualne analize smo identificirali temeljne komponente tega nastavka. Ohišje je iz prozorne plastike, v katero je nameščeno konkavno sferično ogledalo. Na spodnji strani sferičnega ogledala je luknja, pod katero je nameščena konveksna leča. Nad sferičnim ogledalom se nahaja še dodatna leča, nam neznanih lastnosti, ki deluje kot širokokotna leča. Nad širokokotno lečo se nahaja še dodatno plosko ogledalo. Skico modela tega nastavka ter primer zajete slike lahko vidimo na sliki 2.



Slika 2: Skiciran model nastavka s sferičnim zrcalom (levo) in primer zajete slike kalibracijske šahovnice skozi ta nastavek (desno).

Naloga širokokotne leče je preslikava okolice na konkavno sferično ogledalo. Odboj te preslikave vidimo v ploskem ogledalu skozi luknjo na dnu konkavnega ogledala. Tu nameščena konveksna leča poveča sliko s ploskega ogledala. Slika, ki jo ustvari ta nastavek je kolobar (*angl. annulus*) 360° panoramske slike. Notranji krog kolobarja je nepopačen odboj objektiva, medtem ko je v zunanjem delu kolobarja viden ostanek nastavka.

3. METODE ZA RAZVIJANJE PANORAMSKE SLIKE

V tem prispevku smo implementirali in preizkusili dve metodi za razvoj panoramske slike iz leče s sferičnim zrcalom, in sicer a) transformacijo iz polarnega koordinatnega sistema v kartezični koordinatni sistem ter b) transformacijo iz logaritemsko-polarnega koordinatnega sistema v kartezični koordinatni sistem. Učinkovitost in natančnost posamezne transformacije smo ovrednotili na sliki zajete kalibracijske šahovnice, ki smo jo podaljšali v trak in jo ovili okoli nastavka z lečo.

3.1 Transformacija iz polarnega v kartezični koordinatni sistem

Oblika zajete slike nas je napeljala na idejo, da bi uporabili transformacijo iz polarnega koordinatnega sistema v kartezični koordinatni sistem [2]. Razvijanje panoramske slike smo izvedli po enačbi (1), kjer smo središče in polmer določili ročno:

$$r = \sqrt{(x - x_{center})^2 + (y - y_{center})^2}; \quad \varphi = \text{arc tg} \left(\frac{y - y_{center}}{x - x_{center}} \right). \quad (1)$$

Pri razvijanju razumemo višino kot razliko polmerov med skrajno zunanjo in skrajno notranjo krožnico kolobarja s panoramsko sliko, kjer še imamo informacije iz okolice. Širino slike pa določimo z velikostjo kota, za katerega se premikamo po kolobarju tako da naredimo en obhod.

Rezultat takšne transformacije je na primeru razvijanja kalibracijske šahovnice prikazan na sliki 3. Opazimo lahko, da se razmerje med posameznimi kvadrati po višini spreminja. Popačenje, ki ga opazimo, je bilo vpeljšano bodisi zaradi neidealnega konveksnega zrcala ali pa zaradi neznane leče nad njim.



Slika 3: Rezultat transformacije iz polarnega v kartezični koordinatni sistem.

3.2 Transformacija iz logaritemsko-polarnega v kartezični koordinatni sistem

Stopnjevanje popačenja nas je napeljalo na idejo, da bi za razvijanje panoramske slike uporabili transformacijo iz logaritemsko-polarnega prostora v kartezični prostor [3]. Razvijanje smo dosegli z modifikacijo transformacije iz polarnega v kartezični prostor, in sicer z dodatnim logaritmiranjem (glej še enačbo (1)):

$$r' = \log(r); \quad \varphi = \varphi . \quad (2)$$

Transformacija logaritemsko-polarnega v kartezični koordinatni sistem je predstavljena z enačbo (3), kjer Δp opisuje spremembo razmerja pri upoštevanem skalirnem faktorju p , predstavljenim z enačbo (4):

$$r'' = \log \sqrt{(e^{r' \Delta p} \cos \varphi)^2 + (e^{r' \Delta p} \sin \varphi)^2}; \quad \varphi = \arctg \frac{(e^{r' \Delta p} \sin \varphi)}{(e^{r' \Delta p} \cos \varphi)}, \quad (3)$$

$$\Delta p = \frac{\log(rp)}{rp}. \quad (4)$$

Rezultat takšne transformacije je na primeru razvijanja kalibracijske šahovnice prikazan na sliki 4. Vizualna ocenitev pokaže, da je rezultat izboljššan glede na sliko 3. Kljub temu pa še vedno opazimo razliko v razmerjih med stranicami kvadratov, kot je razvidno iz slike 4. Za potrditev tega smo opravili še kalibracijo.



Slika 4: Rezultat transformacije iz logaritemsko-polarnega v kartezični koordinatni sistem brez skalirnega faktorja (skalirni faktor p je enak 1).

Kalibracijo smo izvedli tako, da smo zajeli sliko kalibracijske šahovnice pri različnih ločljivostih. Na ta način smo dobili več slik različnih velikosti. Nato smo za vsako sliko izračunali razmerje med stranicama izbranega kvadrata. S tem eksperimentom smo potrdili, da se razmerje spreminja linearno z velikostjo slike. To ugotovitev smo potem uporabili pri izračunu skalirnega faktorja. Na sliki 5 je prikazan primer rezultata te transformacije, kjer smo uporabili skalirni faktor.



Slika 5: Rezultat transformacije iz logaritemsko-polarnega v kartezični koordinatni sistem s skalirnim faktorjem. Skalirni faktor p je enak 2,6417.

3.3 Izboljšava – Interpolacija slike s pomočjo b-zlepkov

Postopek razvijanja panoramske slike preslika in raztegne piksel iz zajete v razvito sliko (glej prejšnje podpoglavje), pri čemur na dobljeni sliki opazimo t.i. efekt kockanja (*angl. pixelation*) posameznih delov slike. Do tega pride zato, ker se manjša količina razpoložljivih informacij bližje kot smo skrajnemu notranjemu krogu kolobarja. Ta efekt odpravimo s postopkom interpolacije. Za izboljšanje vizualne kvalitete slike smo torej manjkajoče informacije interpolirali s pomočjo kubičnih b-zlepkov [4].

Interpolacijo smo izvedli po enačbi (5), kjer so mejni koeficienti definirani z enačbo (6), beta funkcija pa je definirana z enačbo (7):

$$f(x, y) = \sum_{k=k_1}^{(k_1+n+1)} \sum_{l=l_1}^{(l_1+n+1)} b(k, l) \beta^3(x-k) \beta^3(y-l) . \quad (5)$$

$$k_1 = \left[x - \frac{n+1}{2} \right]; \quad l_1 = \left[y - \frac{n+1}{2} \right] . \quad (6)$$

$$\beta^3(z) = \begin{cases} \frac{2}{3} - |z|^2 + \frac{|z|^3}{2} & 0 \leq |z| \leq 1 \\ \frac{(2-|z|)^3}{6} & 1 \leq |z| \leq 2 \\ 0, & 2 \leq |z| \end{cases} \quad (7)$$

Ker uporabljamo kubične b-zlepke je njihova stopnja n enaka 3. Funkcija $b(k, l)$ vrača vrednost piksla originalne slike nad katero izvajamo interpolacijo.

Primer rezultata z opisano interpolacijo je predstavljen na sliki 6. Na tem rezultatu lahko opazimo zveznejše prehode med piksli kot na prejšnjih slikah (glej slike 2 do 4) in tudi t.i. efekt kockanja praktično ni več opazen. Vse to naredi sliko bolj gladko.



Slika 6: Rezultat transformacije iz logaritemsko-polarnega v kartezični prostor s skalirnim faktorjem in z uporabo b-zlepkov.

4. REZULTATI

V tem poglavju bomo prikazali nekaj rezultatov, dobljenih z našim postopkom.

Slike smo zajemali s pomočjo mobilnega telefona Samsung Galaxy S2, ki je imel nameščen prirejen nastavek za zajem panoramskih slik (glej opis v poglavju 2). Za lažjo orientacijo nastavka v prostoru smo privzeli, da je "pogled naprej" definiran v smeri višine mobilnega telefona (tj., leva stran ovitka z nastavkom, ki je viden na sliki 1). S tem smo fiksirali začetni kot za razvijanje zajete slike. Opazovalec torej vidi "pogled naprej" v središču razvite slike, medtem ko je rotacija po horizontu za $\pm 180^\circ$ vidna v skrajnem levem oz. desnem robu slike. V našem primeru lahko ločimo pogleda, saj je pri pogledu "nazaj" še viden del ovitka mobilnega telefona, ki še pade v vidni kot nastavka. Razliko v perspektivi pri razvijanju slike lahko opazimo med slikama 7 in 8 ter slikama 9 in 10.



Slika 7: Slika drvarnice, navidezni pogled nazaj: transformacija iz polarnega v kartezični prostor, brez interpolacije (levi del slike) in interpolirana z b-lepki (desni del slike).



Slika 8: Slika drvarnice, navidezni pogled nazaj: transformacija iz logaritemsko-polarnega v kartezični prostor s skalirnim faktorjem, brez interpolacije (levi del slike) in interpolirana z b-lepki (desni del slike).



Slika 9: Slika ceste iz strehe avtomobila, navidezni pogled naprej: transformacija iz polarnega v kartezični prostor, brez interpolacije (levi del slike) in interpolirana z b-lepki (desni del slike).



Slika 10: Slika ceste iz strehe avtomobila, navidezni pogled naprej: transformacija iz logaritemsko-polarnega v kartezični prostor s skalirnim faktorjem, brez interpolacije (levo) in interpolirana z b-lepki (desno).

Pri zajemanju slike z nastavkom lahko pride tudi do težav pri nepravilni izostritvi zajete slike. Med zajemom moramo poskrbeti da bo točka izostritve predstavljena iz središča

zajete slike na sredinsko krožnico kolobarja. V nasprotnem primeru nastavljammo fokus izostritve na zrcaljeno sliko zaslonke, namesto na projekcijo slike iz nastavka.

S transformacijo iz logaritemsko-polarnega v kartezični prostor ter opravljeno kalibracijo za skalirni faktor smo uspešno razvili 360° panoramsko sliko. S pomočjo interpolacije pa smo zmanjšali efekt kockanja, kar daje opazovalcu vizualno prijaznejšo sliko.

Kljub kalibraciji in interpolaciji lahko v slikah s kalibracijsko šahovnico (glej slike 2 do 6) še vedno opazimo popačenja, ki se manifestirajo v nepopolno ravnih horizontalnih linijah na kalibracijski šahovnici. Razlog je nepopolni krog okoli nastavka za zajem, v katerega je bila šahovnica zvita. Posledično rob šahovnice ni bil ves čas enakomerno oddaljen od središča. Do podobnega popačenja lahko pride tudi, če ni bil pravilno določen center razvitega kolobarja.

5. DISKUSIJA IN SKLEP

V prispevku smo predstavili razvijanje 360° panoramske slike iz nastavka s sferičnim ogledalom. Preizkusili smo dve metodi razvijanja, in sicer transformacijo iz polarnega v kartezični prostor in transformacijo iz logaritemsko-polarnega v kartezični prostor, kvaliteto zajetih slik pa smo skušali izboljšati z interpolacijo na osnovi b-zlepkov.

Prednost pristopa transformacije iz logaritemsko-polarnega v kartezični prostor je možnost kalibriranja. S tem lahko hitro priredimo metodo na drugačno varianto nastavka, ali pa se prilagodimo neznanim lečam ob menjavi kamere. Z lastnim razvijanjem osnovne slike pa lahko vpeljemo še interpolacijo za izboljšanje vizualne kvalitete zajete slike. Potrebno pa je omeniti, da se t.i. efekt kockanja manjša s povečevanjem ločljivosti slike. Slednje dobro opazimo na slikah 11 in 12. Na osnovi tega ugotavljamo, da pri dovolj visoki ločljivosti izboljšanje ni več očitno ter zato seveda ni več smiselno porabljati dodaten čas za izračun interpolacije.



Slika 11: Izrez slike, zajete pri ločljivosti 640 x 480 pikslov, pri čemer je bila slika obdelana ter povečana za lažjo primerjavo (od leve proti desni): s transformacijo iz polarnega v kartezični prostor s faktorjem povečave 1,9 (brez interpolacije); kot v prejšnji varianti, a s pomočjo interpolacije z b-zlepki; s transformacijo iz logaritemsko-polarnega v kartezični prostor s faktorjem povečave 3,6 (brez interpolacije); kot v prejšnji varianti, a s pomočjo interpolacije z b-zlepki.



Slika 12: Izrez slike zajete pri ločljivosti 3264 x 2448 pikselov, pri čemer je bila slika obdelana ter povečana za lažjo primerjavo (od leve proti desni): s transformacijo iz polarnega v kartezični prostor s faktorjem povečave 1,0 (brez interpolacije); kot v prejšnji varianti, a s pomočjo interpolacije z b-zlepki; s transformacijo iz logaritmsko-polarnega v kartezični prostor s faktorjem povečave 1,6 (brez interpolacije); kot v prejšnji varianti, a s pomočjo interpolacije z b-zlepki.

Delo zaključimo s smernicami za nadaljnje delo. V nadaljnjih raziskavah želimo kvantitativno ovrednotiti rezultate, s čimer bi lahko tudi natančneje izvedli kalibracijo. Možen način za objektivnejše vrednotenje rezultatov bi bil s pomočjo nadgrajene kalibracijske šahovnice, ki bi vsebovala interferiran vzorec dveh sinusnih krivulj. Dekompozicijo takšne šahovnice bi enostavno opravili v frekvenčnem prostoru. Na ta način bi lahko kvantitativno primerjali razvito sliko z generiranim originalom na šahovnici. V prihodnosti pa bomo raziskali tudi druge metode za interpolacijo slike, s pomočjo katerih bi prišli še do ostrejših slik.

LITERATURA

1. <http://www.kogeto.com/say-hello-to-dot>
Nastavek Kogeto dot
2. http://en.wikipedia.org/wiki/Polar_coordinate_system
Polarni koordinatni sistem
3. J. Sarvaiya, S. Patnaik, K. Kothari (2012), Image Registration Using Log Polar Transform and Phase Correlation to Recover Higher Scale, *Journal of Pattern Recognition Research*, vol. 7, str. 90-105.
4. S. Šprager (2008), Diplomsko delo: Elastično poravnavanje ultrazvočnih posnetkov, Maribor: Fakulteta za Elektrotehniko, Računalništvo in Informatiko, podpoglavje 4.1.