



Univerza v Mariboru

Fakulteta za elektrotehniko,
računalništvo in informatiko

Zbornik strokovne konference

ROSUS 2018

Računalniška obdelava slik in
njena uporaba v Sloveniji 2018

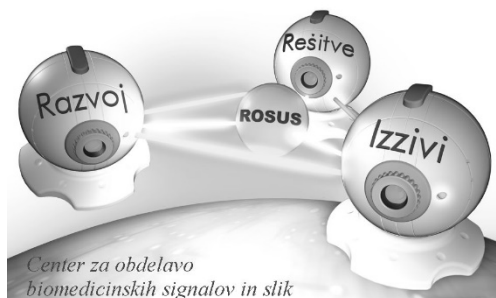
Urednik:
Božidar Potočnik



Maribor, 22. marec 2018



Univerzitetna založba
Univerze v Mariboru



ROSUS 2018 - Računalniška obdelava slik in njena uporaba v Sloveniji 2018

Zbornik 13. strokovne konference

Urednik:

izr. prof. dr. Božidar Potočnik

Maribor, 22. marec 2018

Naslov: ROSUS 2018: Računalniška obdelava slik in njena uporaba v Sloveniji 2018

Podnaslov: Zbornik 13. strokovne konference

Urednik: izr. prof. dr. Božidar Potočnik (Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko)

Strokovna recenzija: red. prof. dr. Aleš Holobar (Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko)
izr. prof. dr. Simon Dobrišek (Univerza v Ljubljani, Fakulteta za elektrotehniko)

Tehnični urednik: Gašper Sedej (Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko)

Oblikovanje ovitka: Gašper Sedej

Konferenca: Strokovna konferenca ROSUS 2018: Računalniška obdelava slik in njena uporaba v Sloveniji 2018

Datum konference: 22. marec 2018

Kraj konference: Maribor

Grafične priloge: Avtorji

Izdajateljica:

Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko
Koroška cesta 46, 2000 Maribor, Slovenija
tel. +386 2 220 70 00, faks +386 2 220 72 72
<http://feri.um.si>, feri@um.si

Založnik:

Univerzitetna založba Univerze v Mariboru
Slomškov trg 15, 2000 Maribor, Slovenija
tel. +386 2 250 42 42, fax +386 2 252 32 45
<http://press.um.si>, zalozba@um.si

Izdaja: Prva izdaja

Dostopno na: <http://press.um.si/index.php/ump/catalog/book/322>
<https://rosus.feri.um.si/rosus2018>

Izid: Maribor, marec 2018

© Univerzitetna založba Univerze v Mariboru

Vse pravice pridržane. Brez pisnega dovoljenja založnika je prepovedano reproduciranje, distribuiranje, predelava ali druga uporaba tega dela ali njegovih delov v kakršnemkoli obsegu ali postopku, vključno s fotokopiranjem, tiskanjem ali shranjevanjem v elektronski obliki. Prispevki predstavljajo stališča avtorjev, ki niso nujno usklajena s stališči organizatorja, programskega odbora in urednika zbornika, zato ne sprejemajo nobene formalne odgovornosti zaradi morebitnih avtorjevih napak, netočnosti in neustrezne rabe virov.

CIP - Kataložni zapis o publikaciji
Univerzitetna knjižnica Maribor

004.932(497.4)(082)(0.034.2)

STROKOVNA konferenca ROSUS (13 ; 2018 ; Maribor)

ROSUS 2018 [Elektronski vir] : računalniška obdelava slik in njena uporaba v Sloveniji 2018 : zbornik 13. strokovne konference / urednik Božidar Potočnik. - 1. izd. - Maribor : Univerzitetna založba Univerze, 2018

Način dostopa (URL): <http://press.um.si/index.php/ump/catalog/book/322>

ISBN 978-961-286-141-4 (pdf)

doi: 10.18690/978-961-286-141-4

1. Ov. stv. nasl. 2. Potočnik, Božidar

COBISS.SI-ID 94131713

ISBN: 978-961-286-141-4

DOI: <https://doi.org/10.18690/978-961-286-141-4>

Cena: Brezplačen izvod

Odgovorna oseba založnika: red. prof. dr. Žan Jan Oplotnik, prorektor Univerze v Mariboru

ROSUS 2018 – Računalniška obdelava slik in njena uporaba v Sloveniji 2018

Zbornik 13. strokovne konference

BOŽIDAR POTOČNIK

Povzetek: ROSUS 2018 – Računalniška obdelava slik in njena uporaba v Sloveniji 2018 je strokovna računalniška konferenca, ki jo od leta 2006 naprej vsako leto organizira Inštitut za računalništvo iz Fakultete za elektrotehniko, računalništvo in informatiko, Univerze v Mariboru. Konferenca povezuje strokovnjake in raziskovalce s področij digitalne obdelave slik in strojnega vida z uporabniki tega znanja, pri čemer uporabniki prihajajo iz raznovrstnih industrijskih okolij, biomedicine, športa, zabavništva in sorodnih področij. Zbornik konference ROSUS 2018 združuje 13 strokovnih prispevkov šestindvajsetih avtorjev, od tega 2 vabljeni predavanji ter 5 demonstracijskih prispevkov. Prispevki podajajo najnovejše dosežke slovenskih strokovnjakov s področij digitalne obdelave slik in strojnega vida, osvetljujejo pa tudi trende in novosti na omenjenih strokovnih področjih. Velik poudarek prispevkov je na promoviranju ekonomske koristnosti aplikacij računalniške obdelave slik in vida v slovenskem prostoru. Takšne računalniške aplikacije zaradi visoke natančnosti, robustnosti in izjemnih hitrosti pri obdelovanju informacij nudijo namreč nove priložnosti za uveljavitev na trgu visokih tehnologij.

Ključne besede: računalniška obdelava slik • strojni vid • biomedicina • industrijske aplikacije • prenos znanja

NASLOV UREDNIKA: dr. Božidar Potočnik, izredni profesor, Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko, Koroška cesta 46, 2000 Maribor, Slovenija, e-pošta: bozidar.potocnik@um.si.

KAZALO

VABLJENA PREDAVANJA	1
Klemen Forstnerič, Rok Ajdnik, Reveel, San Francisco, ZDA <i>Od 20 tisoč do 20 milijonov</i>	3
Andrej Vovk, MF, Ljubljana <i>Izdelava modela človeškega organa za tridimenzionalno tiskanje na osnovi magnetno-resonančnih slik</i>	10
DEMONSTRACIJSKA SEKCIJA	
Zajem slikovnega materiala: Praktični nasveti strokovnjakov	17
Jure Škrabar, Kolektor Vision, Ljubljana <i>Tehnologije zajema slikovnih informacij</i>	19
Rok Mihelj, Peter Lepej, VISTION, Slovenska Bistrica <i>Uporaba tržno dostopnih ali namenskih sistemov strojnega vida</i>	23
Jernej Kranjec, FERI, Maribor <i>Naredi sam: multispektralne in globinske kamere</i>	29
Jure Škrabar, Kolektor Vision, Ljubljana <i>Zajem slikovnih informacij: od zasnove do izvedbe</i>	31
Damjan Vrenčur, Kolektor Vision, Ljubljana <i>Primer iz prakse: Umeščanje skena krožne izvrtine v absolutni koordinatni sistem</i>	34

DOPOLDANSKA SEKCIJA

Industrijske aplikacije

Medicinske in biomedicinske aplikacije

Drugo

Študentske aplikacije 41

Blaž Pongrac, Danijel Šipoš, Dušan Gleich, Luka Ažman,

FERI, Maribor, Magneti, Ljubljana

Avtomatizirana vizualna kontrola magnetov..... 43

Tadej Ciglarič, Žiga Emeršič, Peter Peer, Blaž Meden,

FRI, Ljubljana

Towards improving face deidentification pipeline..... 49

Marko Vačovnik, Janez Perš,

Gorenje, Velenje, FE, Ljubljana

Sistem za testiranje uporabniškega vmesnika vgradnega računalniškega sistema..... 56

Rok Košir, Luc Horvat, Peter Peer,

FRI, Ljubljana, FF, Maribor

Vpliv navidezne resničnosti na človekove zaznave 63

Grega Vrbančič, Vili Podgorelec,

FERI, Maribor

Klasifikacija slik EEG signalov z uporabo globokega učenja 71

Miha Grgič Jelen, Borut Batagelj,

FRI, Ljubljana

Prepoznavna tekačev na fotografijah 79

Spoštovani!

Po dvanajstih konferencah ROSUS 2006–2017 želimo tudi s konferenco ROSUS 2018 nadaljevati s promoviranjem pomembnosti ekonomske koristi 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. Novost letošnje konference je demonstracijska sekcija z naslovom »Zajem slikovnega materiala: Praktični nasveti strokovnjakov«, ki smo jo organizirali v sodelovanju s podjetjem Kolektor Vision in podjetjem VISTION. Demonstrirali bomo, da avtomatska obdelava v industriji lahko zaradi svoje natančnosti in hitrosti prinaša velike ekonomske koristi, hkrati pa nakazali, 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.

Naj sklenemo uvodne misli še s pristrčno zahvalo Javnemu skladu Republike Slovenije za podjetništvo, ki je v okviru konference ROSUS 2018 predstavil zanimive finančne instrumente za spodbujanje prenosa tehnoloških rešitev v podjetniško sfero. Na koncu želimo izpostaviti še medijskega pokrovitelja revijo IRT3000, ki je intenzivno promovirala konferenco ROSUS 2018 ter pomen strojnega vida v slovenskem prostoru.

Božidar Potočnik
predsednik konference
ROSUS 2018



Programski odbor ROSUS 2018:

dr. Aleš Holobar
dr. Damjan Zazula
dr. Boris Cigale
dr. Marjan Mernik
dr. Peter Peer
dr. Tomaž Tomažič
dr. Matjaž Colnarič
dr. Slobodan Ribarič
dr. Vojko Flis
dr. Saša Divjak

dr. Stanislav Kovačič
dr. Zdravko Kačič
dr. Aleš Leonardis
dr. Simon Dobrišek
dr. Mitja Lenič
dr. Franc Solina
dr. Andrej Šoštarič
dr. Franjo Pernuš
dr. Jure Skvarč
dr. Božidar Potočnik

Organizacijski odbor 2018:

Gašper Sedej
Jurij Munda
Martin Šavc
Jernej Kranjec
dr. Danilo Korže
dr. Aleš Holobar
dr. Božidar Potočnik

POKROVITELJI



ROSUS 2018
<http://rosus.feri.um.si>

VABLJENA PREDAVANJA

OD 20 TISOČ DO 20 MILIJONOV

Klemen Forstnerič, Rok Ajdnik

Reveel Technologies, Inc.

E-pošta: developers@reveel.it

URL: <https://reveel.it/>

POVZETEK: *Prodreti v Silicijevo dolino s kakršnokoli idejo iz Slovenije je težko, še težje pa je razviti sistem za prepoznavo več kot milijona slik. Vsak algoritem ima skromne začetke, ki jih spremljajo neuspehi. Ponavadi razvoj takšnih algoritmov temelji na izkustveni metodi. Tudi Reveel ni tukaj nobena izjema. Razvili smo algoritem, ki je deloval odlično na bazi 20000 slik, pojavili so se pa problemi s hitrostjo strežniške poizvedbe ter obdelovanjem sočasnih poizvedb, ki so onemogočali razširitev baze na milijon slik. Izkazalo se je, da smo lahko z rešitvami iz področja porazdeljenih sistemov rešili ta problem.*

1. UVOD

Pri Reveel-u [1] gradimo rešitve, ki omogočajo založnikom in oglaševalcem, narediti njihove vizualne medije interaktivne. Ena izmed rešitev je sistem za prepoznavo slik, ki omogoča, da povežemo fizični (tiskani mediji) in digitalni svet (mobilna aplikacija, spletna stran). Tiskani mediji so sami po sebi neinteraktivni, s pomočjo prepoznave slik v tiskanih medijih pa uporabnik dobi interaktivno vsebino na spletu pri čemer uporablja svojo priljubljeno mobilno napravo ali računalnik. Seveda se problem, ki ga rešujemo sedaj razlikuje od problema in ideje zaradi katere smo ustanovili podjetje. Prvotna ideja je bila, kako poenostaviti nakupovanje preko televizije. Razvijali smo, mobilno aplikacijo, ki bi uporabnikom omogočala, da med gledanjem filmov ali TV oddaj hitro in enostavno kupijo karkoli vidijo na ekranu. Tehnični problem, ki smo se ga lotili je bil, kako z uporabo senzorjev v pametnem telefonu odkriti, kaj uporabnik gleda na televiziji, platnu ali monitorju.

2. ZAČETKI

Med raziskovanjem smo ugotovili, da lahko problem rešimo na tri različne načine. Od uporabnika bi lahko zahtevali, da sam poišče in izbere vsebino, katero gleda. Tako bi problem rešili izključno s standardnimi rešitvami uporabniških vmesnikov. Takšna rešitev bi zahtevala najmanj razvojnega časa, ampak bi z vedno več vsebine v sistemu postala mučna za uporabnika. Naslednja rešitev, ki smo jo raziskali je uporaba avdio prepoznave, ki bi omogočala prepoznavo video vsebin preko njihovih avdio zapisov. Prednost tega pristopa je manjša baza začetnih podatkov, saj je avdio zapis videa manjši kot njegov video zapis. Druga prednost z uporabniške izkušnje je lažja uporaba, saj lahko mobilno napravo držimo v udobnem položaju medtem ko prepoznavava avdio. Pri video prepoznavi moramo

kamero mobilne naprave usmeriti proti videu, da lahko le-ta uspešno izvede prepoznavo. Zadnja rešitev pa je video prepoznavna. Glavna prednost tega pristopa je, da ima veliko večji potencial za prihodnost, saj ni omejen samo na video vsebino ampak lahko tehnologija razpozna tudi revije in ostale tiskane medije. Medtem ko imata oba pristopa, avdio in video razpoznavna, svoje prednosti in slabosti, je bil glavni faktor zakaj smo se odločili za video razpoznavo, da bi se odtujili od naše konkurence.

2.1 Digitalni vodni žig

Prvotna ideja je bila uporaba digitalnih vodnih žigov - "digital watermark" [2] v videih. V video smo zakodirali digitalne vodne žige, s pomočjo kamere na mobilni napravi zajeli sliko, dekodirali digitalni vodni žig in tako prepoznali video. Digitalni vodni žig je postopek, kjer v celotno sliko zapišemo podatke. Pri tem je pomembno, da je popačenje neopazno človeškemu očesu, vendar vidno dekodirnemu algoritmu, ki iz slike izračuna N bitno kodo, ki unikatno identificira vsebino. Seveda se digitalnega vodnega žiga ne zaznava na originalni sliki, ampak na sliki, ki je zajeta s kamero mobilne naprave in vsebuje to originalno sliko. Zato mora algoritem imeti visoko toleranco na šum. Visoko toleranco na šum je možno doseči z uporabo visoke redundance podatkov med kodiranjem digitalnega vodnega žiga, ampak to pa pomeni, da smo omejeni s količino podatkov, ki se lahko zakodirajo v sliko. Vnašanje digitalnega vodnega žiga v video vsebine poteka tako da skozi vse slike videa zakodiramo N bitno kodo, kjer prvih X bitov unikatno identificira video vsebino, drugih Y bitov pa časovno lokacijo v vsebini. Ob zajemanju slike videa preko mobilne naprave se najprej izvede registracija robov televizije ali monitorja, perspektivna korekcija zajete slike in nato ekstrakcija podatkov digitalnega vodnega žiga. Naš prvi prototip je bil algoritem opisan v članku Meerwald et.al. [3], s katerim smo lahko zakodirali 56 bitov v posamezno sliko, kjer je prvih 32 bitov identificiralo vsebino, preostalih 24 bitov pa je predstavljalo časovno lokacijo v vsebini. Algoritem za vstavljanje digitalnega vodnega žiga uporablja diskretno valjčno transformacijo s katero ob uporabi "spread spectrum" kodiranja zakodira podatke v originalno sliko. Ob testiranju smo ugotovili, da je učinkovitost algoritma zelo odvisna od uspešnosti registracije robov televizije ali monitorja. Registracijo robov smo izvajali tako, da smo uporabljali kombinacijo več algoritmov. Najprej smo izvedli Canny za detekcijo robov na sliki, Hough transformacijo za iskanje premic, nato pa smo poiskali najbolj primerna oglišča glede na kriterij, da morajo imeti presečišča notranji kot okoli 90 stopinj. Preden bi lahko nadaljevali razvoj in raziskave smo ugotovili, da naša rešitev ni najbolj primerna za trg v katerem smo želeli lansirati našo rešitev. Rešitev zahteva spreminjanje video vsebine, saj moramo vanjo vstaviti digitalni vodni žig, kar pa bi pomenilo, da bi morali zakupiti licenco za vsako vsebino v katero bi želeli vstaviti digitalni vodni žig. S stališča marketinga in prodaje se je ta pristop izkazal za neizvedljivega oziroma zelo dragega.

2.2 Značilke, gručenje in uteži

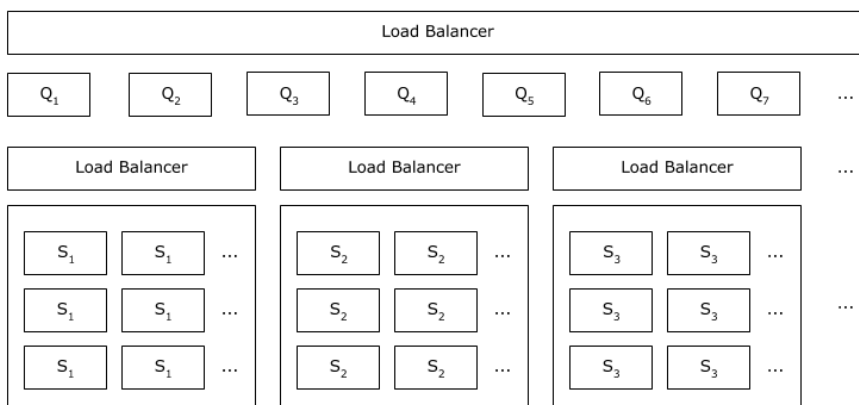
Zaradi zahtev trga smo bili primorani spremeniti naš pristop in ugotoviti kako prepoznati vsebino ne da bi jo spreminjali. Med raziskovanjem smo naleteli na področje algoritmov s skupnim imenom "Content-based Image Retrieval" ali s kratico CBIR [4]. Ti algoritmi

delujejo tako, da najprej ustvarijo bazo slik in nato omogočajo iskanje nad to bazo. Pri iskanju najdejo v bazi sliko, ki se najbolj ujema z iskano sliko. Pri implementaciji naše CBIR rešitve smo se odločili uporabiti pristop, kjer smo iz slik izločili značilke in nato uporabili te značilke pri gradnji baze in iskanju. Pri izbiri značilk smo iskali takšne, ki bodo robustne na popačenja in šum, ki nastanejo na sliki, ko iskano sliko zajamemo s kamero mobilne naprave. Druga zahteva, ki smo jo imeli zaradi omejitev prepoznavanja videov pa je bila hitrost algoritma. Celotna poizvedba se mora na strežniku izvesti hitro, da zagotovimo, da bo video vsebina prepoznana v doglednem času za uporabnika. Zaradi teh omejitev smo se odločili za uporabo značilk ORB [5]. Algoritem je dovolj robusten na transformacije in šum ter kljub vsemu zelo hiter. Predvsem pa nas je prepričalo dejstvo, da algoritem ni licenciran in se ga lahko prosto uporablja. Nato je nastopil drug problem in sicer kako uspešno izvesti iskanje nad bazo značilk. Če poenostavimo, imamo dve množici značilk, prvo množico predstavljajo značilke, katere smo izračunali iz vseh slik, ki jih imamo v bazi, drugo množico pa predstavljajo značilke, katere smo izračunali iz slike s katero izvajamo iskanje. V praksi druga množica vsebuje med 1000 in 2000 značilk, prva množica pa okoli 28 milijonov značilk. Hitro se je izkazalo, da je enostavno linearno iskanje skozi bazo neizvedljivo, predvsem zaradi časovnih omejitev, ki jih imamo pri poizvedbah. Ob raziskovanju rešitev, ki zadostujejo vsem omejitvam smo naleteli na skupino algoritmov FLANN [6], ki omogočajo hitro iskanje skozi velike baze binarnih ali vektorskih podatkov. Značilke ORB so v binarnem formatu, zato smo se odločili za uporabo "Hierarchical Clustering" dreves [7]. S pomočjo FLANN pristopa smo lahko hitro poiskali vse značilke, ki se najbolj ujemajo z iskano množico značilk. Pri tem pa smo naleteli na drug problem, ki nastane pri zajemu slike z mobilno napravo - da se večina značilk originalne slike spremeni kljub robustnosti ORB algoritma. Ta problem nas je privedel do ideje o implementaciji gručenja, kjer smo vse značilke posplošili na nek vnaprej definiran nabor značilk. Ta pristop je pospešil iskanje skozi indeks FLANN in rešil problem značilk, ki so popačene zaradi šuma in artefaktov. Po več mesecih poskusov in neuspehov smo našli članek Sivic et. al. [8], ki nas je usmeril v pravo smer in postavil temelje trenutnem algoritmu. Članek opisuje podoben pristop, kot smo ga sami ubrali vendar predstavi nekaj ključnih rešitev, ki doprinesejo k delujočemu sistemu za razpoznavo. Članek opisuje indeks FLANN, ki je zgrajen iz milijona vnaprej izbranih značilk, s katerim se zgradi podatkovna baza. Med iskanjem se z uporabo indeksa FLANN iz baze poišče podobne slike, ki se nato s TF-IDF [9] utežmi zmanjša na manjšo množico. Nad to množico se nato še linearno izvede geometrična verifikacija ključnih točk z algoritmom RANSAC [10]. S tem algoritmom lahko v manj kot 300 milisekundah izvedemo iskanje skozi bazo 20000 do 50000 slik, odvisno od vsebine slik. Ta sistem za prepoznavo slik smo uporabljali v našem produktu dobro leto in pol, dokler se niso pojavile prepreke zaradi katerih smo morali nadgraditi naš sistem.

3. NADGRADNJA SISTEMA

Sistem, ki je bil načrtovan in implementiran pod velikim časovnim pritiskom zaradi pridobivanja investicije, je sčasoma začel kazati znake degradacije. Sprva smo prepoznavali samo video vsebine, tekom delovanja pa smo v sistem vnašali vse več slik strani revij in ostalih tiskanih medijev, ki pa imajo drugačne karakteristike kot video.

Pojavili so se problemi z dostopnostjo storitve v času povečanega obiska uporabnikov ter v primeru dodajanja nove vsebine v sistem. Časi obdelave posameznega zahtevka so se iz manj kot 300 milisekund povišali na 5 in več sekund, kar je za uporabnika nesprejemljivo. Potrebno je bilo drastično zmanjšati iskalni prostor po katerem išče strežnik, obenem pa obdržati vse obstoječe slike v sistemu. Ena izmed rešitev za ta problem, je implementacija distribuiranega sistema, pri katerem vsi strežniki nimajo dostopnih vseh podatkov v sistemu. Ločili smo strežnik na dva dela, in sicer na strežnik, ki sprejema poizvedbe od uporabnika - “query”, ter na strežnik, ki hrani trenutno dostopne slike - “storage”. Strežnikov “storage” je običajno več, in vsak ima le del vseh slik trenutno v sistemu. S tem pristopom učinkovito zmanjšamo čas, ki je potreben za obdelavo poizvedbe. Ko dobimo poizvedbo, se ta pošlje na “query”, ki dalje paralelno posreduje to poizvedbo s pomočjo protokola RPC [11] na vse “storage” strežnike. Vsak izmed “storage” strežnikov vrne identifikator najboljše slike, ki jo ima - v primeru da ta obstaja - ter numerično utež, ki ponazarja, kako prepričan je, da je ta slika prava. Strežnik “query” nato rangira rezultate in vrne uporabniku najboljšega. Tak pristop se, sodeč po testih obnese bistveno bolje od starega pristopa. Vseeno pa ima še en problem. V primeru, da zaradi katerega koli razloga en strežnik postane nedostopen, uporabniki več ne morejo dostopati do dela slik. Rešitev je, da vsakega izmed “storage” strežnikov še dodatno N-krat repliciramo s čimer povečamo redundanco za ta del slik v primeru, da katerikoli od teh strežnikov postane nedostopen. S tem pristopom resda učinkovito rešimo problem z nedostopnostjo, potrebno pa je poudariti, da pride s tem do dodatnih problemov. Pomembno je namreč, da se vsi strežniki, ki hranijo določen del slik, strinjajo o tem kakšen je ta del, ki ga hranijo. Pridemo torej do problema konsenza, ki je znan problem v distribuiranih sistemih. Obstaja ogromno algoritmov, ki rešujejo ta problem, najbolj znan med njimi je Paxos [12], ki pa je znan po tem, da ga je v realnem svetu zelo težavno implementirati. Zaradi tega dejstva, smo se odločili, da uporabimo algoritem Raft [13], ki je bil načrtovan tako, da je enostaven za implementacijo. Arhitekturno gledano smo naš sistem posodobili iz monolitičnega, ki je vseboval funkcionalnost “query” in “storage” v arhitekturo, ki je prikazana na Slika 1.



Slika 1: Arhitektura nadgrajenega sistema za prepoznavo slik, kjer so Q ti. query strežniki in S ti. storage strežniki.

4. PRIHODNJI IZZIVI

Kljub vsem izboljšavam imamo še veliko za postoriti, preden bo naš sistem zmozel prepoznati 20 milijonov slik. Od razvoja dodatnih funkcij, ki bodo sistemu razširile zmožljivost, do izboljšav že obstoječih algoritmov. V nadaljevanju bomo predstavili nekaj problemov, ki jih še nismo rešili in opisali naše pristope k njihovim rešitvam. Glede na to, da se Reveel vseskozi prilagaja različnim trgom v katere vstopamo, se bodo zahteve sistema za prepoznavo nedvomno spremenile, kar pomeni, da se bomo v prihodnosti srečali s problemi, na katere še nismo pomislili.

4.1 Predpomnjenje

Naš sistem v povprečju posamezen zahtevek obdela v 300 milisekundah. Čeprav je to zelo hitro, je smiselno, da vsebino, do katere uporabniki pogosteje dostopajo, shranimo v predpomnilniku, saj lahko na tak način do nje dostopamo še hitreje in zmanjšamo obremenitev strežnikov. Načinov kako zasnovati predpomnilnik je več in ena izmed teh je uporaba zgoščevalnih funkcij - "hashing functions". V našem primeru imamo namen uporabiti percepcijske zgoščevalne funkcije - "perceptual hashing functions" [14], ki so zmožne zmanjšati dimenzionalnost slik na nekaj zlogov in vseeno ohraniti osnovne informacije o sliki. To pomeni, da si bosta rezultata zgoščevalne funkcije dveh podobnih slik prav tako podobna. Z uporabo zgoščevalnih funkcij in dreves "Hierarchical Clustering" lahko nato zgradimo predpomnilnik, ki bo lahko hitro poiskal, če se kakšna poizvedba ujema s prejšnjimi poizvedbami in namesto, da ponovimo iskalni algoritem lahko vnaprej vnemo rezultat, ki se je izračunal pri prejšnji poizvedbi.

4.2 Geometrična verifikacija

Analiza korakov našega algoritma za prepoznavo nam je pokazala, da se vsaj 50% časa ob vsaki poizvedbi porabi pri geometrični verifikaciji ključnih točk, ki jo rešujemo z algoritmom RANSAC. To pomeni, da je izboljšava tega koraka ključna za celostno pohitritev algoritma. Od leta 1980, ko je izšel članek za RANSAC, se je pojavilo ogromno izboljšav, ena izmed boljših trenutno je Graph-Cut RANSAC [15]. GC-RANSAC uporablja principe iz teorije grafov in v enem izmed korakov iz ključnih točk zgradi graf pretočnega omrežja oz. "flow network" nad katerim nato izvede reze s čimer pohitri iskanje vključenih točk oz. "inliers". Sodeč po članku, pričakujemo približno 40% pohitritev na račun zamenjave algoritma.

5. ZAKLJUČEK

Štiri leta nazaj smo začeli z enostavno idejo, kako poenostaviti nakupovanje gledalcem malih in velikih ekranov. Tekom let se je ta ideja spremenila mnogokrat in z njo se je tudi spremenila naša rešitev. Dokaj hitro smo ugotovili, da je neuspeh naš najboljši učitelj in da bomo do prave rešitve prišli samo, če smo pripravljeni mnogokrat spodleteti.

LITERATURA

1. <https://reveel.it/>
Reveel - Make all your media interactive so you can invite audiences to explore, learn and buy.
2. L. K. Saini, V. Shirvastava (2014), A Survey of Digital Watermarking Techniques and its Applications, *International Journal of Computer Science Trends and Technology*, vol. 2, str. 70-73.
3. P. Meerwald, A. Uhl (2010), Watermark Detection for Video Bookmarking Using Mobile Phone Camera, *Communications and Multimedia Security, Lecture Notes in Computer Science*, vol. 6109, str. 64-74.
4. W. Zhou, H. Li, Q. Tian (2017), Recent Advance in Content-based Image Retrieval: A Literature Survey, <https://arxiv.org/abs/1706.06064v2>.
5. E. Rublee, V. Rabaud, K. Konolige, G. Bradski (2011), ORB: an efficient alternative to SIFT or SURF, *Proceedings of the IEEE International Conference on Computer Vision*, str. 2564-2571.
6. M. Muja, D. G. Lowe (2009), Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration, *International Conference on Computer Vision Theory and Applications*.
7. M. Muja, D. G. Lowe (2012), Fast Matching of Binary Features, *Conference on Computer and Robot Vision*.
8. J. Sivic, A. Zisserman (2003), Video Google: a text retrieval approach to object matching in videos, *Proceedings of the Ninth IEEE International Conference on Computer Vision*, vol. 2.
9. H. Wu, R. Luk, K. Wong, K. Kwok (2008), Interpreting TF-IDF term weights as making relevance decisions, *ACM Transactions on Information Systems*, vol. 26.
10. S. Choi, T. Kim, W. Yu (2009), Performance Evaluation of RANSAC Family, *Proceedings of the British Machine Vision Conference*, vol. 24, str. 1-12.
11. <https://tools.ietf.org/html/rfc5531>
RFC 5531 - RPC: Remote Procedure Call Protocol Specification Version 2
12. L. Lamport (1998), The Part-Time Parliament, *ACM Transactions on Computer Systems* 16, vol. 2, str. 133-169.

13. D. Ongaro, J. Ousterhout (2014), In search of an understandable consensus algorithm, *Proceeding USENIX ATC'14 Proceedings of the 2014 USENIX conference on USENIX Annual Technical Conference*, str. 305-320.
14. C. Zauner (2010), Implementation and Benchmarking of Perceptual Image Hash Functions, *Master's thesis, Upper Austria University of Applied Sciences, Hagenberg Campus*.
15. D. Barath, J. Matas (2017), Graph-Cut RANSAC, <https://arxiv.org/abs/1706.00984>.

IZDELAVA MODELA ČLOVEŠKEGA ORGANA ZA TRIDIMENZIONALNO TISKANJE NA PODLAGI MAGNETNO-RESONANČNIH SLIK

Andrej Vovk

Center za klinično fiziologijo
Medicinska fakulteta, Univerza v Ljubljani
E-pošta: andrej.vovk@mf.uni-lj.si
URL: <http://www.mf.uni-lj.si/>

POVZETEK: *Mnoge bolnišnice po svetu že uporabljajo tridimenzionalne (3D) modele za načrtovanje in optimizacijo kirurškega pristopa, trening izvedbe operacije, klinično izobraževanje ter za personalizirane implante. Izdelava natisnjenih tridimenzionalnih modelov je kompleksen postopek in zahteva znanje iz različnih področij. Prvi korak je zajem slike s čim boljšo prostorsko ločljivostjo in ustrezno modalnostjo, da dosežemo dober kontrast opazovanega organa glede na sosednja tkiva. V medicini se za vpogled v strukturo človeškega telesa uporabljajo različne tehnike. Zaradi neinvazivnosti in prostorske natančnosti je najprimernejše orodje za slikanje mehkih tkiv magnetno resonančni tomograf, ki omogoča zajem prostorskega elementa od velikosti 0.5mm v vseh treh ravninah. Z računalniško obdelavo lahko izoliramo in prikažemo posamezni organ ali del telesa. Te metode razmejivne oz. klasifikacije posameznih prostorskih elementov (vokselov) imenujemo segmentacija in ta je lahko avtomatska, polavtomatska ali ročna. Glede na podobnost struktur med posameznimi preiskovanci lahko pri segmentacijah uporabljamo tudi anatomske predloge, ki služijo kot pomoč pri določevanju verjetnosti pripadnosti posameznega voksla različnim strukturam. Nadaljni korak pri pripravi modela za 3D tiskanje je pretvorba segmentiranega volumna v mrežni oziroma ploskovni model, ter po potrebi še obdelava v posebnih oblikovalskih programih. Glede na izbrano tehnologijo tiskanja se v naslednjem koraku ob določanju poti tiskanja, definira še debelino tiskanja rezin, debelino sten, hitrosti tiskanja, ali so potrebne podporne strukture za previsne dele, uporaba različnih barv in materialov, itd. Kvaliteta končnega izdelka 3D tiskanja je odvisna od vseh omenjenih korakov, ter nenaadnje tudi od kvalitete in tehnologije 3D tiskalnika. Različne tehnologije 3D tiskanja namreč omogočajo uporabo različnih bolj ali manj biološko kompatibilnih materialov.*

1. UVOD

Začetki 3D tiskanja sežejo v začetek 80-ih let 20. stoletja, ko je Chuck Hull v ZDA patentiral stereolitografski postopek. Šele zadnja leta je tehnologija ob zmogljivejših računalnikih, s potekom prvih patentov tiskanja v treh dimenzijah, iznajdbo novih metod in materialov za tiskalnice 3D modelov, ter z izvedbo namiznih tiskalnikov, dosegla velik

vzpon in s cenovno ugodnimi rešitvami dosegla široko dostopnost. Sprva se je 3D tiskanje, zaradi visoke cene naprav, uporabljalo le v industriji za hitro izdelavo prototipov. Z razvojem novih metod tiskanja z različnimi materiali so področja uporabe od tekstilne in obutvene industrije, avtomobilske industrije, gradbeništva in arhitekture, modnega oblikovalstva, robotike, izobraževanja in znanosti, farmacije, do različnih medicinskih implantov in protez, ter celo tiskanja živih tkiv in hrane.

Čeprav se zdi znanstvena fantastika, je podjetje Organovo (Univerza Missouri, ZDA) že leta 2009 razvilo Bioprinter in natisnilo žile iz celic pacienta, Švedsko podjetje CELLINK tiska ušesa z živim tkivom, Nemško podjetje Laser Center iz Hannoverja je 2010 razvilo lasersko tiskanje kožnih celic.

Mnoge bolnišnice po svetu že uporabljajo tridimenzionalne (3D) modele za načrtovanje in optimizacijo kirurškega pristopa, trening izvedbe operacije, klinično izobraževanje ter za personalizirane implante. Tehnologija tomografskega slikanja in nato 3D tiskanja omogočata tudi izdelavo personaliziranih implantov. Prvi znani uspešni primeri so vsaditev titanijeve spodnje čeljusti v Belgiji (2012)[1], implantata pelvisa iz titanove zlitine v Veliki Britaniji in plastične traheje pri ameriškem dojenčku (2013)[2]. Ne konča pa se zgodba samo s tiskanjem nadomestnih personaliziranih kosti; 3D natisnjene mrežaste titanijeve nadomestne kosti namreč tudi prevlečejo z bioaktivnimi agenti, ki pospešujejo rast kosti. Največja uporaba personaliziranih 3d natisnjenih modelov se predvideva za slušne pripomočke in v dentalni industriji [3].

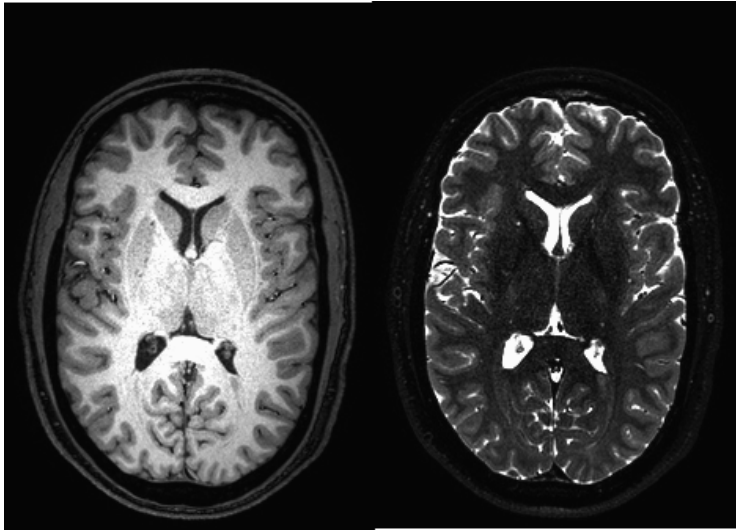
Še veliko zanimivih primerov uporabe 3d tiskalnikov na področju medicine je opisanih v preglednih člankih kot sta Three-Dimensional Printing and Medical Imaging: A Review of the Methods and Application (Marro, 2016)[4] in članek Medical Applications for 3D Printing: Current and Projected Uses (Ventola, 2014)[5].

2. IZDELAVA 3D MODELA ČLOVEŠKEGA ORGANA ZA UČNE NAMENE

Proces izdelave 3D modela zajema naslednje korake: zajem 3D slik, segmentacija, pretvorba v ploskovni model, določanje specifik tiskanja in samo 3D tiskanje.

2.1 Zajem 3D slik

V medicini se za natančen vpogled v strukturo človeškega telesa uporabljata tomografski tehniki računalniška tomografija (CT) in magnetna resonanca (MR). CT je primeren predvsem za slikanje kosti in nekaterih s kontrastom poudarjenih mehkih tkiv zaradi hitrosti slikanja. Poleg tega, da je MRI neinvazivna metoda (za razliko od CT-ja), omogoča slikanje z boljšo resolucijo od 0.5mm x 0.5mm x 0.5mm navzgor in z različnimi modalnostmi lahko brez kontrastnega sredstva poudarimo različne strukture. Seveda lahko tudi pri MR slikanju uporabimo kontrastna sredstva za slikanje posebnih podrobnosti npr. ožilje.



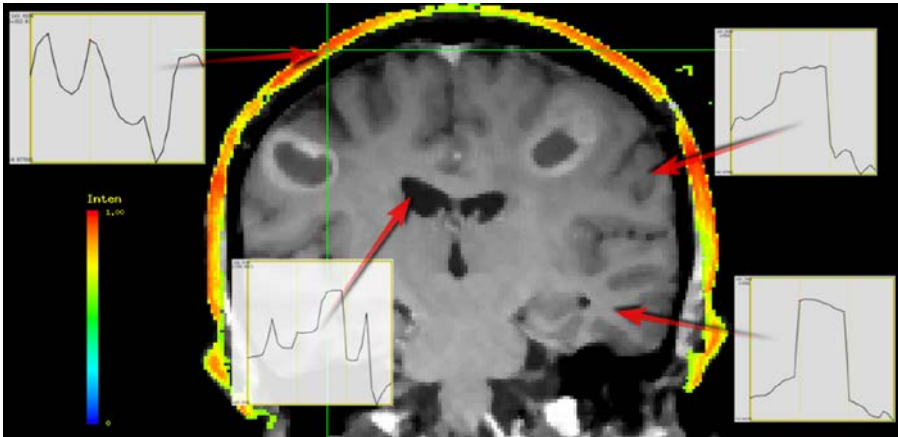
Slika 1: MR slika človeških možganov T1 modalnosti (desno) in T2 modalnosti (levo).

2.2 Segmentacija

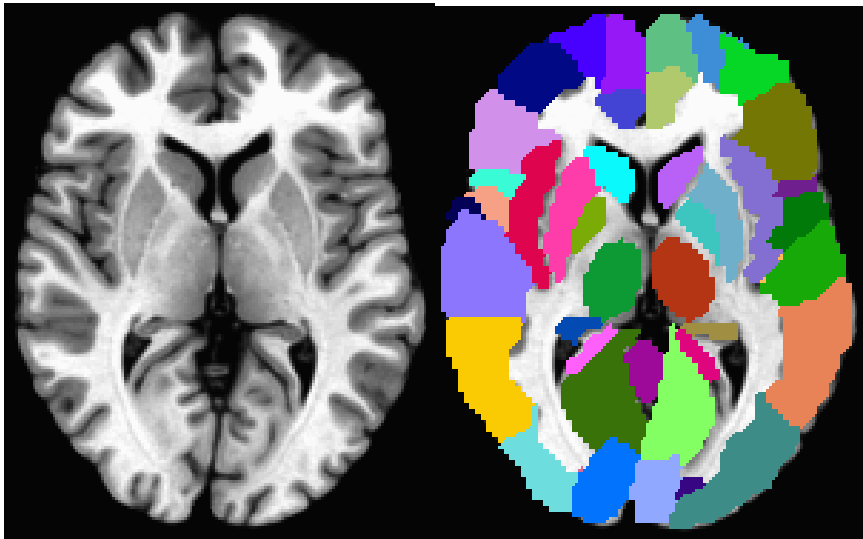
Iz 3D slik - volumnov, nato z eno od segmentacijskih metod razmejimo volumen na področja; oziroma 3D točke - voksle, ki sestavljajo volumen, glede na odtenke sivine porazdelimo v različna področja. Segmentacija je ključni del pri prepoznavanju in določevanju mej naše opazovane strukture. Segmentacijske algoritme lahko razdelimo na 8 kategorij [6]: metode pragov, metode rasti območij, metode uvrščanj (klasifikacije), metode razvrščanj, model Markovskih naključnih polj, nevronske mreže, deformacijske modele in metode, ki uporabljajo prostorske atlase. V praksi se običajno uporabljajo kombinacije teh pristopov, ki so združene v učinkovitejši segmentacijski algoritem.

Način segmentacije z razvrščanjem lahko izvedemo tudi z metodo pridobivanja informacij iz 3D teksture slike na posamezni lokaciji in na več prostorskih lestvicah okoli te lokacije. Tak skupek informacij o teksturah na več prostorskih lestvicah imenujemo podpis voksla. [7]

Če nam metoda slikanja zagotovi dober kontrast opazovane strukture, lahko metodo segmentacije izvedemo enostavno z določevanjem spodnjega in zgornjega nivoja sivine slike. Na tak način nam npr. odprtokodni programski paket 3Dslicer omogoča enostavno segmentacijo kosti iz CT slik.



Slika 2: Obarvana maska lobanje, temelji na podobnosti (korelaciji) podpisov s podpisom v izbranem vokslu v središču zelenega križa. Rdeče puščice kažejo vokse izvorov posameznih podpisov.



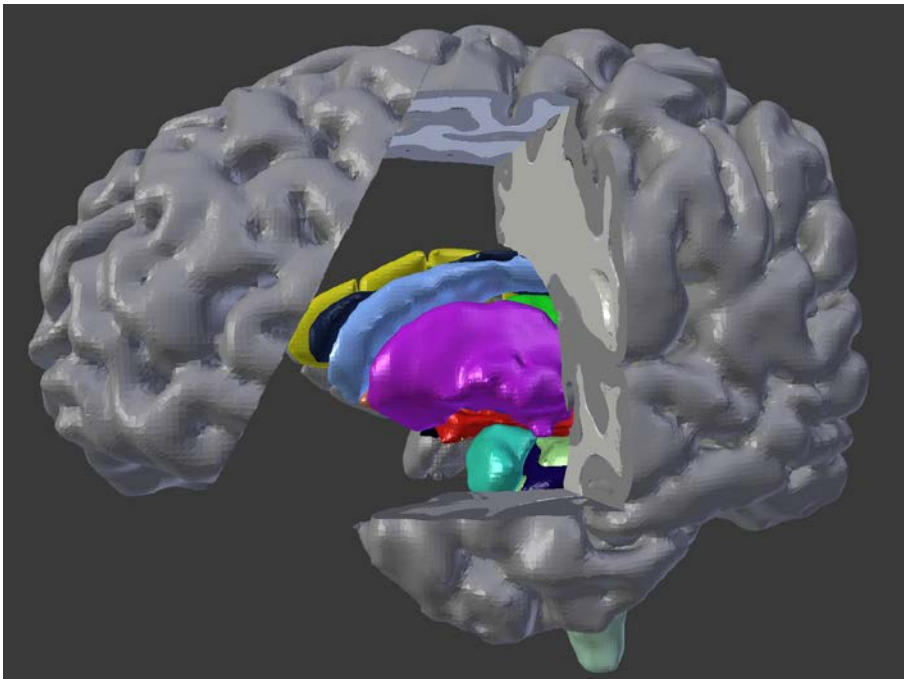
Slika 3: Talairch predloga (levo), MNI152 atlas poravnan na Talairch predlogo (desno).

Strukturno zahtevni organi, kot so možgani, zahtevajo kompleksnejše algoritme segmentacije, ki vključujejo tudi poravnavo slik z verjetnostnimi atlasii. Ti atlasii so izračunani kot povprečje večjega števila normalnih možganov. Poleg Talairach Atlasa, ki sicer "bazira" na posmrtni disekciji enih samih možganov, se za celotne možgane največkrat uporablja poravnava na MNI152 ali MNI 305 standardni prostor. Obstajajo pa še mnogi atlasii, ki na različne načine razmejujejo možgane; kortikalne subkortikalne

strukture – Harvard-Oxford atlas, glede na celične in mielinske strukture – Julich histološki atlas, razmejevanje beline glede na mielinske poti – JHU atlas, talamična področja, področja malih možganov, itd. Obstajajo tudi atlasi za različne starosti možganov, saj vemo, da se možgani razvijajo še v najstniškem obdobju in za natančno segmentacijo je nujna poravnava na primerno predlogo.

V našem centru uporabljamo za segmentacijo možganov odprtokodni programski paket FreeSurfer [8], ki vsebuje nabor programskih orodij za raziskovanje anatomije možganske skorje in subkortikalnih struktur. Proces segmentacije volumnov uporablja, po začetnih popravkih nehomogenosti in variacije intenzivnosti, večdimenzionalno nelinearno poravnavo na atlas MNI305. Končna segmentacija se izvede na podlagi predvidenih individualnih karakteristik iz atlasa in specifičnih izmerjenih vrednosti. Prednost programa FreeSurfer je avtomatiziranost pri visoki stopnji ponovljivosti, vendar postopek segmentacije traja na večprocesorskem računalniku od 10 do 40 ur.

2.3 Pretvorba v ploskovni model



Slika 4: Obdelava ploskovnega objekta v programu Blender.

Segmentirane strukture/volumne, ki so sestavljena iz vokslov nato v naslednjem koraku pretvorimo v ploskovni model, kakršnega potrebujemo za 3D tiskanje. Ploskovni model namreč omogoča hitrejše tiskanje in bolj ekonomično izrabo materiala za tiskanje. Programski paket FreeSurfer vsebuje enostavno orodje `mri_tessellate`, ki na robnih točkah izbrane strukture naredi mrežo oziroma ploskve. To ploskovno strukturo lahko nato še

zgladimo z ukazom `mris_smooth`. Malo kompleksnejši program za izdelavo ploskovnih struktur je `3dVol2Surf` iz prav tako odprtokodnega programskega paketa AFNI [9]. Ta program omogoča tudi določevanje skrajnih meja ploskovnega objekta, kar je zelo uporabno pri tiskanju sosednjih objektov, oz. objektov, ki se držijo skupaj. Na mrežnih strukturah lahko namreč zaradi ostrih robov ali lukenj pride do nepravilnosti, ki določijo ploskev skozi drugo strukturo. Prav tako lahko ploskev zaide v drug objekt pri glajenju površinskega modela.

Nadaljno obdelavo in popravke nepravilnosti kot so nezaključene ploskve ali preostri robovi smo izvedli v programu Blender.

Tudi program Blender je odprtokodni in prostodostopen na spletu. Je sicer zelo kompleksen program, saj je namenjen tudi kreiranju animacij, vendar na spletu obstaja mnogo učnih naptokov kako se lotiti obdelave ploskovnih objektov za 3D tiskanje.

2.4 3D tiskanje

Ko imamo model pripravljen in shranjen ponavadi v najbolj razširjenem standardnem STL formatu, je potrebno iz ploskev določiti pot tiskanja, debelino sten, pri previsnih predelih določiti podporne strukture in še kakšne specifikke za posamezne tiskalnike. Nekateri 3D tiskalniki imajo svoja programska orodja za pretvorbo iz STL v jezik tiskanja – G-code, obstaja pa tudi več odprtokodnih programov, ki podpirajo najrazličnejše 3D tiskalnike: Slic3R, Cura, IceSL, itd. Tudi na tem koraku se razvijajo novi pametni algoritmi, ki sami predlagajo glede na željeno hitrost tiskanja različne debeline rezin in različno gosto polnilo po zahtevnosti odsekov modela.

Kakorkoli fascinantno izgleda rojstvo iz digitalne v fizično obliko, je podobno kot v “realnem” življenju lahko zadnji korak zelo naporen in zakompliciran. Pri ekstruzijskih 3D tiskalnikih z nalagalno metodo tiskanja je pomembno, da je umerjena in ravna nalagalna površina, saj moramo zagotoviti, da se prvi sloj dobro sprime s podlago, drugače se tiskan model kasneje lahko odlepi in ob prerivanju tega po tiskalniku nam zraste ob nedokončanem modelu “grmovje”. Za boljši oprijem modela z nalagalno površino imajo danes že skoraj vsi 3D tiskalniki ogrevano nalagalno površino. Težave nam lahko povroči tudi zamašitev brizgalne šobe. Največkrat se to zgodi zaradi neenakomerne debeline materiala za tiskanje, ki ga kupimo na kolutih, seveda po različno ugodnih cenah. Nenazadnje je potrebno tiskalnik tudi vzdrževati, čistiti tako brizgalne šobe kot nalagalno površino, umerjati nalagalno površino, shranjevati material za tiskanje v primernem prostoru, ob primernem zračenju skrbeti za konstantno temperaturo. In dejstvo, da ekstruzijsko tiskanje traja dlje kot si mislite in kot napove tiskalnik ob začetku tiskanja, dostikrat pripelje do tega, da je tiskanje nenadzorovano, zaradi česar marsikatero omenjeno nezgodo prepozno opazimo.

3. ZAKLJUČEK

Zaradi popularizacije 3D tiskalnikov se ti nenehno izboljšujejo in končni korak postaja tudi pri cenovno najbolj ugodnih ekstruzijskih tiskalnikih čedalje prijetnejši. Področje 3D tiskanja je že tako popularno na najrazličnejših področjih, da obstaja celo revija z imenom 3DPrinting in Medicine (<http://threeedmedprint.springeropen.com/>). Ekonomisti in poznavalci razvijajočih tehnologij uvrščajo 3D tiskanje ob umetni inteligenci, nanotehnologiji, kvantnem procesiranju, biotehnologijam, ter robotiki, v dogajajočo se četrto industrijsko revolucijo. (The Fourth Industrial Revolution, professor Klaus Schwab). Čeprav je v našem prostoru mogoče premalo slišati o uporabi tehnologije 3D tiskanja in je zaenkrat to še prednost držnih, je po drugi strani vzpodbudno slišati, da se že v nekaterih osnovnih šolah učenci spoznavajo s to tehnologijo.

LITERATURA

- [1] Transplant jaw made by 3D printer claimed as first - BBC News 2015; Available from: <http://www.bbc.com/news/technology-16907104>.
- [2] Bartlett S, 2016. Printing organs on demand. *Lancet Respir Med.* 1(9):684.
- [3] 3D Printing in Medical and Dental Markets: An Opportunity Analysis and Ten-Year Forecast, Smartech; 2015. Available from: <http://smartechnpublishing.com/reports/3d-printing-in-medical-and-dental-markets-an-opportunity-analysis-and-ten-y>.
- [4] Marro A, et al 2016. Three-Dimensional Printing and Medical Imaging: A Review of the Methods and Application. *Curr Probl Diagn Radiol.*
- [5] Ventola CL, 2014. Medical Applications for 3D Printing: Current and Projected Uses. P&T.
- [6] Pham DL, et al 2000. Current methods in medical image segmentation. *Annu Rev Biomed Eng.*
- [7] Vovk A, et al 2011. Uporaba statističnih podpisov za izdelavo verjetnostnih map zdravih in obolenih možganov - nov diagnostični pripomoček?. *Zdravniški vestnik.*
- [8] Fischl, B., Dale, A.M., 2000. Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc Natl Acad Sci U S A* 97, 11050-11055.
- [9] Cox, R.W., 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res* 29, 162-173.

DEMONSTRACIJSKA SEKCIJA
Zajem slikovnega materiala: Praktični nasveti
strokovnjakov

TEHNOLOGIJE ZAJEMA SLIKOVNIH INFORMACIJ

Jure Škrabar

Kolektor Group d.o.o.

E-pošta: jure.skrabar@kolektor.com

URL: <http://www.kolektorvision.com>

POVZETEK: *Slikovna informacija, kot podlaga za robustno izvedbo kvalitetnih merilnih sistemov, je ena izmed najpomembnejših sklopov izgradnje kvalitetnih merilnih sistemov. V sklopu delavnice je narejen kratek pregled različnih tehnologij zajema (2D, 3D) s poudarkom na razumevanju prednosti in slabosti posamezne tehnologije in njene uporabe v praksi.*

1. UVOD

Zajem slikovne informacije je najpomembnejši korak pri izgradnji sistemov strojnega vida v industrijskem okolju. Ponudniki komponent razširjajo nabor izdelkov z novimi tehnologijami in metodami zajema, nadgrajujejo obstoječe ter s tem višajo prag zmožljivosti. Ravno širok nabor različnih izdelkov z oglaševanimi zmožljivostmi, katerih v večini primerov niti ni mogoče primerjati med seboj, lahko hitro zavedejo uporabnika, da izbere komponento, katera ne ustreza specifični aplikaciji. Iz tega razloga je v nadaljevanju prispevka narejen kratek pregled tehnologij zajema slikovnih informacij, z namenom, da uporabnik izbere tehnologijo, katera najbolj ustreza specifični aplikaciji.

2. TEHNOLOGIJE 2D ZAJEMA

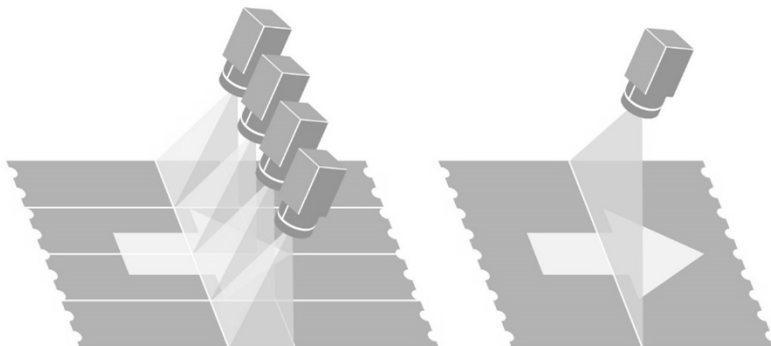
2.1 Matrični zajem slikovnih informacij v primerjavi z linijskim

Najbolj razširjena in iz stališča uporabe najbolj enostavna metoda zajema slikovnih informacij je s t.i. matričnimi senzorji. V primerjavi z linijskimi senzorji so ti precej bolj intuitivni za uporabo, saj so po načinu zajema najbližje človeški percepciji. Kljub enostavnosti metode zajema pa se sami senzorji lahko močno razlikujejo po lastnostih. Pri izbiri moramo biti pozorni predvsem na ločljivost, velikost slikovnega elementa in optično velikost senzorja, tehnologijo senzorja (CCD, CMOS) ter tipa zaklopke.

Linijski senzorji uporabljajo eno ali v izjemnih primerih več vrstični zajem slikovne informacije, katero kamere same že združujejo v 2D matriko podatkov oz. sliko. Slabost takšnega tipa zajema je, da se mora predmet zajema premikati pred kamero, da le-ta lahko zajame celoten objekt. Potrebna je dodatna sinhronizacija zajema ter premikanja objekta, kar v večini primerov tudi občutno poveča ceno končne aplikacije.

Velikost posameznega slikovnega elementa na linijskih senzorjih je v primerjavi z matričnimi senzorji večja, kar posledično prinese tudi boljšo občutljivost, zaradi enostavne

strukture sensorja pa so možne tudi večje frekvence zajema ter večje ločljivosti, le-te pa posledično prinesejo tudi večjo optično velikost sensorja. Linijski sensorji so najbolj primerni za kontrolo kontinuiranih, površinskih izdelkov ter bolje izkoristijo razpoložljiv prostor za vgradnjo.

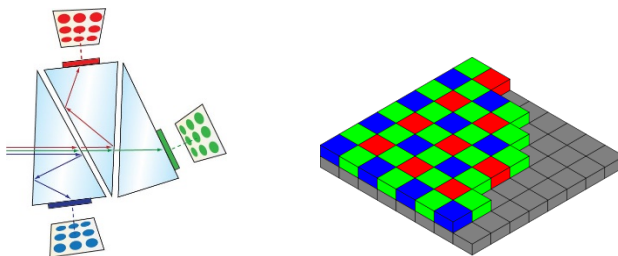


Slika 1: Matrični zajem 4x2k pix (levo) v primerjavi z linijskim zajemom 8k pix slikovne informacije (desno).

Ne glede na tip zajema slikovne informacije je ločljivost sensorja primarno pogojena s potrebno občutljivostjo za specifično aplikacijo. Po načelu izgradnje robustnih merilnih sistemov se predvidi občutljivost enote sensorja v velikostnem rangju 1/10 najmanjše značilke, katero mora sistem strojnega vida razločevati.

2.2 Monokromatski zajem slike v primerjavi z barvnim

V veliki večini aplikacij sistemov strojnega vida se uporabljajo monokromatski sensorji. Razlog za to je, da se uparja barvo osvetljevanja z optimalnim izražanjem značilke, kar posledično pomeni, da delujemo v omejenem spektru valovne dolžine svetlobe. V kolikor je za potrebe aplikacije nujno potrebna barvna informacija, le-te pa ni mogoče izraziti z uporabo fizičnih metod osvetljevanja ter filtriranja, se za izvedbo uporabljajo barvni sensorji.



Slika 2: Prizma in ločevanje na 3 sensorje pri 3CCD sensorjih (levo) ter Bayer vzorec mikrofiltrnov na barvnem sensorju.

Najboljša rekonstrukcija barv oz. zajem ter ločevanja posameznih spektrov je izvedljiva s 3CCD sensorji, kjer je za vsako izmed barv v barvnem spektru (RGB) uporabljen svoj sensor. Cenovno bolj učinkovita metoda je z uporabo t.i. Bayer vzorca, kjer ima

posamezen slikovni element dograjen tudi filter za barvo. Razporeditev mikro filtrov je povzeta po občutljivosti človeškega očesa (50% zelena, 25% rdeča, 25% modra). Barva posameznega slikovnega elementa je interpolirana na podlagi vrednosti sosednjih elementov, kar posledično niža realno efektivno ločljivost optičnega sistema za zajem slikovnih informacij.

2.3 Druge tehnologije zajema 2D slikovnih informacij

S spremembo tehnologije ter področja zaznavanja slikovnega elementa v spektru valovnih dolžin se izvajajo tudi meritve izven vidnega spektra (UV, NIR, IR, X-ray, ...). Posebnost so multi in hiperspektralni senzorji. Hiperspektralni senzorji združujejo digitalni zajem slik in spektroskopijo – za vsak slikovni element na sliki zajame hiperspektralni senzor svetlobno sevanje za veliko število spektralnih pasov. Vsak slikovni element tako vsebuje zvezen spekter (radiacije in odbite svetlobe) - informacija, na osnovi katere lahko prepoznamo objekt glede na njegove spektralne lastnosti.

Poleg vidnega dela spektra zajamejo hiperspektralne kamere tudi precejšen del IR spektra. Za veliko aplikacij so ravno značilnosti, kako objekt absorbira oz. odbija svetlobo v IR delu spektra ključne, da karakteriziramo oz. klasificiramo objekt. Pri aplikacijah kontrole kakovosti v industriji prihaja interes po tovrstnih kamerah iz farmacevtske industrije, industrije predelave hrane, industriji proizvodnje tankih filmov ipd.

Informacija pridobljena s takšnim senzorjem oz. t.i. hiper kocka nosi za vsak slikovni element tudi pripadajoči spektralni odziv, kar posledično pomeni ogromne količine podatkov za zajem, prenos in obdelavo.

3. TEHNOLOGIJE 3D ZAJEMA

3.1 Laserska triangulacija

Laserska triangulacija je najbolj razširjena metoda zajema 3D podatkov. Pri tej metodi se kot svetilni vir uporablja laserski projektor (črta), na katero je usmerjen pogled senzorja pod triangulacijskim kotom. Čeprav se v zadnjem obdobju na tržišču pojavljajo komponente, katerih izhodna informacija je oblak točk, se v večini primerov višinska informacija interpretira s sivinsko sliko, saj je le-ta bolj enostavna za obdelavo z algoritmi 2D obdelave slikovnih informacij.

Slabost laserske triangulacije je občutljivost odboja laserske črte ob robovih merjenja, senčenje kot posledica triangulacijskega kota ter nehomogenost laserske črte same. Zasnova merilnega sistema mora biti narejena na podlagi zavedanja omejitev, tako npr. s tem principom ni možna meritev precizna meritev izvrtine izdelka orientirane normalno na pregledovano površino.

3.2 Metode za zajem oblaka točk

Napredek na področju procesne zmogljivosti vgradnih naprav je v zadnjem obdobju privedel do komercializacije produktov s tehnologijami aktivnega ter pasivnega sterea,

Time of Flight, projekcijskih metod ter Light-field kamer. Vsaka izmed naštetih ima svoje prednosti in slabosti, med njimi so za izpostaviti projekcijske metode, katere dosegajo največjo občutljivost in natančnost.

Kljub izvedljivosti pridobivanja oblakov točk pa še vedno ostaja problematična manipulacija in obdelava podatkov znotraj njih. V tem trenutku je le-ta pogosto iterativna ter posledično časovno potratna. S konstantnim napredkom zmogljivosti procesnih enot je za pričakovati napredek tudi na tem področju.

4. ZAKLJUČEK

Tehnologije zajema slikovnih informacij so temelj robustne aplikacije sistemov strojnega vida. V sicer množici ponujenih komponent na tržišču se še vedno pogosto srečujemo z dilemo, ali je izbrana komponenta primerna. Poznavanje ozadja delovanja zajema slikovne informacije lahko v takšnih situacijah pomaga pri pravi izbiri, saj je končni cilj delujoča aplikacija v industrijskem okolju brez dodatnih posegov v sistem tekom in preko predvidene življenjske dobe delovanja.

LITERATURA

1. B. Dawson (2015), Advanced Color Machine Vision and Applications, Vision Conference Boston, MA
2. J. Škrabar (2017), Konstrukcija optičnih merilnih sistemov, Interna publikacija Kolektor Orodjarna d.o.o. – PC Vision

UPORABA TRŽNO DOSTOPNIH ALI NAMENSKIH SISTEMOV STROJNEGA VIDA

Rok Mihelj, Peter Lepej

VISTION d.o.o.

E-pošta: info@vition.si

URL: <http://www.vition.si>

POVZETEK: *Predstavljena bo primerjava med sistemi pametnih kamer, in namenskimi rešitvami strojnega vida, pomembnost osvetlitvenih tehnik pri izvedbi rešitev in končna podpora stranki. Pametne kamere dandanes že zajemajo vrsto knjižnic, ki jih je enostavno implementirati v proizvodni proces. Vsi ti sistemi se uporabljajo za enostavnejša merjenja ali odkrivanja napak, v primeru zahtevnejših nalog pa je za to potreben individualen pristop z namenskimi sistemi. Glede na današnjo zahtevnost proizvodnih procesov in izdelkov se je potrebno pravilno lotiti reševanja problematike procesa. Po končani izvedbi sistema pa naloga integratorja še ni končana, saj je potrebno uporabniku nuditi popolno tehnično podporo, to pa v svetu strojnega vida ni enostavna naloga.*

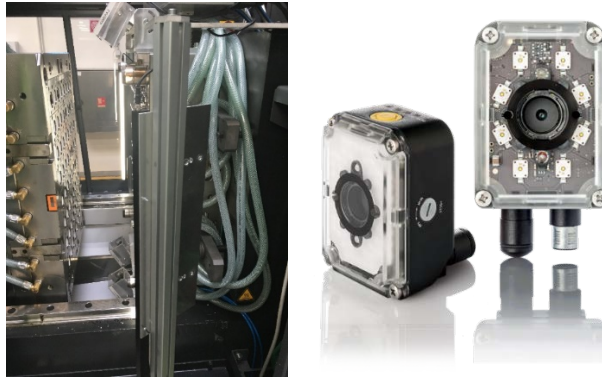
1. UVOD

VISTION sestavlja mlada, razvojna, dinamična in motivirana ekipa, ki je sinonim za ustvarjanje novih inovativnih rešitev na področju strojnega vida. Ukvarja se z integracijo znanih svetovnih ponudnikov slednjega in izdeluje namenske sisteme s svojo programsko opremo in svetili. Na trgu se želi uveljaviti kot najbolj inovativni ponudnik strojnega vida.

2. PRAVILNA IZBIRA SISTEMA STROJNEGA VIDA

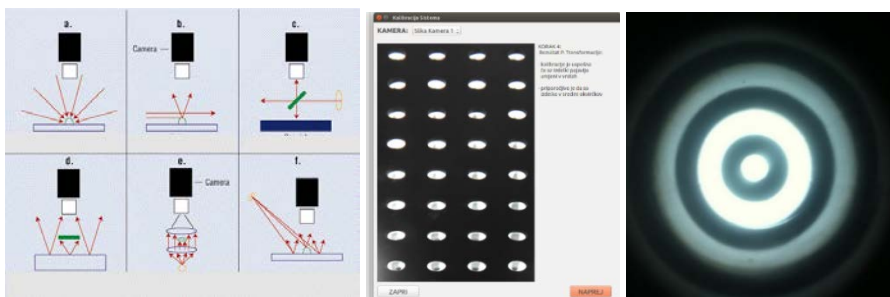
Pred nakupom strojnega vida je potrebno jasno določiti in opredeliti napake na izdelku, saj bodo slednje predstavljale temelj za izbiro strojne in programske opreme. Po natančni opredelitvi si je potrebno ogledati delujoči proizvodni proces, ter določiti prostor za umestitev opreme. Pri integraciji sistemov strojnega vida se lahko hitro soočimo s kar nekaj težavami, na primer tresljaji stroja, nenatančna pozicija kosa, prisotnost srha, omejena izbira svetil ... Vse te težave lahko privedejo do nepravilnega delovanja sistema ali do pojava nestabilnost, to pa pomeni, da je lahko zaznava nepravilna - kot slabe kose zazna tudi te, ki so brez napak ali obratno. Velikokrat uporabniki kar sami kupijo tako

imenovane »pametne kamere«, ki že vsebuje vrsto programskih knjižnic in jih skušajo integrirati v proizvodni proces. Kaj kmalu pa se zgodi, da sistem zaradi nezkušenosti uporabnika ali nepravilne zasnove sistema, ne deluje pravilno. Primerneje je, da podjetja poiščejo zunanjo pomoč integratorjev strojnega vida že v začetku.



Slika 1: Levo slika prikazuje integracijo sistema MOLDEYE INSPECTOR, desno slika pa prikazuje pametne kamere DATALOGIC.

Vsak potencialni naročnik sistema se mora zavedati, da sistema strojnega vida ne predstavljajo samo kamere, računalniki, objektivni in krmilja. Zelo pomembno vlogo (skoraj 50 %) predstavlja pravilna izbira osvetlitvene tehnike. Pravilna izbira svetil je tako ključnega pomena za uspešen razvoj sistema strojnega vida. Glavni namen osvetlitve je vzpostavitev svetlosti površine kosa, ki se mora razlikovati od ozadja. V primeru, da je kos težko ločiti od ozadja, je naloga strojnega vida zelo težka ali celo neizvedljiva. Drugi namen osvetlitve je zmanjšanje odseva in tretji izboljšanje kontrasta slike s osvetlitveno tehniko. Obstaja več vrst osvetlitvenih tehnik, kot so (Slik 2 levo): tehnika presvetlitve (Slika 2 desno), poševna/direktna, ko-aksialna osvetlitev in druge. Slika 1 prikazuje primer namenskega sistema, kjer so bila svetila prilagojena specifičnemu stroju (Slika 2 sredina) in desno vidimo pametne kamere z integrirano osvetlitvijo, kar predstavlja neuspešno rešitev v tem primeru.



Slika 2: Leva slika prikazuje tehnike osvetlitve strojnega vida - a.: Osvetlitev z sprednje strani, b.: Osvetlitev z stranske strani velik kot, c.: Ko-aksialno osvetljevanje, d.: Difuzna osvetlitev z zadnje strani, e.: Direktna osvetlitev z zadnje strani, f.: Osvetlitev z stranske strani mali kot . Srednja slika prikazuje pravilno kontrastno razliko kosa in ozadja sistema MOLDEYE INSPECTOR , na desni sliki pa je predstavljen primer presvetlitvene tehnike.

VISTION ponuja – poleg rešitev individualnih rešitev strojnega, še namenska svetila, ki so popolnoma prilagojena glede na uporabnikovo aplikacijo. Ključ do popolne aplikacije strojnega vida namreč temelji na individualnem sestavu kakršnega koli elementa. Ta pristop je ključen dejavnik za doseganje stabilnih in ponovljivih rezultatov obdelave slik. Svetila definiramo glede na zahteve izdelka:

- **SVETLOBNI SPEKTER:** 1. Rdečo, Zeleno, Modro, Belo barvo, 2. IR spekter, 3. UV spekter, ...
- **RAZLIČNE DOLŽINE IN ŠIRINE:** Glede na objekt lahko izbiramo različne dolžine in osvetljevalne površine. Vsaka svetilka se lahko prilagodi glede na zahteve uporabnikove aplikacije (Slika 3).
- **MOČI:** Moč svetilke je potrebno definirati v zgodnji fazi procesa izdelave svetil, ki je odvisna od osvetljevalne površine, odmika svetilke od objekta in materiala objekta. Na tej točki je priporočljivo testiranje za določitev najbolj optimalne osvetlitve.
- **OHIŠJA SVETIL:** Ohišja so lahko iz nerjavečega in aluminijastega materiala in s tem prilagojena zahtevam posamezne proizvodnje.
- **RAZPRŠENOST:** Na voljo so različna razpršitvena stekla, od matiranega za zelo odbojne površine, do prozornega za mat površine.



Slika 3: Prikazujeta svetila podjetja VISTION, ki so bila namensko izdelana glede na zahteve uporabnika.

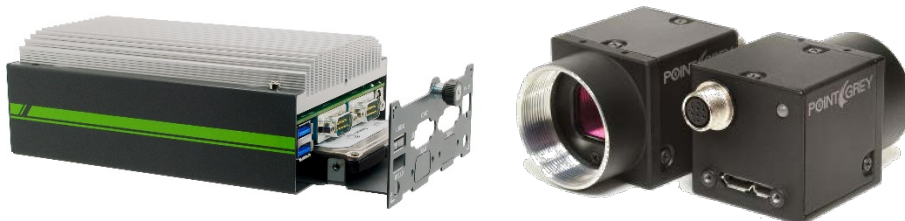
Ob vseh omenjenih svetilih, pa pomembno vlogo igra tudi svetloba, ki prihaja iz okolice. Proizvodnje obratujejo podnevi in ponoči, v oblačnem, deževnem ali sončnem vremenu, zato prihaja do različnih svetlobnih pogojev v sami notranjosti. Ti pogoji so za strojni vid

popolnoma nestabilni, kar lahko vodi do nestabilnosti sistema. Zaradi tega je najbolje, da se – kadar je to mogoče – ustvari lastne pogoje sistema strojnega vida; to pomeni, da se vzpostavi v prostoru s kontroliranimi pogoji.

V ponudbi obstaja kar nekaj različic sistemov strojnega vida, ki se med seboj razlikujejo po fleksibilnosti, zmogljivosti in ceni. Sistemi se lahko razdelijo v 2 kategoriji:

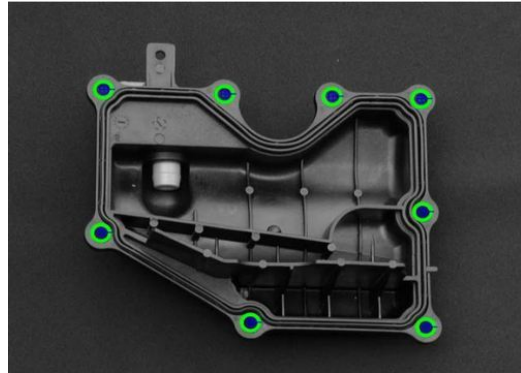
- Procesiranje z zunanjim industrijskim računalnikom
- Procesiranje neposredno na pametni kameri.

Zunanji industrijski računalnik: Klasični sistemi strojnega vida so sestavljeni iz industrijskega računalnika, ki upravlja in komunicira s perifernimi napravami, kot so kamere, svetila, itd. Ti sistemi omogočajo visoko računalniško moč, ter fleksibilnost, po slabost pa je, da zavzamejo več prostora kot pametne kamere. Priporočajo se tako za enostavne kot za zelo kompleksne aplikacije, kjer je potrebno opraviti kompleksnejšo nalogo z visoko hitrostjo obdelave podatkov.



Slika 4: Prikazuje zunanji industrijski računalnik, na katerega se priklopijo kamere, svetila, ekrani, itd.

Pametne kamere: Preprostejši in cenovno bolj dostopni sistemi temeljijo na pametnih ali inteligentnih kamerah, ki vsebujejo standardno optiko (pogosto z fiksno goriščno razdaljo leče) in svetila. Uporabljajo se za enostavne aplikacije kot je na primer prisotnost kovinskega vložka v izdelku. Že vnaprej imajo pripravljene programske knjižnice, ki jih posameznik grafično programira.



Slika 5: Leva slika prikazuje pametno kamero, desna pa zaznave vložkov.

Zmogljivost programske opreme: Če je strojna oprema označena kot "srce" strojnega vida, potem lahko rečemo da je programska oprema njena "kri". Programska oprema vsebuje dve ključni komponenti: algoritme za obdelavo slik in uporabniški vmesnik. Algoritmi programske opreme poiščejo potrebne informacije izdelka, z uporabniškim vmesnikom pa se te algoritme specifikira.



Slika 6: LEVA slika prikazuje uporabniški vmesnik podjetja VISION, DESNA pa algoritme za obdelavo slik.

Programska oprema strojnega vida se lahko uporablja v širokem spektru industrijskih aplikacij, predvsem za analiziranje in procesiranje slike, nadzor avtonomnih industrijskih vozil in druge. Zmožna je obdelovati mnogo stvari, predvsem v primeru, ko gre za namensko razvito programsko opremo. Obstajajo različni algoritmi ali tehnike, ki so sposobne obdelave slik, pregledovanja črtnih kot, OCR, filtriranje podatkov, komunikacijski vmesniki z ostalimi napravami v sistemu ter možnost krmiljenja enostavni procesov avtomatizacije.

3. ZAKLJUČEK

V prispevku smo predstavili le delček uporabe strojnega vida v industriji. Poudariti želimo, da kvalitetne rešitve strojnega vida predstavljajo tesno sodelovanje med naročnikom in integratorjem. Naloga integratorja je, da sistem uspešno razvije, implementira, ter nudi končno podporo; naloga naročnika pa, da mu posreduje čim več informacij o izdelku ter procesu izdelave.

LITERATURA

1. www.datalogic.com/eng/products/industrial-automation/
Pametni senzorji
2. www.vision.si
Programska oprema, svetilke
3. www.ptgrey.com
Kamere
4. <https://www.embedded.com/print/4382211>
Osvetlitvene tehnike

NAREDI SAM: MULTISPEKTRALNE IN GLOBINSKE KAMERE

Jernej Kranjec

Laboratorij za sistemsko programsko opremo

E-pošta: jernej.kranjec@um.si

URL: <https://lspo.feri.um.si/>

POVZETEK: *V tem predavanju bo na praktičen način predstavljeno, kako v svoji »garaži« izdelati cenovno ugoden sistem iz kamere RGB in »domače« multispektralne kamere NIR (Near-Infrared), ki ga lahko uporabimo za sinhrono zajemanje scene iz dveh pogledov. Demonstrirano bo tudi, kako sistem poravnati s posebej načrtovanim ohišjem, natisnjenim s tridimenzionalnim tiskalnikom.*

1. UVOD

V svetu računalništva ne potrebujemo veliko od ideje do eksperimenta. Najpogosteje je naš vložek čas, katerega porabimo za izvedbo programske opreme. Stvari postanejo težje, ko realizacija naše ideje zahtevajo posebno strojno opremo. Na srečo lahko danes s potrošniškim materialom in ustreznim znanjem izdelamo dovolj dobre specializirane merilne inštrumente za testiranje naših idej. Tokrat bomo spoznali, kako prirediti navadno kamero za možnost snemanja še bližnje infra-rdečega spektra.

2. ZAKAJ MULTISPEKTRALNE SLIKE

Eden od primerov, kjer lahko uporabimo multispektralno kamero z zmožnostjo zajemanja bližnje infra-rdečega spektra, je v analizi rastlinskega življenja. Z računanjem razmerja med rdečo (ali zeleno) in infra-rdečo pasovno širino barvnega spektra lahko dokaj natančno določimo, ali imamo v objektivu zajet klorofil. Več kot to, če imamo dovolj dobro opremo, lahko tudi ločimo med vrstami rastlin in celo ugotovimo njihovo zdravje. Seveda nismo omejeni samo na detekcijo rastlin, z različnimi valovnimi dolžinami lahko razlikujemo tudi med naravnimi in umetnimi materiali, izboljšamo kakovost slike v oteženih okoljih ali celo zajamemo biomedicinske signale.

3. KAKO DO MULTISPEKTRALNIH SLIK

Osnovni element naše multispektralne kamere je senzorski element vsake digitalne kamere, saj je kot tak občutljiv na večji razpon valovnih dolžin svetlobe kot pa človeško oko. Medtem ko lahko z očesom zaznavamo svetlobo med 380 nm in 750 nm je tipičen senzor v digitalni kameri občutljiv na valovne dolžine med 300 nm in 900 nm. Neželene

barve so tako odstranjene iz slike s pomočjo fizičnih filtrov v objektivu ali pa pred samim senzorjem. Največkrat so v obliki stekla na katerega je nanosena plast za infra-rdečo svetlobo odbojnega materiala, katerega lahko brez večjih težav odstranimo.

Na ta način dobimo slike z mešanico rdeče in infra-rdeče barve v rdečih pikah, kar brez posebne obdelave daje pogled presvetljene slike. Da se izognemo temu oz. če želimo zajeti spektralno čistejšo sliko, lahko odstranjen filter zamenjamo z lastnim filtrom s specifično karakteristiko.

Zaradi načina izdelave senzorjev, katera vključuje nanosa Bayerjevega filtra na sam senzor za produkcijo barvnih slik, moramo za zajem več kot treh barvnih spektrov hkrati uporabiti več sočasnih kamer. Za usmerjenost kamer ter poravnavo slike poskrbimo že z namestitvijo kamer na posebna ohišja. Porabimo lahko že narejena ohišja za zajem stereo posnetkov s športnimi kamerami, ali pa uporabimo današnjo dostopnost 3D tiskalnikov in izdelamo svoje ohišje v primeru uporabe manj popularnih modelov kamer.

4. ZAKLJUČEK

Zagotovo se natančnost in kvaliteta tako izdelanih instrumentov ne more primerjati s profesionalnimi produkti, nam pa omogoča hiter in cenovno ugoden način preizkušanja idej.

ZAJEM SLIKOVNIH INFORMACIJ: OD ZASNOVE DO IZVEDBE

Jure Škrabar

Kolektor Group d.o.o.

E-pošta: jure.skrabar@kolektor.com

URL: <http://www.kolektorvision.com>

POVZETEK: *Zasnova optičnih merilnih sistemov je kompleksna naloga, ki zahteva poznavanje gradnikov s področja strojništva, elektrotehnike ter programske opreme. Za doseganje zahtevanih tehničnih lastnosti v celotnem življenjskem obdobju optičnega sistema/naprave je ta korak izredno pomemben. V sklopu delavnice so predstavljeni potrebni koraki za vzpostavitev optičnega merilnega sistema od zasnove, do končne izvedbe na praktičnem primeru.*

1. UVOD

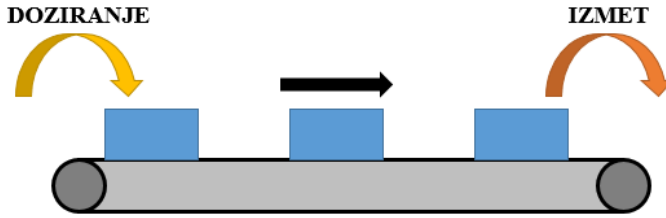
Samo razumevanje delovanja senzorjev za zajem slikovnih informacij je dobra podlaga za konstrukcijo optičnih merilnih sistemov in naprav. V obzir je potrebno vzeti ne samo izbiro komponent za vzpostavitev zahtevanega delovanja merilnega sistema temveč tudi variacijo vhodnega materiala oz. komponent, katere so predmet kontrole. Optični merilni sistemi so mehatronski produkt, sinergija delovanja in znanj elektrotehnike, strojništva, optike ter programske opreme.

2. KONTROLA IZDELKA – MERITEV IZVRTINE NA TEKOČEM TRAKU

2.1 Predstavitev problematike

Za lažjo predstavo cikla demonstracijske aplikacije si predstavljajmo aluminijast ulitek, z izvrtino na zgornji površini. Izdelek se premika po tekočem traku, pozicioniran v prečni smeri in naloga optičnega merilnega sistema je, da zazna nepravilno obdelano izvrtino.

Izdelek je na površini kontrole grobo obdelan, s površinskimi razami od obdelave s postopkom rezkanja, izvrtina pa je v vrhnjem delu posneta. Napaka se formira kot odstopanje premera izvrtine v rangi +/- 0,1 mm. Izdelek leži na tekočem traku z ravnino kontrole paralelno z naležno površino.

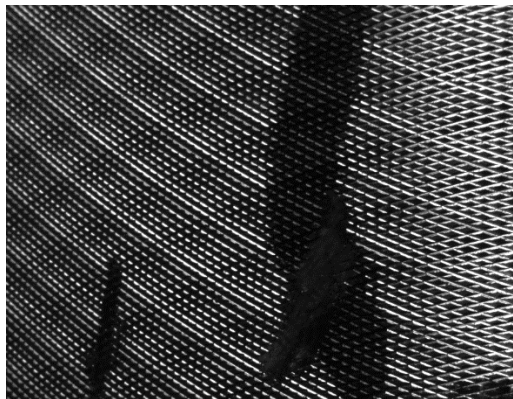


Slika 1: Manipulacija izdelka

2.2 Zasnova merilnega sistema

Z obzirom na predstavljeno problematiko si v sosledju sledijo koraki zasnove:

- Izbor metode za zajem slikovne informacije.
- Izbor osvetlitvenega principa ter eksperimentalna potrditev.
- Izbor potrebne elektronske periferije ter vmesnikov.
- Definicija metode za programsko analizo vhodnih slikovnih informacij.
- Definicija mehanskih vmesnikov.
- Definicija komunikacijskih vmesnikov.
- Končni popis izbrane opreme ter potrditev zasnove merilne naprave.
- Končna potrditev merilnega principa z MSA analizo.



Slika 2: Obdelana površina izdelka

Vsak izmed zapisanih korakov je obsežen, a z izbiro prave metode za zajem slikovne informacije se doseže v 90% primerov zadostno robustnost za stabilno delovanje merilne naprave v celotnem življenjskem obdobju.

3. ZAKLJUČEK

Tekom delavnice je predstavljena vsaka izmed naštetih točk, na praktičnih primerih prikazan optimalen izbor metode zajema slikovnih informacij ter optične opreme (2D, 3D), elektronskih vmesnikov ter periferije (proženje zajema slike, sledenje objektom po tekočem traku) ter nenazadnje končna validacija merilne naprave z MSA analizo. Le-ta nam potrdi pravilno izbiro celotne verige merilnega sistema, saj vsak izmed členov doprinese k končnemu rezultatu.

LITERATURA

1. A. Hornberg (2002), Handbook of Machine Vision, Wiley-VCH Verlag GmbH & Co. KGaA

PRIMER IZ PRAKSE: UMEŠČANJE SKENA KROŽNE IZVRTINE V ABSOLUTNI KOORDINATNI SISTEM

Damjan Vrenčur

Kolektor Orodjarna d.o.o.

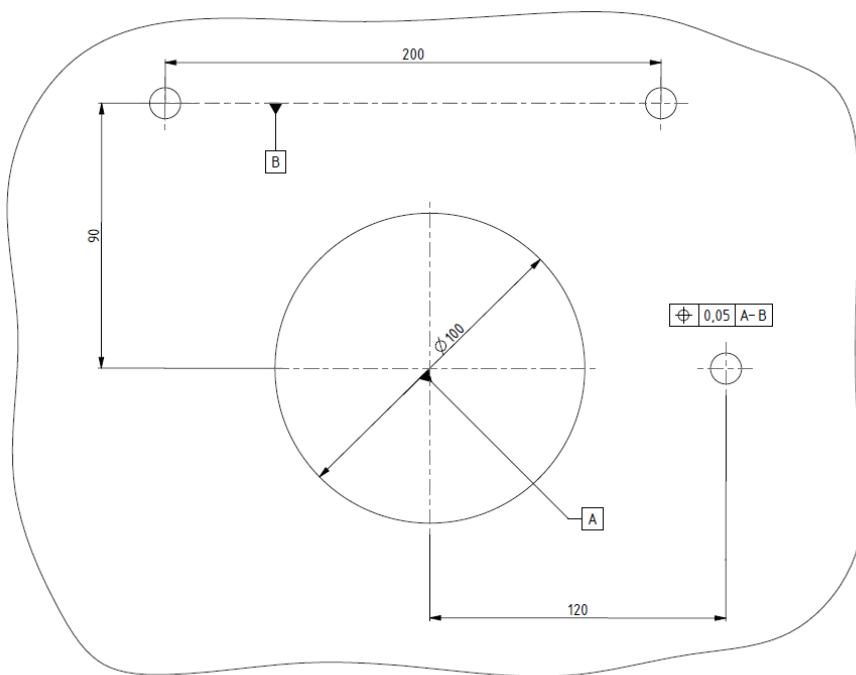
E-pošta: vision@kolektor.com

URL: <http://www.kolektorvision.com>

POVZETEK: *Pri merjenju dimenzij ulitkov, ki se uporabljajo v avtomobilski industriji, se ključna meritev pozicije pogosto veže na kakšno od krožno obdelanih izvrtin večjih dimenzij. Izkazuje se, da je rotacijski sken fizikalno ena bolj primernih metod za optično detekcijo lastnosti takšnih izvrtin. V predavanju si bomo ogledali izzive, ki jih takšen pristop prinaša ter eno od implementiranih rešitev, ki smo jo izdelali v Kolektor Vision v letu 2017.*

1. UVOD

Mnogo izdelkov livarske industrije za potrebe proizvodnje pogonskih delov avtomobilov (ohišij menjalnikov, delilnikov navora, motorja ipd.) je načrtovanih okoli centralne izvrtine – izvrtine, skozi katero je v končnem sestavu speljana uležajena pogonska gred. Centralna izvrtina je običajno opremljena tudi s tesnilom gredi, zato je njen premer pogosto velik v primerjavi z ostalimi merjenimi značilkami na ulitku, poleg tega pa so tolerance in ujemi lahko tudi do tri velikostne stopnje manjši od značilnih dimenzij izdelka. Narava uporabe izdelka običajno zahteva postavitev vsaj ene tolerančne baze v središče centralne izvrtine. Primer tehnične risbe takšnega izdelka kaže slika 1.

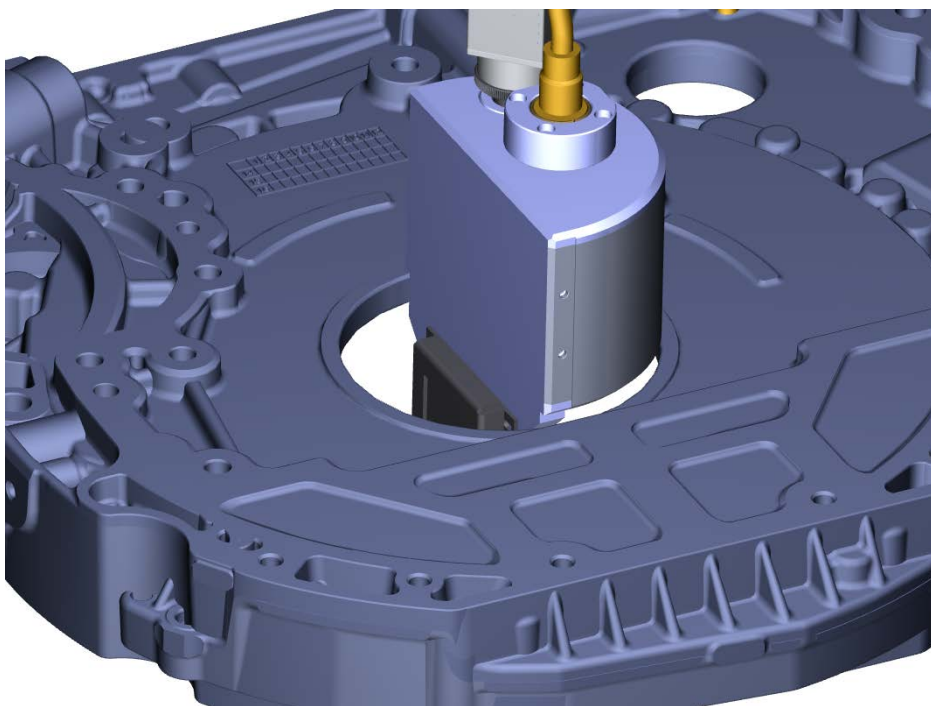


Slika 1: Idealiziran primer tehnične risbe, kjer je pozicija izvrtine na desni vezana na dve virtualni bazi: središče izvrtine z večjim premerom ter nosilko središč dveh manjših izvrtin zgoraj.

Ob načrtovanju merilne naprave za takšne tipe izdelkov se kot poseben problem zastavlja naloga uspešne realizacije baz. Najočitnejša rešitev – 3D skeniranje celotnega izdelka ter analiza tako pridobljenega oblaka točk – običajno ne pride v poštev zaradi časovnih omejitev takta proizvodne linije. Združevanje večih kamer v skupni koordinatni sistem omogoča hitrejšo kontrolo, a s seboj prinaša dodatne izzive. Prvi je potreba po izdelavi posebnih kalibrskih kosov, drugi pa dilema, kako dovolj natančno zaznati obliko ter pozicijo centralne izvrtine. V nadaljevanju se bomo osredotočili na problem centralne izvrtine.

2. CENTRALNA IZVRTINA – METODE OPTIČNE DETEKCIJE

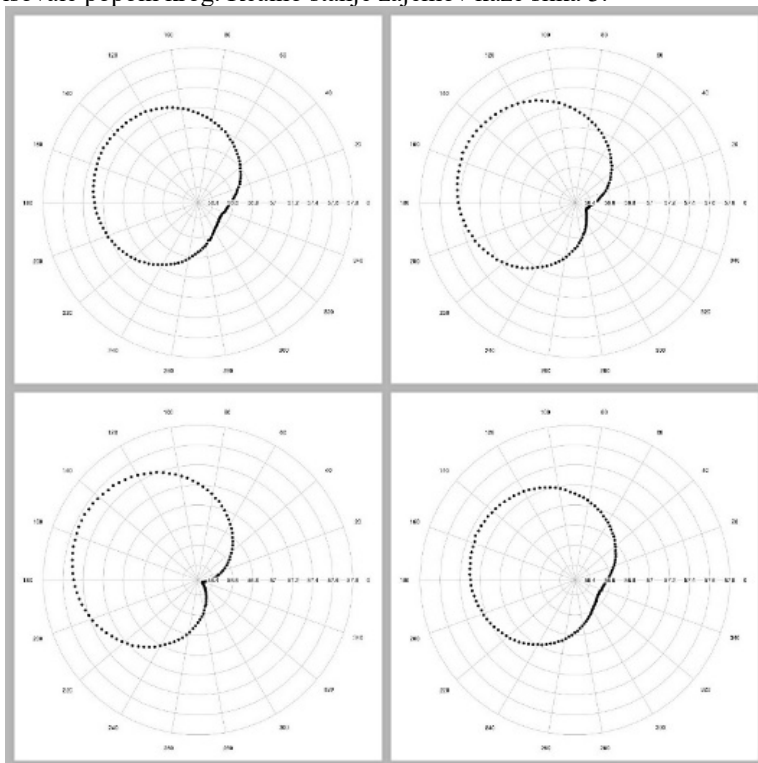
Zanesljivo ter natančno posneti obliko izvrtine z večjim premerom v tolerancah, ki so tri velikostne rede manjše od njenih značilnih dimenzij (recimo $100\text{ mm} \pm 0.1\text{ mm}$) je z enim posnetkom na meji optično mogočega. Mehanske omejitve natančnosti postavitve izdelka, neenakomerna obdelava kontroliranih površin, lom svetlobe na robovih, inherentne optične distorzije leče ter učinkovita ločljivost celotnega optičnega sistema predstavljajo ovire, ki jih je izredno zahtevno preseči. Zato smo se v Kolektor Vision odločili uporabiti drugačen pristop: razvili smo lastno lasersko triangulacijsko glavo, ki z rotacijskim gibom znotraj izvrtine posname njeno obliko (Slika 2). Takšen način detekcije omogoča mnogo večjo natančnost kot metoda enega posnetka, zahteva sicer več časa, a je za razliko od 3D skeniranja izdelka običajno še vedno v časovnih okvirih, ki so sprejemljivi za naročnika.



Slika 2: Triangulacijska glava v centralni izvrtini ulitka med zajemom podatkov.

3. CENTRALNA IZVRTINA – ANALIZA

V idealnem primeru, ko je krožni tek osi rotacijske skenirne glave ničeln, ko je optična pot laserja poravnana radialno ter, ko ni optičnih distorzij, pričakujemo, da bodo zajete točke opisovale popoln krog. Realno stanje zajemov kaže slika 3.

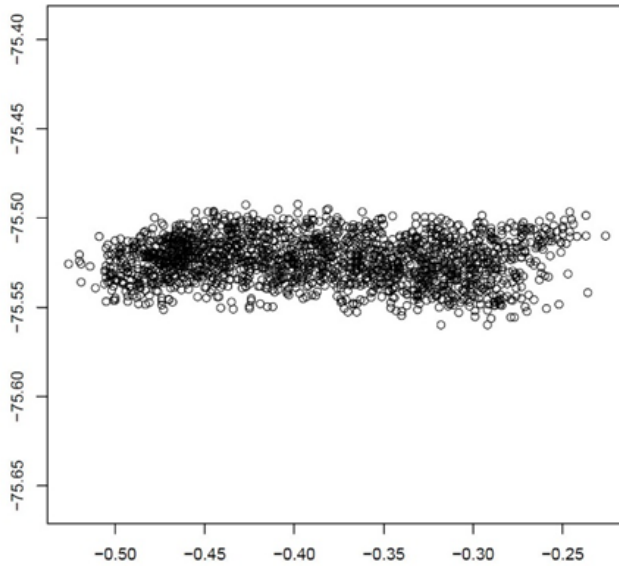


Slika 3: Prikaz podatkov iz štirih skenov centralne izvrtine istega kosa.

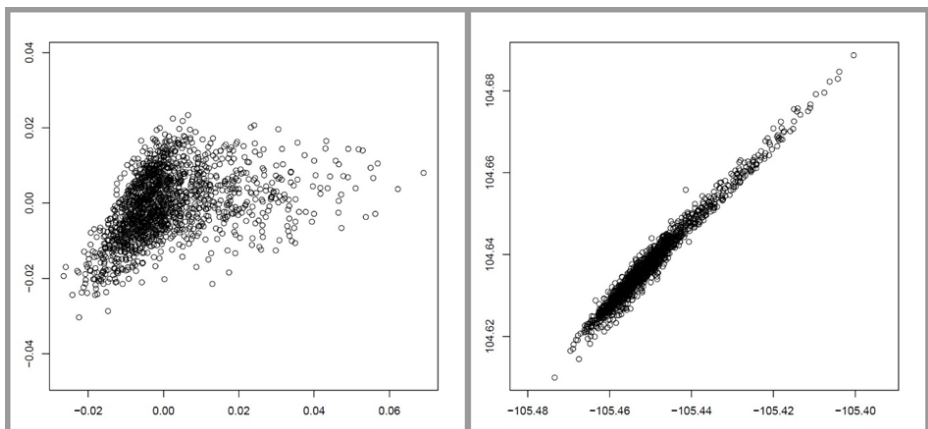
Postavlja se vprašanje, kako iz takšnih podatkov dovolj natančno določiti središče izvrtine, saj se le to uporablja kot baza za vse ostale (vezane) meritve.

Izkaže se, da omogočajo pogosto uporabljane metode, ki temeljijo na metodi najmanjših kvadratov oziroma na presekih eksaktnih rešitev različno filtriranih točk dovolj veliko natančnost ter ponovljivost za določanje premera izvrtine, vendar so se pri določanju položaja središča te iste metode izkazale za premalo natančne (slika 4).

Zato smo v sodelovanju s podjetjem Ektimo d.o.o. v Kolektor Vision razvili lastno metodo, prilagojeno posebnostim pri nas izdelane skenirne glave. Rezultate korelacije med teoretičnimi ter izmerjenimi položaji središča prikazujeta grafa na sliki 5.



Slika 4: Korelacija med x-koordinato središča centralne izvrtine, pridobljene s pomočjo ene od klasičnih metod prilagajanja kroga skeniranim podatkom ter x-koordinato, izračunano iz znanih značilnih mer kalibrskega vzorca. Napaka, ki jo naredimo na tak način, je velikostnega reda dveh desetink milimetra, kar je nesprejemljivo veliko.



Slika 5: Korelacija med izmerjenima in realnima koordinatama središča centralne izvrtine. Pričakovana maksimalna napaka v izmerjeni poziciji je ob uporabi lastne metode izračuna središča v rangu nekaj stotink milimetra.

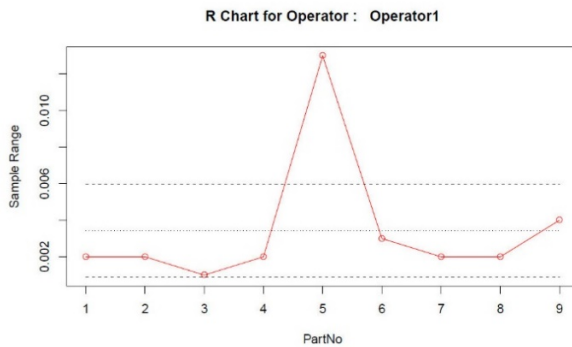
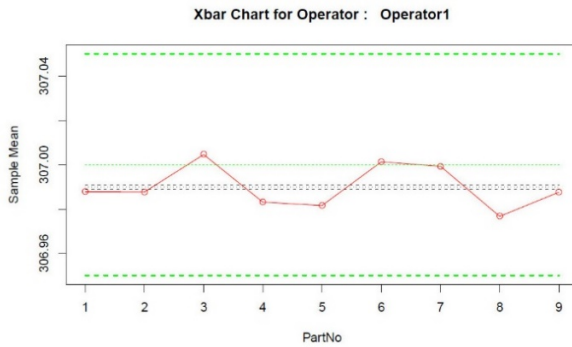
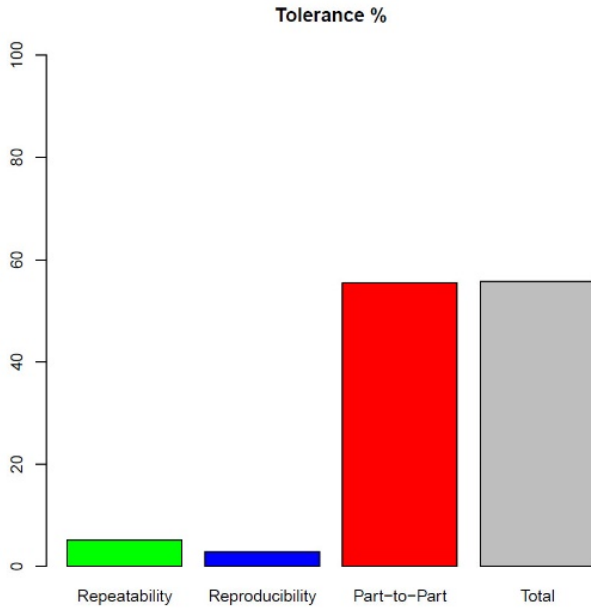
4. ZAKLJUČEK

Izkazalo se je, da lahko s tako umerjenim strojem merimo nekatere parametre ulitkov s tolerancami v stotinkah milimetra ter ob tem dosežemo visoko ponovljivost meritev, kot to kažejo slike 6, 7 in 8.

Gauge R&R

	stdDev	%SudyVar	%Tolerance
Repeatability	0.00085	9.12	5.09
Reproducibility	0.00047	5.08	2.84
Part-to-Part	0.00925	99.45	55.48
Total	0.00930	100.00	55.79

Number of distinct categories: 13.47



Slike 6, 7 in 8: Rezultati MSA analize na enem od merjenih parametrov na ulitkih.

DOPOLDANSKA SEKCIJA
Industrijske aplikacije
Medicinske in biomedicinske aplikacije
Drugo
Študentske aplikacije

AVTOMATIZIRANA VIZUALNA KONTROLA MAGNETOV

Blaž Pongrac, Danijel Šipoš, Dušan Gleich

**Laboratorij za obdelavo signalov in daljinska vodenja
Fakulteta za elektrotehniko, računalništvo in informatiko
E-pošta: blaz.pongrac1@um.si
URL: <http://feri.um.si>**

Luka Ažman

**Magneti Ljubljana, d.d.
E-pošta: luka.azman@magneti.si
URL: <http://www.magneti.si>**

POVZETEK: *Predlagana je bila rešitev avtomatizirane vizualne kontrole magnetov enega tipa. Kontrola se izvaja z merjenjem več parametrov: dolžine in premere magneta, eliptičnosti in okrušenosti magneta. Prve tri merimo z upragovanjem, okrušenost magneta pa s pomočjo Prewitt-ove metode za detekcijo robov.*

1. UVOD

Magneti Ljubljana d.d. se ukvarja s proizvodnjo magnetov in magnetnih sistemov. Proizvodnja magnetov je sestavljena iz več postopkov: izdelava magneta, brušenje magneta, kontrola, magnetizacija in pakiranje. V prvih dveh postopkih je magnet zaradi samih postopkov in manipulacije podvržen mehanskim poškodbam. Iz tega razloga je potrebno pred magnetizacijo preveriti dimenzijsko ustreznost magneta in morebitne poškodbe ter neustrezne magnetne izločiti iz nadaljnje obdelave.

Trenutni postopek kontrole opravlja človek. Preverjanje dimenzij magneta je kvantitativno dobro definirano. Za to se uporabljajo preverjena orodja (mikrometer), sam postopek pa ima relativno majhen merilni pogrešek. Vizualna kontrola zajema preverjanje eliptičnosti magneta, konusne oblike in okruškov magneta. Problem trenutne vizualne kontrole je zagotoviti kvantitativno merilo kvalitete, saj lahko več kontrolorjev različno oceni določen magnet (ocena je subjektivna). Z vpeljavo avtomatiziranega postopka kontrole s pomočjo strojnega vida lahko pohitrimo in pocenimo postopek vizualne kontrole, prav tako pa lahko v veliki meri odpravimo subjektivnost končne ocene.

Predlagana rešitev uporablja tehnologijo strojnega vida. Ta se vedno pogosteje uporablja za namene vizualne kontrole, ne glede na panogo ali tip proizvodnje. Omogoča nam kreiranje kvantitativnega merila kakovosti izdelka, ki temelji na vnaprej določeni vzorčni

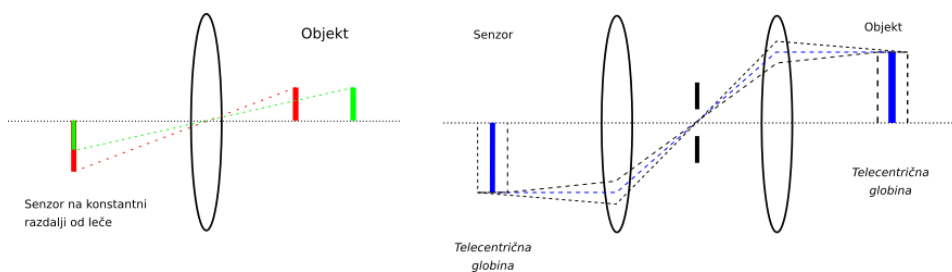
množici. Gre tudi za hitre sisteme, ki lahko več meritev opravijo v veliko krajšem času, kot lahko to storijo kontrolorji.

Sama rešitev je prilagojena točno določenemu tipu magneta, vendar je tehnologija lahko univerzalna oz. je prilagodljiva drugim tipom. Potrebno je kreirati novo vzorčno množico in preko nje nov kvantitativni kriterij kvalitete.

2. PROBLEM VIZUALNE KONTROLE MAGNETOV

Postopek avtomatizirane vizualne kontrole je prilagojen za valjčne magnete. Poteka v treh stopnjah oz. postajah. V prvi stopnji se preverja plašč magneta, v drugi in tretji pa obe čelni ploskvi. Preverjanje plašča magneta poteka s pomočjo vrtenja le-tega in zajemanja petih fotografij površine plašča. S tem je zagotovljeno zajemanje celotne površine plašča, saj je pri petih fotografijah zagotovljeno delno prekrivanje. Iz fotografij plašča se določijo dolžina magneta in okrušitve. Druga in tretja postaja preverita prednjo in zadnjo čelno ploskev. Določijo se premer magneta in eliptičnost, iz teh dveh pa posredno tudi konusna oblika, ter okrušitve.

Vsaka postaja ima po dve kameri, vsaka kamera zajame sivinsko fotografijo z dvema magnetoma (avtomat ima štiri linije za kontrolo magnetov). Uporabljena je čelna osvetlitev s pomočjo LED, ki delujejo v režimu bliskavice. Pri zajemu fotografij se pojavi več različnih pogreškov: pogrešek zaradi gretja kamere, pogrešek leče in odstopanje globinske ostrine. Vsi ti pa negativno vplivajo na merjenje dimenzij magneta. Te probleme rešujejo telecentrične leče.

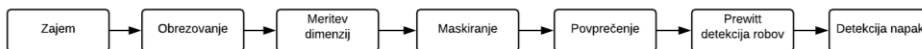


Slika 1: Primerjava navadne in telecentrične leče

Slika 1 prikazuje primerjavo med delovanjem navadne leče in uporabljene bilateralne telecentrične leče, kot je to predstavljeno v [1]. Navadna leča preslika objekt na senzor z različno globinsko ostrino – če je objekt bližje leči, je tudi večji. Pri bilateralni telecentrični leči pa sta uporabljeni dve leči in zaslonka, s pomočjo katerih se ustvari telecentrični globini na strani objekta in senzorju. Iz slike 1 vidimo, da je v območju telecentrične globine globinska ostrina enaka, kar pa odpravi večino težav pri merjenju dimenzij na 2D fotografiji. Zagotoviti je potrebno, da se magnet vedno nahaja znotraj območja telecentrične globine.

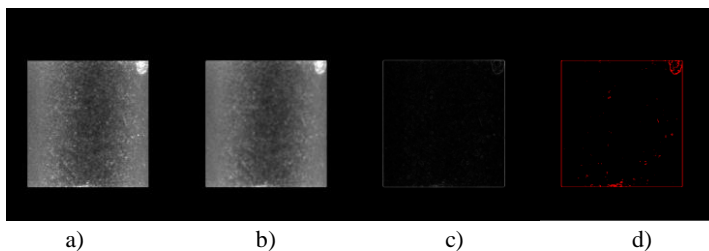
2.1 Predlagana rešitev

Izdelana sta bila dva ločena algoritma, ki se pa po svoji zasnovi ne razlikujeta. Oba temeljita na detekciji robov s pomočjo Prewitt-ovega operatorja. Slika 2 prikazuje splošno blokovno shemo.

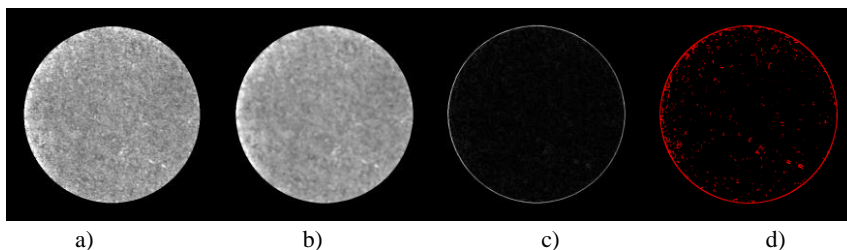


Slika 2: Blokovna shema algoritma

Po zajemu fotografije je potrebno le to obrezati, saj se z eno kamero zajema fotografija dveh magnetov. Na tako pripravljene fotografiji se izvede meritev dimenzij, na podlagi teh pa s pomočjo maskiranja izločimo ozadje iz nadaljnje obdelave. Preden določimo robove na magnetu, fotografijo le tega povprečimo. Prewitt-ov gradientni operator, ki predstavlja magnitudo gradienta, je določljiv s konvolucijskima maskama, predstavljenima v [2], kateri pa ne vnašata poudarke na slikovne točke v okolici središča maske. Zato je ta algoritem detekcije robov način, kako se izognemo računanju gradienta v okolici slikovne točke. Odzive detekcije robov s Prewitt-ovim operatorjem prikazujeta sliki 3 in 4. Detektirani robovi se nato upragujejo, s pomočjo merjenja površine napake pa se le-ta določi. Mejna vrednost napake je določena s pomočjo vzorčne množice. Vzorčno množico so pripravili izkušeni kontrolorji magnetov.

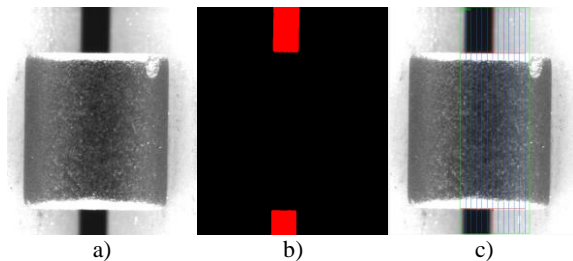


Slika 3: Detekcija napak na plašču: a) obrezana slika, b) povprečenje, c) Prewitt detekcija robov, d) upragovanje najbolj očitnih robov



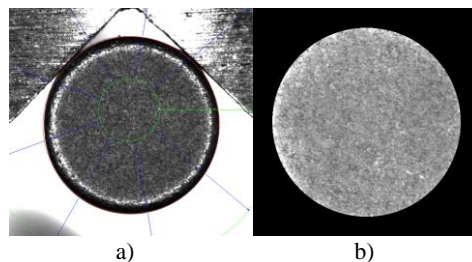
Slika 4: Detekcija napak na čelni plošči: a) obrezana slika, b) povprečenje, c) Prewitt detekcija robov, d) upragovanje najbolj očitnih robov

Algoritma se razlikujeta le v detekciji položaja magneta in v določevanju dimenzij. Za določitev položaja magneta pri preverjanju plašča se uporablja upragovanje. S tem se določi središče mase zajetih temnih področij, ki se nahajajo nad in pod magnetom ter so posledica razmika med valjčki, ki vrtijo magnete, ter zgornji in spodnji rob magneta. S temi dimenzijami se določuje področje izreza. Posamezne stopnje algoritma prikazuje slika 5. Dolžino magneta določujemo s pomočjo upragovanja. Po večih vertikalnih premicah algoritem določuje prestop praga in smer prestopa praga vrednosti slikovne točke. Na ta način algoritem določi dve premici, ki pomenita zgornji in spodnji rob magneta, in oddaljenost med njima. S pomočjo kalibracije pa se nato izračuna dolžina, odstopanje od želene dolžine, iz slednjega pa se določi napaka v dolžini.



Slika 5: Določevanje izreza in dimenzij na plašču: a) obrezana slika, b) iskanje središča magneta z iskanjem središča mase reže, c) merjenje dolžine

Položaj magneta pri preverjanju čelne plošče se določi z detekcijo kroga na fotografiji. S tem se določi center kroga, premer kroga in odklon premera kroga. Podobno kot merjenje dolžine, tudi detekcija robov kroga poteka tako, da merimo vrednosti slikovnih točk in po prestopu praga v zeleni smeri določimo oddaljenost roba od središča območja iskanja. Po določitvi pragovnih točk na robu kroga, se nato izračunata center kroga in premer. Oba se uporabita za izrez kroga in nadaljnjo procesiranje. Določitev izreza kroga je prikazana na sliki 6.

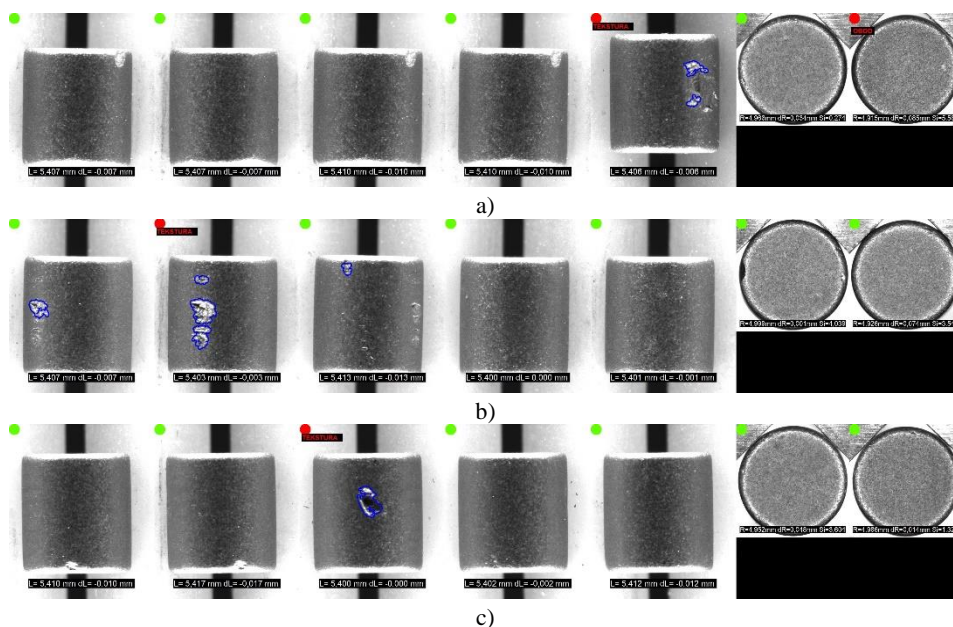


Slika 6: Detekcija kroga: a) upragovanje roba po daljicah b) izrez

Konusna oblika magneta se detektira posredno s pomočjo prevelikega odstopanja premera, ne glede na čelno ploskev. Eliptičnost pa se določi s pomočjo najmanjšega kvadratičnega pogreška oz. odklona premera. Detektiran in v fotografijo vrisan krog ima

idealno obliko. Algoritem detekcije kroga po vrisu le-tega preveri odstopanja idealne oblike od dejanskih točk, nato pa izračuna minimalni kvadratični pogrešek. Ker je maksimalni odklon premera magneta poznan, je mejna vrednost izmerjenega odklona določljiva.

Na sliki 7 je prikazano nekaj primerov rezultatov. Na končni sliko je izrisanih vseh pet fotografij plašča in obe čelni plosči. Na vsaki fotografiji pa je izpisana dimenzija in njeno odstopanje od želene vrednosti.



Slika 7: Primeri rezultatov

3. ZAKLJUČEK

Načrtovano je, da bi predlagana rešitev avtomatizirane vizualne kontrole magnetov pohitrila in pocenila postopek kontrole, hkrati pa bi dvignila nivo kvalitete končnih izdelkov. Predlagan algoritem je načrtovan tako, da je univerzalen in robusten, čeprav je v samem postopku več spremenljivk (vsak magnet ima lahko različen položaj v trenutnem času, npr. pri vrtenju magnetov na prvi postaji pride do premikov magneta – položaj magneta moramo detektirati). Po prvih testih se je izkazalo, da lahko v samem algoritmu pride do nekaterih težav. Največ težav se lahko pojavi z osvetlitvijo in umazanijo, ki jo sami magneti vnašajo v sistem (zapacajo se lahko vrtljivi valčki, leče, itd.). Kljub temu se je algoritem izkazal za robustnega na testnem vzorcu.

LITERATURA

1. B. Pan, L. Yu, D. Wu (2013), High-Accuracy 2D Digital Image Correlation Measurements with Bilateral Telecentric Lenses: Error Analysis and Experimental Verification, *Experimental Mechanics* (2013) 53, Society for Experimental Mechanics, str. 1719-1733
2. G.T. Shrivakshan, C. Chandrasekar, A Comparison of various Edge Detection Techniques used in Image Processing, *IJCSI International Journal of Computer Science Issues*, Vol. 9, Issue 5 (2012), *IJCSI International Journal of Computer Science Issues*, str. 269-276

TOWARDS IMPROVING FACE DEIDENTIFICATION PIPELINE

Tadej Ciglarič, Žiga Emeršič, Peter Peer, Blaž Meden

Faculty of Computer and Information Science, University of Ljubljana

tc2922@student.uni-lj.si,
{ziga.emersic, peter.peer, blaz.meden}@fri.uni-lj.si

ABSTRACT: *Face deidentification is an important part of privacy and security domains. Deidentification methods that rely on image blurring, pixelization or black-boxes were replaced in recent years with approaches based on formal anonymity models that provide privacy guaranties and at the same time aim at retaining certain characteristics of the data. However, current state-of-the-art pipeline we developed in earlier work, still suffers from sometimes erroneous face detection and color discrepancies in the replaced faces. We made improvements to an existing deidentifications pipeline by replacing face detector and implementing color correction on replaced faces. We also changed blending of replacement face to the original image in order to reduce amount of visual artifacts. Resulting pipeline misses less faces and creates more natural looking face replacements.*

1. INTRODUCTION

Nowadays large mount of videos are recorded every day. Sharing videos databases may infringe personal rights - videos can contain personal information such as faces. In such cases personal information needs to be deidentified, while still retaining certain characteristics, e.g. person's pose and gender. This creates the need for automatic deidentification. To tackle this problem we improve on the recently introduced deidentification pipeline [1], namely with improvements to detecting faces and making face replacement artifacts less noticeable.

1.1 Deidentification Pipeline

The recently introduced pipeline [1] consists of multiple image processing modules implemented in Python and Keras deep learning framework [2]. The overview of the pipeline improvements is illustrated in Figure 1, together with added improvements presented in this paper.

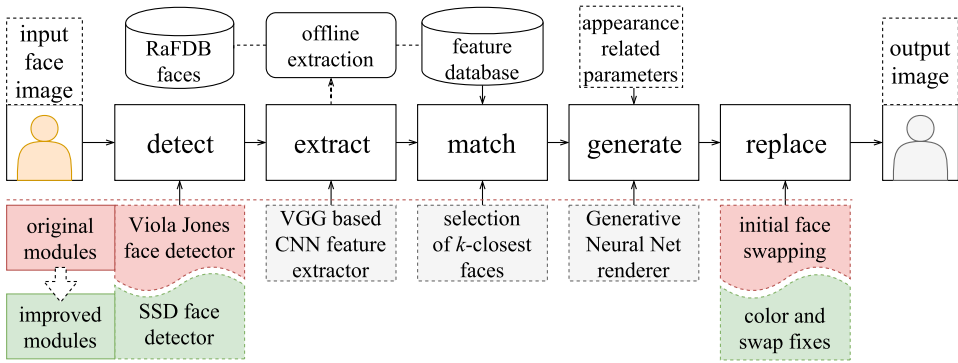


Figure 1: Overview of the deidentification pipeline from [1] with added module improvements. Viola-Jones face detector was replaced with CNN-based Single Shot Detector (SSD) trained for face detection. Face swapping module was updated with more advanced masking system with skin color corrections.

Pipeline begins with a face detection algorithm. The original pipeline uses Viola-Jones [3] face detector. Detected faces are then compared to database of predefined identities. Features, extracted by VGG [4] neural network are used for comparison. In the matching phase, k -closest identities are chosen to be combined to create a replacement face. This way the replaced face will be roughly similar to original one, but not enough to allow reliable identity recognition.

Data about chosen identities is fed into the generative neural network which generates replacement face, combining features of the chosen identities. The generating network has additional, appearance based inputs, which could control expression or any other visual aspects of the generated face to match the original one, thus preserving non-identity related attributes and enabling data utility in the deidentification process.

Finally original face must be replaced with generated one. On both original and generated face landmarks are detected. Detected points are used to estimate homography between original and generated face using RANSAC algorithm. Generated face is then warped according to estimated transformation. Gaussian mask is used to blend central part of the generated face to the original face – this is intended to suppress the sharp edges during face merging. The mask is originally combined with HSV threshold based skin segmentation in order to merge only facial regions, excluding hair and background.

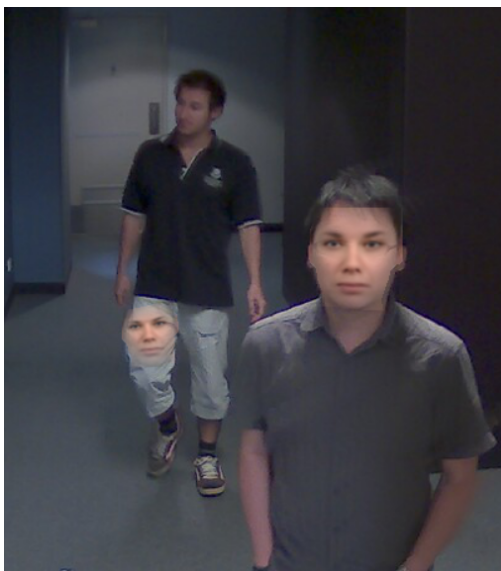


Figure 2: Failures of Viola-Jones face detector.



Figure 3: Visible face replacement artifacts.

2. PROBLEMS AND IMPROVEMENTS

Original pipeline implementation suffered from two major issues: face mis-detections and suboptimal face replacement. We addressed both of these problems and in this section we provide descriptions of the solutions to these problems and visual illustrations of some examples.

2.1 Improving Face Detector

Viola-Jones face detector, used in the original pipeline, is fast but suffer from mis-detections, when small or non-frontal faces are present in video sequences. In Figure 2 an example of the right person being correctly detected and deidentified, while face of the left person is not being detected. In this particular frame there is also a false detection in the knee area of the left person.

In order to avoid this kind of mis-detections we replaced the Viola-Jones-based face detector with a CNN-based approach. We selected Single Shot multibox Detector (SSD) proposed by Liu et al. [5]. The idea behind SSD is that it produces a fixed-size collection of bounding boxes and scores for the presence of object class instances in those boxes. This is followed by a non-maximum suppression step to produce the final detections. Detector uses VGG architecture [6] as a base model. Dense layers at the end of the original VGG are replaced with multi-scale convolutional layers, which predict the offsets to default boxes of different scales and aspect ratios and

their associated confidences.

We used MobileNet implementation from GitHub [7] which was adapted and trained for the task of face detection. It offers improved robustness compared to Viola-Jones. However faces are not always as well localized as Viola-Jones detections.

2.2 Improving Face Replacement Module

Next we have improved face replacement module. Using color-based skin segmentation and Gaussian kernel often creates more or less visible artifacts. Also there are visible sharp edges around replaced face. An extreme case can be seen in Figure 3.

Instead we try to estimate face area more precisely. We calculate convex hull of detected face landmarks on original face. Points of convex hull are used to render a polygon to use as a mask. It is than slightly blurred to prevent creation of sharp edges between original image and replaced face. Figure 4 shows the original replacement procedure and improved technique and visualises final deidentification result for both cases.

Generated faces often differ in color from originals. This is especially evident for different skin colors or in case of poor lighting conditions. Generated faces always have white skin and seem to be in bright environment. Examples are shown in Figure 5.

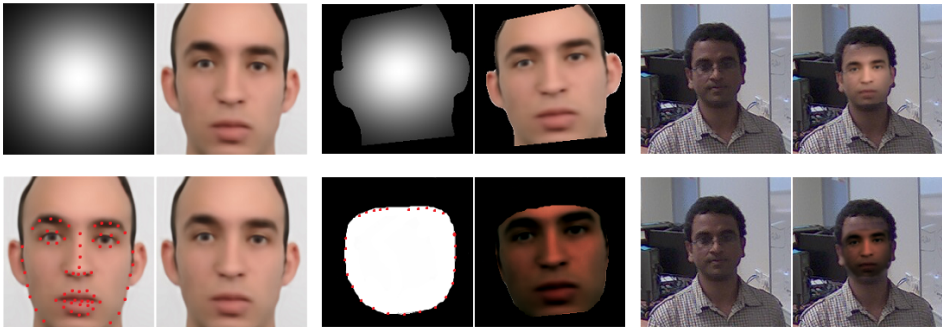


Figure 4: Top row shows the original pipeline with hard masks for face replacement [1]. The bottom line shows the improved replacement, based on the detected face landmarks.

3. QUALITATIVE RESULTS

Evaluation of existing and modified pipeline was done on Chokepoint dataset [8]. It contains images from surveillance-like scenarios. Cameras are placed at "chokepoint" locations, where many people pass by.

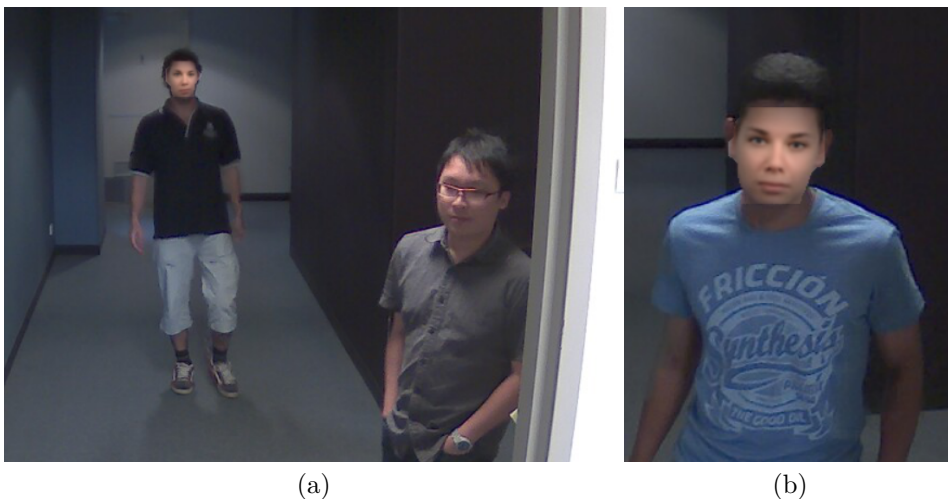


Figure 5: Two examples of color artifacts, present in the original implementation. (a) Generated face is not adapted to dark environment. (b) Deviating skin color effect of replaced face.

Replacing the Viola-Jones with SSD results in less missed face detections and therefore more complete deidentification, however the face localization is not that accurate than in Viola-Jones, which can result in face landmark detection failures in the final step of the pipeline. Modified face replacement module creates more natural looking faces as shown in Figure 6 (a) and (b).

SSD face detector we used does not seem to localize faces as well as Viola-Jones. Often this causes landmarking to fail. In such case the replacement face is not aligned with the head. Two examples of this are shown in Figure 6 (c) and (d). Nevertheless, the overall results are significantly better compared to using Viola-Jones.

4. CONCLUSION

We made two significant improvements to an existing deidentification pipeline. Significantly less faces are missed and replacements look more natural. Landmarking failures could be addressed by picking another face detector or landmarking model. Another possibility would be to use detections from SSD to retrain the landmarker. Future work will also include hair segmentation and forehead deidentification.



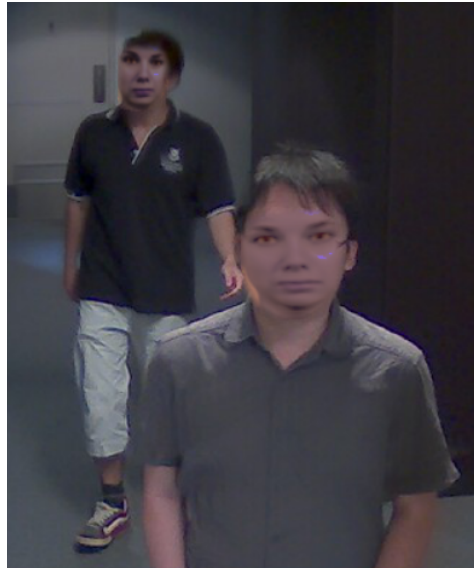
(a)



(a)



(c)



(d)

Figure 6: Four deidentified examples using improved pipeline. In top row (a) and (b) show a good and the most common examples of the deidentification. In bottom two failures are shown. (c) An example of misaligned face. (d) An example of one well deidentified face and one bad alignment in the back.

REFERENCES

- [1] B. Meden, R. C. Malli, S. Fabijan, H. K. Ekenel, V. Štruc, and P. Peer, “Face deidentification with generative deep neural networks,” *IET Signal Processing*, vol. 11, no. 9, pp. 1046–1054, 2017.
- [2] F. Chollet *et al.*, “Keras,” <https://github.com/fchollet/keras>, 2015.
- [3] P. Viola and M. J. Jones, “Robust real-time face detection,” *International journal of computer vision*, vol. 57, no. 2, pp. 137–154, 2004.
- [4] O. M. Parkhi, A. Vedaldi, A. Zisserman *et al.*, “Deep face recognition.” in *BMVC*, vol. 1, no. 3, 2015, p. 6.
- [5] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. E. Reed, C. Fu, and A. C. Berg, “SSD: single shot multibox detector,” *CoRR*, vol. abs/1512.02325, 2015. [Online]. Available: <http://arxiv.org/abs/1512.02325>
- [6] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *CoRR*, vol. abs/1409.1556, 2014.
- [7] “A mobilenet ssd face detector,” <https://github.com/yeephycho/tensorflow-face-detection>.
- [8] Y. Wong, S. Chen, S. Mau, C. Sanderson, and B. C. Lovell, “Patch-based probabilistic image quality assessment for face selection and improved video-based face recognition,” in *IEEE Biometrics Workshop, Computer Vision and Pattern Recognition (CVPR) Workshops*. IEEE, June 2011, pp. 81–88.

SISTEM ZA TESTIRANJE UPORABNIŠKEGA VMESNIKA VGRADNEGA RAČUNALNIŠKEGA SISTEMA

Marko Vačovnik¹, Janez Perš²

¹Gorenje d. d., Razvoj elektronike

E-pošta: marko.vacovnik@gorenje.com

²Laboratorij za strojno inteligenco

Fakulteta za elektrotehniko, Univerza v Ljubljani

POVZETEK: Članek opisuje avtomatski sistem za testiranje uporabniškega vmesnika. Pri tem se osredotočamo na preverjanje vmesnikov vgradnih računalniških sistemov. Glavni cilj aplikacije je preverjanje pravilnosti prikaza receptov in nastavitvev za avtomatsko pečenje na vgradni pečici. Za preverjanje pravilne umestitve sličic jedi je uporabljen primerjalnik deskriptorjev na osnovi histograma orientiranih gradientov, za preverjanje izpisanega besedila pa odprtokodno orodje za optično razpoznavanje znakov – Tesseract-ocr.

1. UVOD

Uporabniški vmesniki so že dalj časa eno najbolj vročih področij razvoja elektronskih naprav, saj za končnega uporabnika predstavljajo edini način interakcije z napravo. Uporabniki si pri uporabi želijo barvite in pestre interakcije, zato lahko kompleksnost uporabniškega vmesnika hitro naraste. Kompleksnejša izvedba botruje večji verjetnosti pojavljanja napak, zaradi česar nastane večja potreba po testiranju. To pa pomeni vse višjo proizvodno ceno izdelka.

Testiranje uporabniških vmesnikov je dolgotrajen postopek, ki ga je potrebno v polnosti izvesti ob vsaki spremembi programske ali strojne opreme. Testiranje samo predstavlja kar 50-60% celotne cene razvoja programske opreme, zato je smiselno čim večji del testov avtomatizirati.

1.1 Avtomatsko testiranje uporabniških vmesnikov

Področje bi lahko razdelili na dva pola. Pri tem so predstavniki prvega pola invazivni, predstavniki drugega pa neinvazivni pristopi. Razlika med njima je to, da prvi neposredno posegajo v programsko kodo uporabniškega vmesnika, drugi pa ne. Iz tega je jasno, da morajo invazivni teči znotraj platforme, ki teste v ozadju izvaja, medtem ko lahko neinvazivni tečejo tudi na drugi napravi v kolikor je ta povezana s testirano napravo.

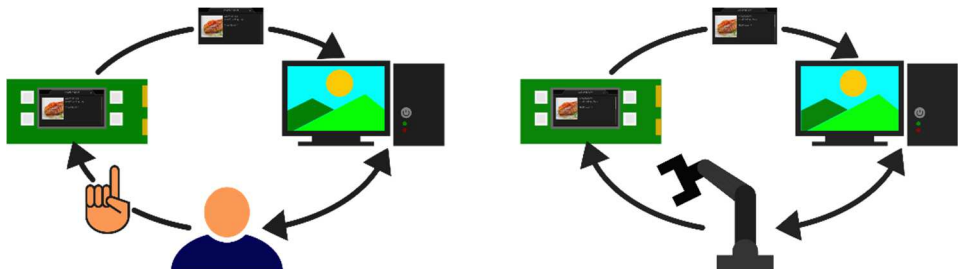
Neinvazivne pristope lahko razdelimo še na pristope za preverjanje s pomočjo skript in postopke posnemi-ponovi (ang. *Record-playback*). Pri prvih strokovnjak, ki ustvarja testne scenarije napiše program, ki komunicira direktno z uporabniškim vmesnikom, pri drugih pa poteka v prvi fazi snemanje uporabnikovih interakcij z vmesnikom. Celotna seja interakcije se posname in se lahko kasneje ponovi avtomatsko.

Naš cilj je v začetku bil avtomatizacija testnih postopkov za testiranje uporabniškega vmesnika pečice, ki so najbolj zamudni in periodični. Preverjanje receptov, prikazanih na zaslonu uporabniškega vmesnika pečice je zelo dolgotrajen postopek, saj uporabniški vmesnik vsebuje na šestih različnih modelih skupno preko 300 receptov jedi v tridesetih različnih jezikih. Pri tem je potrebno pozorno prebrati celotno besedilo celotnega recepta – zank za znakom. Poleg receptov je časovno potratno tudi preverjanje pravilnosti prikazanega predvidenega časa pečenja na zaslonu z nastavitvami za pečenje jedi, za katere so na voljo recepti. Pri tem je potrebno spreminjati količino ter zeleno intenziteto zapečenosti jedi, s čimer se spreminja tudi predviden čas pečenja.

Ker je elektronika pečice »svet zase«, pridejo v poštev samo algoritmi za avtomatsko testiranje, ki temeljijo na neinvazivnih pristopih. Interakcijo uporabnika s pečico je s pomočjo testne naprave – v našem primeru PC – nemogoče zaznati, zato odpadejo tudi metode *posnemi in ponovi*. Za naš primer je praktično izvedljiva samo metoda za avtomatsko preverjanje z uporabo skript, saj je scenarij periodičen in enostavno obvladljiv.

1.2 Pregled arhitekture uporabljenega sistema

Naš sistem neposredno iz elektronskega vezja pridobi sliko, prikazano na zaslonu pečice, ki jo ustrezno obdela in preveri, če je prikaz skladen z dokumentacijo. Za predobdelavo in segmentacijo slike je uporabljena odprtokodna knjižnica OpenCV [1], pri čemer je celoten sistem spisan v obliki skript v programskem jeziku Python. S pomočjo optičnega razpoznavalnika znakov preverjamo pravilnost prikazanega besedila, s pomočjo histograma orientiranih gradientov (HOG) pa pravilnost prikazanih sličic. Arhitektura našega sistema je nekako še najbolj podobna arhitekturi, ki jo opisujejo avtorji v delu [2].



Slika 1: Orodje za testiranje uporabniškega vmesnika samo z uporabo računalniškega vida (levo), in popolnoma avtomatizirano testiranje brez potrebe po operaterjevi interakciji (desno).

Na sliki 1 je prikazana shema sistema, kjer je viden potek interakcij med napravami in operaterjem. Ker smo se v tej fazi v glavnem posvečali avtomatizaciji s pomočjo računalniškega vida, sistem še ni v celoti avtomatiziran. Eden od ciljev za prihodnost je v model vključiti še kartezični robotski manipulator, ki bi upravljal z napravo namesto operaterja.

2. ZAJEM IN SEGMENTACIJA SLIKE

2.1 Zajem slike

Kot omenjeno, sliko pridobimo neposredno iz elektronskega vezja pečice. Osebni računalnik, na katerem teče testna aplikacija je s testirano napravo povezan preko protokola RS232, pri čemer je hitrost nastavljena na 2 Mbit/s, brez paritetnega bita in z enim stop bitom. Zajemamo sliko iz zaslona, velikega 480×272 , ki je zapisana v formatu RGB565. Iz tega sledi, da je minimalen čas prenosa slike daljši od 1.3 s.

Ko sliko pridobimo, na testni napravi pretvorimo njen format v RGB888, ker je ta format bolj prikladen za analizo slike.

2.2 Segmentacija slike

Če želimo, da naš sistem za avtomatsko testiranje doseže dober rezultat, je potrebno pridobljeno sliko kvalitetno segmentirati. Segmentacijo izvršimo v petih korakih.

Za namen hitrejše in enostavnejše segmentacije sliko najprej pretvorimo v sivinsko. Sivinsko sliko tvorimo kot kombinacijo vseh treh kanalov RGB slike. Na dobljeni sivinski sliki nato s pomočjo Laplaceovega filtra in upravljanja določimo robove. Za detekcijo robov smo testirali tudi Cannyjev detektor in Sobelov operator, a je pri teh prišlo do zamikanja robnih točk, zato slednjih pristopov v orodju nismo uporabili.

S pozitivnim pragom upravljenjo sliko, ki predstavlja robove s pozitivnim gradientom nato obdelamo tako, da jo pošljemo čez filter, ki iz nje odstrani majhna področja, ki so bila posledica šuma ozadja. To lahko storimo na dva načina. Prvi omogoča odstranjevanje premajhnih področij na podlagi minimalne površine, drugi pa na podlagi minimalne višine in širine. V obeh primerih filter za delovanje potrebuje označena sklenjena področja, ki predstavljajo robove. S pomočjo algoritma za označevanje povezanih komponent [3] poiščemo 8-povezljiva področja, kar nam omogoča lažje in hitrejše filtriranje. Nato na filtrirani sliki preostala področja združimo s pomočjo morfoloških operacij in uokvirjanja. Najprej preostala povezana področja uokvirimo in na sliki izvedemo dilacijo. Nato spet izvedemo uokvirjanje in na koncu še erozijo. Pri tem je zelo pomembna velikost jedra, saj ne želimo med seboj povezati področij, znotraj katerih se nahajajo različne grafike.

Zadnji korak segmentacije je še na enak način poiskati področja, ki so temnejša od ozadja. Na enak način, kot je opisano zgoraj opravimo obdelavo slike, le da pri upravljanju slike

robov uporabimo negativ filtrirane sivinske slike z Laplaceovim filtrom. Nato sliki s pozitivnimi in negativnimi robovi priležemo eno na drugo in od področij, ki se prekrivajo vedno izberemo manjše oz. objeto področje.

Vmesni rezultati posameznih korakov so prikazani na sliki 2. Prva vrstica slik prikazuje prve korake segmentacije zaslona z nastavitvami pečenja, druga pa končne na zaslonu z receptom. Na zadnji sliki je z modro barvo označeno področje na zaslonu, ki je temnejše od ozadja.



Slika 2: Vmesni rezultati posameznih korakov segmentacije. Začetni koraki na sliki prikazujejo segmentacijo zaslona z nastavitvami pečenja (prva vrstica), končni pa segmentacijo zaslona z receptom (druga vrstica).

3. METODE

Ustrezno segmentirano sliko zaslona za tem ustrezno analiziramo. Najprej klasificiramo regije znotraj posameznih področij, nato pa jih ustrezno z vsebino še nadalje segmentiramo ali pa ustrezno preverimo pravilnost njihove vsebine.

3.1 Klasifikacija področij

Klasifikacijo področij se izvaja na podlagi vektorjev značilk, ki predstavljajo posamezno področje. Osnova za gradnjo vektorjev značilk so momenti kotur znotraj področij. Računamo 24 momentov [4] za vsako konturo, nato pa dobljene vektorje povprečimo. Povprečnega predstavnika področja nato še logaritmiramo in tako dobljeni vektor odvajamo. Za še boljšo ločljivost med področji momentom dodamo še nekaj značilk na osnovi analize barv znotraj področij ter dimenzij in položaja področja znotraj slike zaslona.

Področja na podlagi izpeljanih vektorjev značilk delimo v 7 razredov: ikone, slike, drsnike, razmejevalnike, glave/noge, področja z besedilom in področja, ki jih pri pregledovanju ignoriramo.

3.2 Šivanje slik

Večina receptov je predolga, da bi jih bilo mogoče prikazati na enem zaslonu, zato je celoten recept razširjen na več zaslonov, uporabnik pa si lahko celoten recept ogleda s pomočjo drsnika. Drsnik je tako tudi pogoj za zagon algoritmov za šivanje slik v našem sistemu. Algoritem od operaterja zahteva, da mu »pokaže« celotno sliko recepta. To naredi tako, da obdela izrezek zaslona z drsnikom in na podlagi tega določi, kateri del celotnega recepta je že videl. Ko ima shranjene sličice celotnega recepta, jih glede na njim lastno pozicijo drsnika uredi po vrsti od zgornje proti spodnji. Nato se za posamezni sosednji sliki zaslonov izvede iskanje najboljšega prekrivanja s pomočjo RMS razlike področja prekrivanja. Ko so določena optimalna prekrivanja za celotno množico posnetih sličic, se sličice na ta mesta prilepijo ena na drugo.

3.3 Razpoznavanje besedila

Za razpoznavanje besedila na zaslonu pečice je uporabljeno odprtokodno orodje za optično razpoznavanje besedila Tesseract-ocr [5]. Z njim preverimo vsa področja, ki jih je klasifikator razpoznal kot področja z besedilom.

Optični razpoznavnik deluje tako, da najprej sliko po svoje predobdela, na njej poišče vrstice z besedilom ter k njim prileže kubični polinom. Za tem izvede detekcijo fiksne širine znakov in če ta ni detektirana, začne z iskanjem besed pri besedilih s proporcionalno širino znakov. To izvede tako, da najprej s pomočjo točk kandidatk razdeli povezane znake, za tem pa izvede asociiranje zlomljenih znakov na podlagi izpeljanih značilnk in klasifikacije s pomočjo vpogledne tabele (ang. *look-up table*). Na koncu izvede še lingvistično analizo za besede, ki imajo nizko mero zaupanja. Klasifikator je adaptivnega tipa, kar pomeni, da se skozi celotno razpoznavo strani še dodatno uči in nato adaptiran še enkrat preveri celotno besedilo.

Orodje sicer omogoča učenje dodatnih slogov pisav, a smo se kljub temu odločili, da bo hitrejša pot do izboljšanih rezultatov s pomočjo predobdelave sličic z besedilom. Pri tem sta najpomembnejši operaciji izravnava širine razmakov med znaki in besedami in skaliranje besedila na ustrezno velikost.

3.4 Preverjanje slik in ikon

Ustreznost slik in ikon se preverja s pomočjo deskriptorjev na podlagi histogramov orientiranih gradientov [6]. Enostavnejša rešitev se v začetku zdi zgolj razlika med sličico na zaslonu in sličico, ki jo preverjamo iz urejene baze z dokumentacijo, a je celotno preverjanje mnogo bolj robustno, če se temu ozognemo. Zajete in obrezane sličice se namreč po velikosti lahko razlikujejo od originalov, so pa tudi zajete v barvnem formatu RGB565 in pretvorjene v format RGB888, v katerem so izvirmiki. Pri pretvarjanju izvirmikov v format RGB565 se nekaj informacije izgubi in prav lahko bi prišlo do situacije, da bi pri preverjanju napravili napako.

4. REZULTATI

Predstavili bomo rezultate, kot jih v našem sistemu za testiranje uporabniškega vmesnika dosegajo posamezna metode.

4.1 Klasifikator področij

Za izbiro ustreznega klasifikatorja je bila najprej opravljena 20-delna križna validacija na setu 289 izrezkov področij. Testirali smo uspešnost klasifikatorjev na podlagi nevronskega omrežja, Gaussovih mešanic in metode podpornih vektorjev. Kljub temu, da je nevronskega omrežje doseglo najboljši rezultat, zaradi določenih pomanjkljivosti kasneje ni bilo uporabljeno. Za klasifikacijo smo uporabili klasifikator, ki temelji na metodi podpornih vektorjev in ima naučene ločilne meje za en razred proti nemu.

Nadalje je bil klasifikator testiran še na novih 246 različnih sličicah področij. Pri tem je napačno klasificiral 3.16% področij. Od tega je 62.5% področij z besedilom zamenjal za ikone, 25% ikon za besedilo in 12.5% področij s sliko za glavo/nogo.

4.2 Šivanje slik

Ob neustreznem prekrivanju sosednjih izrezkov zaslona lahko šivanje slik odpove. To se zgodi v primeru, da si bo prvih par vrstic slikovnih elementov dovolj podobnih, da bo ujemanje dodeljeno zadnji od njih, pravo ujemanje pa bo, zaradi rahle zatemnitve slike na zgornjem in spodnjem delu zaslona z receptom, prezrto.

4.3 Razpoznavnik besedila

Razpoznavnik besedila smo testirali na štirih recepturah, pri čemer smo vnesli napake v 40 besed. Pri generaciji napak smo pazili, da je bil njihov spekter čim bolj poln. Napake smo vnašali na začetek in na konec besede, menjali smo velike in male črke ter ločila in številke, zamenjali pa smo tudi nekaj celotnih besed.

Razpoznavnik je pri tem napako naredil le v kontekstu »734 g«, kjer je »g« zamenjal za »9«.

4.4 Primerjalnik slik in ikon

Za testiranje primerjalnika slik in ikon smo uporabili testno množico desetih pravilno in desetih napačno umeščenih sličic, pri tem smo izbrali samo najbolj podobne si pare. Primerjalnik se je v vseh primerih odrezal 100% pravilno.

5. NADALJNJE DELO

V prihodnosti imamo namen sistem narediti povsem avtonomen in namesto operaterja uporabiti robotski manipulator. Ker se je za ozko grlo sistema izkazalo šivanje slik imamo

namen nadgraditi algoritme in uporabiti kompleksnejše iskanje ujemanja na podlagi deskriptorjev.

LITERATURA

1. G. Bradski, The OpenCV Library, *Dr. Dobb's Journal of Software Tools*, 2000.
2. Z. Yu, P. Xiao, Y. Wu, B. Liu in L. Wu, A novel automated gui testing echnology based on image recognition, *High Performance Computing and Communications; IEEE 14th International Conference on Smart City; IEEE 2nd International Conference on Data Science and Systems (HPCC/SmartCity/DSS), 2016 IEEE 18th International Conference on Data Science and Systems*, str. 144–149, IEEE, 2016.
3. C. Grana, L. Baraldi in F. Bolelli, Optimized Block-based Connected Components Labeling with Decision Trees, *IEEE TRANSACTIONS ON IMAGE PROCESSING*, vol. 19, no. 6, str. 1596–1609, 2010.
4. C. Steger, On the Calculation of Moments of Polygons, *FGBV-96-04*, Forschungsgruppe Bildverstehen, Informatik IX, Technische Universität München, Orleansstrase 34, 81667 München, Germany, August 1996.
5. R. Smith, An Overview of the Tesseract OCR Engine, *Proc. Ninth Int. Conference on Document Analysis and Recognition (ICDAR)*, str. 629–633, 2007.
6. N. Dalal, B. Triggs, Histograms of oriented gradients for human detection. *Computer Vision and Pattern Recognition, 2005, CVPR 2005, IEEE Computer Society Conference on*, vol. 1, str. 886-893, IEEE, 2005.

VPLIV NAVIDEZNE RESNIČNOSTI NA ČLOVEKOVE ZAZNAVE

Rok Košir¹, Luc Horvat², Peter Peer¹

¹Fakulteta za računalništvo in informatiko, UL

²Filozofska fakulteta, UM

E-pošta: rokew111@gmail.si, horvat.luc@gmail.com,
peter.peer@fri.uni-lj.si

POVZETEK: *V članku je predstavljen prototip sistema, ki se navezuje na akrofobijo (strahom pred višino), njenemu vplivu na človeka in soočanju z njo v varnem okolju. Sistem je realiziran z virtualnimi očali Oculus Rift in napravo za zaznavanje globine in gibanja Kinect, oboje pa je povezano v navidezni svet, ki je narejen v igralnem pogonu Unreal Engine. Za pravilno izkušnjo smo pripravili tudi ustrezno fizično okolje oziroma poligon, s katerim uporabniku predstavimo varno in nadzorovano izkušnjo v navideznem svetu. Uporabnike smo postavili na poligon in preko kvalitativnih eksperimentov ovrednotili odzive ter tako preverili učinkovitost prototipa. Iz pridobljenih rezultatov sklepamo, da s takšnim pristopom lahko pomagamo ljudem, da se lažje soočajo z nevarnimi izkušnjami in fobijami.*

1. UVOD

Akrofobija [1] oz. strah pred višino se pojavi pri približno 2-5 % populacije, dvakrat bolj pogosto pri ženskah kot pri moških. Ljudje, ki imajo diagnozo za fobijo doživljajo panične napade, kar pripelje do povečanja vznemirjenosti. V takem stanju so preveč nestabilni, da bi si lahko sami pomagali. Pri soočanju z višino moramo paziti, saj se brez nadzora srečamo z dejansko nevarnostjo. Cilj našega projekta je bilo pripraviti prototip, s katerim lahko obvladamo takšno izkušnjo ter jo varno predstavimo uporabnikom. Ker se naloga navezuje na počutje posamezne osebe, smo pripravili različne teste, s katerimi se uporabniki srečujejo in analizirali vplive okolja na njih.

Pregledali smo primere obstoječih raziskav in rešitev, ki naslavlja podobno tematiko [2,3,4,5]. Rezultati so pokazali, da se ob uporabi takšnih sistemov poveča samozavest in zmanjša strah uporabnikov. Pri vseh raziskavah lahko opazimo, da zaznavanje globine in gibanja ni bila ena izmed uporabljenih tehnologij. Zato smo se odločili razširiti našo rešitev s to novo dimenzijo. Pri tem smo uporabili igralni pogon Unreal Engine, ki je poskrbel za prikaz navideznega sveta, povezavo navideznih očal Oculus Rift in senzorja Kinect za zaznavanje globine in gibanja in uporabo teh informacij za upravljanje telesa v navideznem svetu.

2. GLAVNA ORODJA IN OPREMA

2.1 Očala Oculus Rift

Oculus Rift so očala za navidezno resničnost, ki jih razvija in proizvaja podjetje Oculus VR (v lasti Facebooka) (slika 1a). Očala se v glavnem uporabljajo za računalniške igre in v simulacijah. V naši rešitvi smo uporabili Oculus Rift Development Kit 2. Oculus Software Development Kit (SDK) nam omogoči integracijo navidezne resničnosti v sistem.

Očala imajo ločljivost 1920×1080 slikovnih pik oz. 960×1080 slikovnih pik na posamezno oko. Omogočajo hitrost osveževanja 75 Hz na sekundo, z zakasnitvijo 20-40 ms. Njihovo vidno polje je 100 stopinj.



Slika 1: a) Oculus Rift, b) Kinect

2.2 Kinect

Podjetje Microsoft je leta 2010 na trg dalo napravo za zaznavanje globine in gibanja za konzolo Xbox 360. Kinect je kombinacija programske (SDK) in strojne opreme (slika 1b), uporabljen kot vhodna naprava za upravljanje virtualnih likov.

Pri projektu smo želeli uporabiti Kinect v1, vendar v uporabljenem igralnem pogonu ni bilo razvite nobene podpore za to različico. Zato smo se odločili za uporabo v2, s pomočjo vtičnika Kinect4Unreal, namenjenemu zaznavanju oseb. Kinect v2 ima nadgrajen senzor za zajem globine in drugačen način zaznavanja imenovan čas letenja (angl. *time-of-flight*), ki odpravlja pomanjkljivosti prve različice (pred tem so uporabljali pristop s strukturirano svetlobo). Zaradi hitrosti procesiranja moramo napravo priključiti na vhod USB 3.0 s pomočjo posebnega pretvornika.

Kinect ima ločljivost barvne kamere 1920×1080 slikovnih pik, ločljivost globinske slike pa 512×424 slikovnih pik. Ponuja nam vodoravno vidno polje velikosti 70 stopinj in navpično vidno polje velikosti 60 stopinj.

2.3 Igralni pogon Unreal Engine

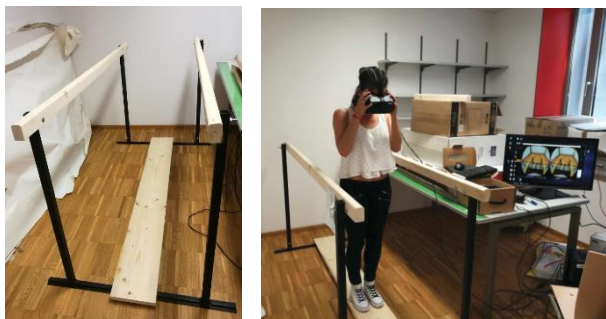
Igralni pogon je razvilo podjetje Epic Games, sprva zgolj za računalniško igro imenovano Unreal. Čeprav je okolje razvito za prvoosebne streljaške igre, se je začel zaradi zmogljivosti uporabljati tudi za druge namene. Pri nalogi uporabimo različico Unreal Engine 4.

2.4 Poligon

Za pravilno delovanje naše rešitve je bilo potrebno urediti prostor in postaviti opremo na ustrezno mesto. Pri tem smo si pomagali z izgradnjo poligona, po katerem so se uporabniki sprehodili med samimi testi (slika 2). Večjo realnost v navideznem svetu lahko namreč dosežemo z dodatnimi fizičnimi elementi, kar pripomore, da uporabniki doživijo izkušnjo na povsem drugi ravni.

Uporabili smo desko ter na njo vezali ograjo, tako da je predstavljal prehod z ograjo. Zgradili smo jo iz železa in lesa in s pametno postavitvijo vijakov poskrbeli za hitro odstranjevanje držal. Poligon je bil dvignjen 3 cm od tal. Deska po kateri so se uporabniki sprehodili je bila dolga 180 cm in široka 20 cm. Ograja pa se je nahajala na višini 110 cm, enake dolžine kot deska ter široka 5 cm. Razdalja med držali je bila 90 cm.

Vse dimenzije poligona v navideznem svetu so se ujemale z dimenzijami fizičnega poligona.



Slika 2: Poligon

3. RAZVOJ NAVIDEZNEGA SVETA

Pred začetkom dela smo načrtali osnove prostora za pripravljene izkušnje. Za pripravo navideznega sveta smo morali pripraviti modele objektov. Igralni pogon nam je omogočil vse logično povezati v primerne scenarije eksperimentov.

3.1 Modeliranje in teksture

Navidezni svet za naš prototip je moral izgledati čim bolj realen. Odločili smo se, da modele pripravimo sami s pomočjo programa Blender. Pri tem pa smo bili pozorni na čim manjši čas upodabljanja elementov.

V programu Blender smo izbrali željeno geometrijsko telo in ga ustrezno preoblikovali. Navidezne elemente smo pripravili po enakih merah kot elemente v realnem svetu. Elementi so v tej fazi brez tekstur, zato jih moramo ustrezno popraviti. Elementom nastavimo povezavo do materiala, prav tako pa za njih pripravimo še shemo mapiranja. Ko smo končali, elemente uvozimo v igralni pogon.

Posameznim elementom smo na podlagi njihovih materialov pripravili teksture s pomočjo programa CrazyBump. Program nam vrne dodatne slike oziroma informacije (angl. *height map*, *bump map*, *normal map*, *displacement map*, *specular map*, *occlusion map*). S pomočjo teh informacij v igralnem pogonu modelom dodamo nov nivo realnosti.

3.2 Igralni pogon

Igralni pogon služi kot osrednje orodje, kjer združimo vse module rešitve. Imamo tudi prostor za manipulacijo objektov. Z uporabo vizualnih skriptnih načrtov (angl. *blueprint*) hitro pripravimo manjše dele kode, kot so recimo teksture. Te pripravimo tako, da vizualne gradnike povežemo ustrezno s slikami in pripravimo material, katerega večemo na uvožene modele.

Modelom v igralnem pogonu prav tako dodamo kolizijo ter jih pravilno postavimo v svetu in tako pripravimo sceno. Za naše teste smo pripravili 2 svetova: notranjega in zunanjega. Prav tako pa smo pripravili nekaj različnih variacij notranjega, kot so brez ograje in s povečano višino.

3.2.1 Potopitev v navidezni svet

Za uporabo v igralnem pogonu je potrebno namestiti Oculus SDK. V igralnem pogonu pripravimo modul kamere, s katero bomo gledali naš svet. Kamero večemo na glavnega karakterja. Nastavimo način premikanja kamere v odvisnosti od pozicije očal in izključimo prvotni senzor gibanja očal.

Pri testiranju moramo biti pozorni, saj moramo ob zagonu usmeriti očala proti Kinectu in jih postaviti v enako višino kot je Kinect. S tem pravilno orientiramo očala in nam jih ni potrebno ponovno umerjati med uporabo v našem svetu.

3.2.2 Zaznavanje gibanja

Kinect smo dodali z namenom, da izboljšamo interaktivnost.

Pred pričetkom uporabe je potrebno na svoj sistem naložiti Kinect SDK. V igralnem pogonu smo med vtičnike dodali Kinect4Unreal, ki nam omogoči uporabo Kinecta.

Zaznavanje se dogaja v vsaki iteraciji programa. Vedno preverimo ali se na zaznani sliki nahaja oseba in si shranimo njeno pozicijo in pozo. To prenesemo na animacijo modela, ki replicira uporabnikovo premikanje. Pri testiranju moramo imeti Kinect na poziciji, iz katere lahko vidi celotno telo človeka. V primeru, da tega ne naredimo, dobimo motnje, saj Kinect aproksimira pozicijo.

3.3 Eksperimenti

3.3.1 Scenarij 1: test različnih višin

Eksperiment odgovarja na vprašanje vpliva višine na dožemanje navideznega sveta. Tako je bilo potrebno pripraviti različne višine mostu, na katerega smo postavili uporabnike. Most je bil steklen. Pripravili smo višini 2 metra in 10 metrov. Pri testu se je spreminjala

višina, izgled prostora je ostal enak (slika 3a). Višini naj bi predstavili izziv uporabniku, saj se naj bi strah z večjo višino stopnjeval.



Slika 3: a) Notranji in b) zunanji navidezni svet

3.3.2 Scenarij 2: test brez ograje

Pri tem testu smo se osredotočili na vpliv ograje pri prečkanju. V prvem testu je soba pripravljena z ograjo, ki jo uporabnik vidi, na realnem poligonu pa se je lahko tudi prime. V tem testu pa je enaka soba pripravljena brez ograje. Ograjo smo odstranili tudi na poligonu, tako da uporabnik ni imel dodatnega oprijema.

3.3.3 Scenarij 3: test zunanje scene

Pri uporabi zunanje scene uporabnika potopimo v popolnoma drug svet v naravi. Podamo mu izkušnjo prečkanja lesenega mostu iz ene pečine na drugo. Pri tem uporabnik nima ograje in se nahaja na 10 metrih nad vodo (slika 3b).

3.3.4 Scenarij 4: dodatni testi

Ko so se uporabniki vživeli v navidezni svet in se v njem dobro počutili, smo jim podali dodatne izzive, kot so stanje na eni nogi, večanje višine in po željah uporabnika na hitro pripravili kakšno dodatno spremembo, recimo spremenili material mostu, zmanjšali svetlobo ipd.

4. REZULTATI

Rezultate smo pridobili s pomočjo vprašalnika, ki smo ga pripravili s psihologom. Pripravljen je tako, da čim bolj povzamemo uporabnikove odzive ob testiranju brez dodatne opreme. Opažanja smo si pisali sproti, prav tako pa smo na koncu testa naredili pregled le-tega z vsakim uporabnikom. Detajli so podani v [6].

Testi so zajemali 13 ljudi, katerih starost je bila med 18 in 49 let. K testom je pristopilo 9 moških in 4 ženske. Izmed 13 ljudi sta bila 2, ki se na višini nista počutila lagodno in 5 s strahom pred višino. 1 izmed njih je imel tudi diagnozo akrofobije. Dva izmed 6 ljudi, ki niso imeli težav z višino, sta že prej uporabljala navidezna očala za zabavo. Vsi ostali pa se nikoli prej niso srečali z navidezno resničnostjo.

4.1 Scenarij 1: test različnih višin

Pri testu nas je zanimalo celostno počutje uporabnika, zato smo si zapisali opažanja pred, med in po izvedenem testu. Pri tem smo bili pozorni na fizične in psihične spremembe. Testi so se izvajali v notranjem svetu.

4.1.1 Občutki na višini 2 metrov

Uporabniki so že na tej višini doživeli opazne spremembe. Najbolj pogoste fizične spremembe so bile: povišan srčni utrip pri 11 uporabnikih, tresoči udi pri 10 ter povišan adrenalin pri 9. Najbolj pogoste psihične spremembe pa so bile: navdušenje in strah pri 11 uporabnikih, zanimanje pri 10, veselje pri 8 ter vrtenje pri 6. Vsi uporabniki so zadan izziv prestali.

4.1.2 Razlike na višini 10 metrov

Pri teh rezultatih smo kategorizirali odzive v 4 sklope: brez sprememb, povečan strah, zabava in brez premikanja. Pričakovali smo, da bo povečana globina prinesla tudi povečan strah. Vendar se je izkazalo, da se je strah povečal samo pri 5 uporabnikih. 3 uporabniki so otrpnili (brez premikanja). Pri 3 ni imelo povečanje višine nobenega vpliva. Pri 2 pa se je spremenilo celo v zabavo.

4.2 Scenarij 2: test brez ograje

Eksperiment brez ograje je potekal v notranjem svetu. Pri testiranju smo uporabnikom odstranili ograjo na fizičnem poligonu, tako smo naredili tudi v navideznem svetu. K testu je pristopilo 11 uporabnikov, od tega 1 ni opravil testa zaradi problemov z ravnotežjem.

Uporabniki so težje prečkali most zaradi manj zaupanja vase, kar se je zgodilo pri 6 uporabnikih, povečanega strahu pri 5 in občutka padanja pri 3.

4.3 Scenarij 3: test zunanje scene

Uporabnike smo postavili tudi v zunanjo sceno, kjer so se nahajali na pečini v naravi in so morali prečkati most. Nahajali so se na višini 10 metrov brez ograje. K testom so pristopili vsi uporabniki, vendar 2 izmed njih samo za pogled v naravo. Večina izmed uporabnikov je občudovala detaljno predstavitev elementov v svetu.

Uporabniki, ki so pristopili k prečkanju so izkušnjo tudi opravili, vendar nekateri bolj počasi; uporabnik z akrofobijo je pristopil, vendar je naredil nekaj korakov in prenehal. Številni odgovori so bili povezani s pomirjenostjo v naravi, iz česar lahko sklepamo, da se z naravnimi izkušnjami lažje soočimo. Pri sami izkušnji so bili uporabniki videti veliko bolj pomirjeni kot pa pri prvih dveh scenarijih. Kar postavlja zanimivo tezo, da so se uporabniki na tej stopnji že navadili na navidezno resničnost in so se zaradi tega lažje soočili s sceno.

4.4 Scenarij 4: dodatni testi

Pri dodatnih testih smo uporabnikom zadali nalogo ali pa spremenili scenarij. Sprva so se kosali s stanjem na eni nogi. 8 uporabnikov je pristopilo k testom in od tega so 3 uspešno prestali nalogo. 3 uporabniki so hoteli preizkusiti večjo višino, zato smo jih postavili na 30 metrov, kjer so bili odzivi podobni kot pri 10 metrih. Pri zunanji sceni pa sta 2 uporabnika preizkusila nočni prehod, kar jima je povečalo strah in sta zgolj z težavo uspela prečkati most.

5. ZAKLJUČEK

Tako kot na vseh področjih ima tehnologija vedno večji vpliv tudi na področju psihologije. Razviti sistem omogoča večji nadzor nad okoljem in fobijo – omogoča možnost definiranja varnih okolij, ki uporabnikom ne predstavljajo nobene prave nevarnosti. Tako pripomorejo k raziskovanju in soočanju s fobijami.

Ugotovili smo, da je pogosto pomembna tudi oprema, saj mora biti odzivna in dobro umerjena. Ob nepravilnostih lahko pride do posledic, kot so vrtoglavica in slabost.

Iz pridobljenih rezultatov sklepamo, da s takšnim sistemom lahko pomagamo ljudem, da se lažje soočajo z nevarnimi izkušnjami in fobijami. Večina uporabnikov resnično verjame, da se nahajajo v pravem in ne navideznem svetu.

Večina uporabnikov se je strinjala, da bi zvok pripomogel k popolnejši izkušnji. Nekateri uporabniki so imeli probleme z umerjanjem očal, kar je pripeljalo do vrtoglavice.

LITERATURA

1. Zitrin, C. M., Klein, D. F., & Woerner, M. G. (1978). Behavior therapy, supportive psychotherapy, imipramine, and phobias. *Archives of General Psychiatry*, 35(3), 307-316.
2. <http://www.vrphobia.com/>.
Virtual reality medical center.
3. Hodges, L. F., Kooper, R., Meyer, T. C., Rothbaum, B. O., Opdyke, D., de Graaff, J. J., ... & North, M. M. (1995). Virtual environments for treating the fear of heights. *Computer*, 28(7), 27-34.
4. Emmelkamp, P. M. G., Krijn, M., Hulsbosch, A. M., De Vries, S., Schuemie, M. J., & Van der Mast, C. A. P. G. (2002). Virtual reality treatment versus exposure in vivo: a comparative evaluation in acrophobia. *Behaviour research and therapy*, 40(5), 509-516.
5. Coelho, C. M., Santos, J. A., Silvério, J., & Silva, C. F. (2006). Virtual reality and acrophobia: one-year follow-up and case study. *CyberPsychology & Behavior*, 9(3), 336-341.

6. Košir, R. (2017). Vpliv navidezne resničnosti na človekove zaznave. Diplomsko delo, Fakulteta za računalništvo in informatiko, UL.

KLASIFIKACIJA SLIK EEG SIGNALOV Z UPORABO GLOBOKEGA UČENJA

Grega Vrbančič, Vili Podgorelec

Inštitut za informatiko

Fakulteta za elektrotehniko, računalništvo in informatiko

E-pošta: grega.vrbancic@um.si

POVZETEK: *Analiza biomedicinskih signalov kot je EEG, za merjenje možganskih aktivnosti omogoča diagnosticiranje različnih kognitivnih nalog in nevroloških motenj. Pogosto so takšni signali pretvorjeni v vizualne predstavitve kot so spektrogrami, ki lahko razkrijejo karakteristične vzorce in služijo kot osnova za klasifikacijo. Za namen klasifikacije EEG signalov subjektov z in brez motoričnih okvar oz. motenj, smo oblikovali novo metodo, ki uporablja spektrogramске slike kot vhod v globoko konvolucijsko nevronsko mrežo, brez kakršnihkoli predhodnih izbir ali ekstrakcij značilk. Rezultati, pridobljeni brez kakršnegakoli človeškega posega in z uporabo privzetih vrednosti parametrov, so se izkazali, da ne zaostajajo veliko za najsodobnejšimi metodami, ki izkoriščajo uporabo domenskega znanja za analizo EEG signalov.*

1. UVOD

V zadnjih letih je opaziti velik porast raziskav na področju procesiranja in analiziranja elektroencefalografskega (EEG) signala s ciljem, da bi z informacijami zbranimi iz teh signalov lahko diagnosticirali nekatere vrste nevroloških motenj. Nabor možnih diagnoz, ki jih je mogoče diagnosticirati na tak način se iz leta v leto povečuje in pokriva široko področje nevroloških bolezni, od blagih kognitivnih okvar [1], raznih neurodegenerativnih bolezni kot je Alzheimerjeva bolezen [2], pa do najbolj resnih oblik nevroloških motenj, kot so amiotrofična lateralna skleroza ali cerebralna paraliza, kjer je lahko bolnik resno telesno prizadet ali celo popolnoma paraliziran [3].

Osnovo EEG signalov predstavlja električna aktivnost merjena kot napetost na različnih točkah možganov. Te signale, ki so običajno časovno spremenljivi in v naravi nestacionarni lahko raziskujemo z uporabo različnih tehnik obdelave signalov. Čeprav je obseg analize signalov EEG zelo širok, je splošen način tak, da surove signale zajete iz elektrod na lasišču subjekta predhodno obdelamo in filtriramo, da ti postanejo berljivi [4]. Te signale lahko zatem analiziramo in/ali preoblikujemo z namenom pridobitve poglobljenega vpogleda v delovanje možganov. Po preoblikovanju, lahko signale tudi vizualno predstavimo v obliki spektrogramov za vsak posamezen kanal. Takšna vizualna

predstavitev nam lahko razkrije vzorce sprememb v aktivnosti nihanja možganov med opravljanjem vnaprej določenih aktivnosti ali nalog.

Z nedavnim napredkom na področju strojnega učenja, z uporabo globokega učenja in globokih arhitektur nevronske mreže (NN), je mogoče za opravljanje nalog klasifikacije slik, doseči enako dobre ali še boljše rezultate, kot jih dosegamo ljudje [5].

Glede na prej omenjene vzpodbudne rezultate pri klasifikaciji slik, predlagamo novo metodo za klasifikacijo spektrogramskih slik EEG signalov za zaznavo motoričnih motenj oziroma okvar, z uporabo globoke konvolucijske nevronske mreže. Glavna cilja raziskave sta bila raziskati ali se lahko predlagana metoda uporabi za uspešno klasifikacijo nevroloških motenj ter kako ustrezno pretvoriti surove EEG signale v spektrogramske slike, primerne za učenje NN.

Preostali del prispevka je organiziran sledeče. V poglavju 2 je podrobneje predstavljena predlagana metoda, poglavje 3 opisuje izveden eksperiment in njegove rezultate, poglavje 4 pa zaključuje naš prispevek z zaključnimi mislimi ter možnostmi za nadaljnje delo.

2. PREDLAGANA METODA

Predlagana metoda je sestavljena iz več korakov in sicer iz pred-procesiranja signala, pretvorbe signala v spektrograme ter učenja globoke konvolucijske mreže na ustvarjenih spektrogramskih slikah. V nadaljevanju so podrobneje opisani posamezni koraki predlagane metode.

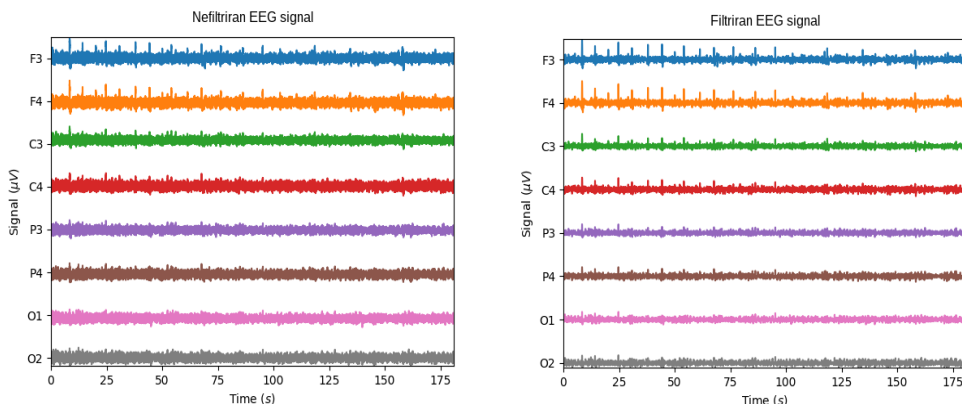
2.1 Pred-procesiranje EEG signala

Zajet EEG signal smo pred-procesirali in vizualizirali s pomočjo odprto-kodne programske platforme CEBL3. Posamezen posnetek EEG signala je dolg 3 minute in je zajet z osmimi elektrodami (C3, C4, F3, F4, O1, O2, P3 in P4) pri hitrosti vzorčenja 256Hz. EEG signalne ritme lahko okarakteriziramo kot [2]:

- delta valovi (0Hz – 4Hz),
- theta valovi (4Hz – 8Hz),
- alpha valovi (8Hz – 14Hz),
- nizki beta valovi (14Hz – 20Hz),
- visoki beta valovi (20Hz – 30Hz) in
- nizki gamma valovi (30Hz – 60Hz).

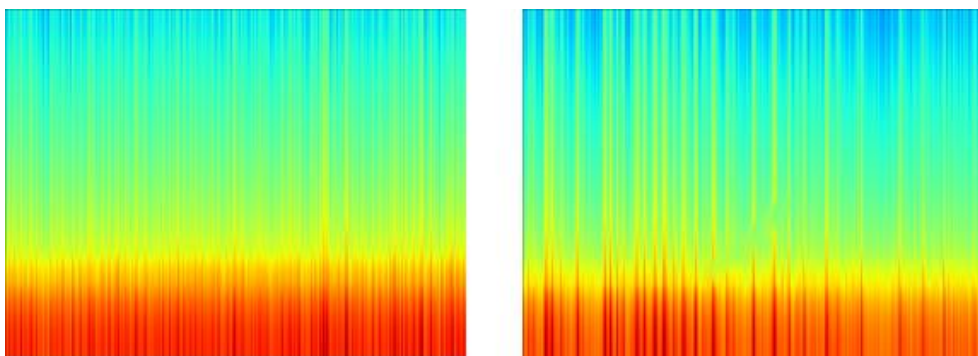
Na vsakem izmed kanalov smo opravili filtriranje v frekvenčnem pasu od 0.5Hz do 7.5Hz z uporabo Butterworth filtra (Slika 1), z namenom odstranitve šumov na nizkih in visokih frekvencah ter za odstranitev drugih artefaktov. Omenjen frekvenčni pas smo izbrali na podlagi že opravljenih raziskav na tem področju [2], [6], katerih izsledki kažejo, da je največ spektralnih sprememb moči mogoče zaznati na nivoju alpha in theta valov, kateri v večini primerov kažejo na neko vrsto možganske patologije. Za pretvorbo

filtriranega EEG signala v časovno-frekvenčno domensko predstavitev smo uporabili Hitro Fourierjevo transformacijo (angl. Fast Fourier Transformation), ki velja za eno najbolj primernih metod za transformacijo signalov med časovno in frekvenčno domeno [7].



Slika 1: Prikaz EEG signala po posameznem kanalu pred (levo) in po (desno) opravljenem filtriranju.

Po transformaciji signala v časovno-frekvenčno domensko predstavitev, smo za vsak posamezen kanal izrisali spektrograme, ki predstavljajo gostoto moči signala pri uporabi logaritemske normalizacije (Slika 2).



Slika 2: Spektrograma EEG signala na kanalu C3. Levo je prikazan spektrogram motene, desno pa nemotene možganske aktivnosti.

2.2 Učenje in klasifikacija z globoko konvolucijsko nevronske mrežo

Zadnjih nekaj let je strojno učenje, še posebej pa globoko učenje, v razcvetu. Globoko učenje je podmnžica splošnejšega področja umetne inteligence, ki mu pravimo strojno

učenje. Do nedavnega je večina strojnega učenja in tehnik procesiranja signalov izkoriščala plitvo strukturirane arhitekture, ki so navadno sestavljene iz enega oziroma največ dveh nivojev nelinearnih transformacij atributov (npr. logistična regresija, modeli z maksimalno entropijo). Slednje so se izkazale za zelo učinkovite pri reševanju lažjih ali zelo omejenih problemov, po drugi strani pa njihovo omejeno modeliranje in predstavitevna moč povzročata težave pri reševanju kompleksnejših problemov iz realnega sveta, ki vsebujejo naravne signale (govor, zvok, slika) [8], [9].

V splošnem je algoritem globokega učenja sestavljen iz hierarhične arhitekture s številnimi nivoji, od katerih vsak izmed njih predstavlja nelinearno enoto za obdelavo informacij. Globoke nevronske mreže (angl. Deep Neural Networks – DNNs), ki uporabljajo globoke arhitekture NN, lahko predstavljajo funkcije z večjo kompleksnostjo, v kolikor je število nivojev in število enot znotraj nivoja povečano [10].

Za namene klasifikacije slik (zaznava in prepoznavna obrazov [11], segmentacija bioloških slik [6], prepoznavna prometnih znakov [12]) se je še za posebej uspešno izkazala arhitektura globokih nevronskih mrež imenovana konvolucijska nevronska mreža (angl. Convolutional Neural Network – CNN), katere delovanje inspirira iz organizacije vizualnih korteksov pri živalih [10].

Standardna arhitektura CNN vsebuje konvolucijske nivoje, organizirane v obliki sklada, katerim opcijsko sledi nivo združevanja (poznani tudi kot nivo podvzorčenja), katerim sledi en ali več polno povezanih nivojev. Konvolucijski nivo je navadno sestavljen iz več polj značilk (angl. Feature maps), z drugačnimi vektorji utežmi, z namenom prepoznave več lastnosti na posamezni lokaciji. Zamikanje vhoda konvolucijskega nivoja, posledično zamakne tudi izhod, sicer pa ga ohrani nespremenjenega. Ko je značilka zaznana, njena točna lokacija postane manj pomembna, dokler njena relativna pozicija glede na ostale ostane približno enaka. Zato je pogosto, da vsakemu konvolucijskemu nivoju sledi dodatna plast, ki opravlja lokalno povprečenje in podvzorčenje. Rezultat tega je manjša ločljivost polja značilk in zmanjšana občutljivost vhoda na zamikanje in izkrivljanja. Sledeči si konvolucijski sloji in sloji pod-vzorčenja se običajno izmenjujejo, kar ima za posledico bi-piramido: na vsakem nivoju je število polj značilk povečano, medtem ko se prostorska ločljivost zmanjša [13]. Omenjenim, izmenjaje sledečim si nivojem, tipično sledi eden ali več polno povezanih nivojev, kateri imajo polne povezave do vseh aktivacij v prejšnjem nivoju, kot pri klasičnih NN.

3. EKSPERIMENT IN REZULTATI

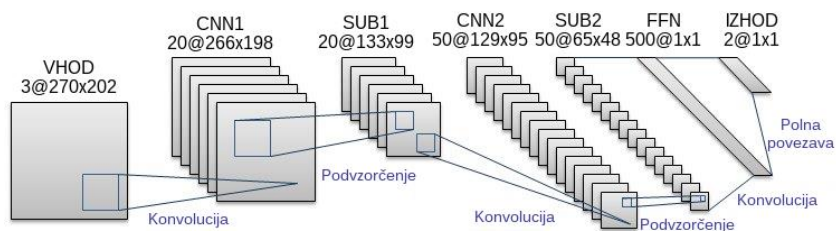
V tem poglavju je podrobneje predstavljena uporabljena podatkovna množica, arhitektura CNN ter nastavitve učenja. V zadnjem podpoglavju so predstavljeni rezultati opravljenega eksperimenta.

3.1 Podatkovna množica

Uporabljena je bila podatkovna množica BCI [14], zbrana s strani Colorado State University. EEG signali so bili zajeti z uporabo g.GAMMASys aktivnih elektrod (8 kanalov), z frekvenco vzorčenja 256Hz in strojnim filtriranjem signala frekvenčnega pasu od 0.5Hz – 100Hz [15]. Dolžina posameznega zajetega posnetka je 3 minute. Podatkovna množica vsebuje posnetke skupno 13 subjektov, od katerih 9 subjektov nima nobenih znanih zdravstvenih težav (t.i. unimpaired), preostali pa trpijo za hudimi motoričnimi motnjami (t.i. impaired).

3.2 Konvolucijska nevronska mreža

Naša CNN arhitektura (Slika 3) temelji na LeCun-ovi petnivojski LeNet arhitekturi [16] z manjšimi prilagoditvami parametrov konvolucijskih nivojev in učnimi parametri. CNN sestoji iz dveh konvolucijskih nivojev z filtrom velikosti 5 x 5 ter zamikom velikosti 1 x 1. Vsakemu izmed njiju sledi nivo podvzorčenja z uporabo maksimizacijske funkcije s filtrom in zamikom velikosti 2 x 2. Na prvem konvolucijskem nivoju je apliciranih 20, na drugem pa 50 filtrov. Za tem skladom nivojev imamo en polno povezan nivo z 500 izhodnimi filtri povezan z izhodnim nivojem, kateri klasificira posamezno sliko kot »impaired« ali »unimpaired«.



Slika 3: Uporabljena arhitektura konvolucijske nevronske mreže.

3.3 Učenje

Zaradi omejenega, malega števila učnih primerkov, smo za učenje CNN uporabili in validacijo uporabili princip »izpusti enega«. Učenje je bilo izvedeno ločeno za vsak kanal v zajetem posnetku EEG signala, s čemer smo ustvarili 8 modelov, ki smo jih nato z uporabo algebralnega kombinacijskega pravila – glasovanje večine, združili v ansambel. Ansambel je podan subjekt klasificiral kot »impaired«, v kolikor ga je več kot polovica kanalov klasificirala kot »impaired«, sicer ga je klasificiral kot »unimpaired«. Učenje je bilo izvedeno z uporabo enega paketa dvanajstih slik, ter eno sliko za testiranje. Za stopnjo učenja smo uporabili vrednost 1×10^{-6} , uteži smo inicializirali z uporabo inicializacijske sheme Xavier [17], za aktivacijsko funkcijo pa smo uporabili ReLU [18]. Za vsako učenje CNN smo izvedli 50 ponovitev z 1 iteracijo.

3.4 Rezultati

Rezultati klasifikacije pridobljeni na testnih primerkih za vsak kanal posamezno in združeno z uporabo glasovanja večine so predstavljeni v tabeli 1. Subjekti označeni s s21 – s25 predstavljajo osebe brez kakršnikoli znanih nevroloških motenj (označeno z F – negativni primeri), medtem ko subjekti označeni s s11 – s16 predstavljajo paciente z motoričnimi okvarami oziroma motnjami (označeno s T – pozitivni primeri).

Tabela 1: Rezultati klasifikacije za posamezen kanal kot tudi končna klasifikacija z uporabo glasovanja večine.

subjekt→ kanal ↓	s21	s24	s23	s27	s28	s22	s26	s20	s25	s11	s13	s15	s16
C3	F	F	F	F	F	F	F	F	F	F	F	F	T
C4	F	F	T	T	F	F	T	F	T	T	F	F	F
F3	T	T	T	F	F	F	F	F	F	F	F	F	F
F4	T	F	F	F	T	F	T	F	F	F	F	F	T
O1	T	T	T	F	T	T	F	F	F	F	F	F	F
O2	F	F	T	F	F	T	F	F	F	F	T	T	T
P3	F	F	T	F	F	F	T	F	T	F	T	F	T
P4	F	T	T	F	F	F	T	T	T	F	T	F	T
razred	F	F	T	F	F	F	F	F	F	F	F	F	T

Kot je razvidno, je naša metoda pravilno klasificirala 8 izmed 9 subjektov brez znanih nevroloških motenj in 1 od 4 pacientov z motoričnimi okvarami oziroma motnjami. S čimer smo dosegli splošno natančnost 69,23%, F1 vrednost 65,64%, občutljivost 25% ter specifičnost 88,89%. V primerjavi z najnaprednejšo klasifikacijsko metodo [15], ki je dosegla splošno natančnost okrog 74%, lahko ugotovimo, da naša predlagana metoda ne zaostaja veliko (manj kot 5%). Še posebej, če ob tem upoštevamo popolnoma samodejno pretvorbo EEG signalov v spektrogramske slike ter klasifikacijo brez domenskega znanja ter ročne izbire ali ekstrakcije značilk.

4. ZAKLJUČEK

V tem članku smo predstavili proces transformacije EEG signala v slike spektrogramov z uporabo Butterworth filtriranja frekvenčnega pasu od 0.5Hz – 7.5Hz, Hitro Fourierjevo transformacijo in logaritmično normalizacijo za predstavitev gostote moči transformiranega signala. Tako ustvarjene slike, smo zatem uporabili za učenje naše CNN z uporabo principa »izpusti enega«. Pridobljeni rezultati so obetajoči, še posebej če upoštevamo popolno avtomatizirano naravo predlagane metode.

V prihodnosti, bi želeli naše delo razširiti z uporabo večjih podatkovnih množic, različnih časovno-frekvenčnih domenskih transformacij signalov, različnih CNN arhitektur in nastavitve parametrov učenja.

LITERATURA

- [1] M. Kashefpoor, H. Rabbani, and M. Barekatin, "Automatic Diagnosis of Mild Cognitive Impairment Using Electroencephalogram Spectral Features.," *J. Med. Signals Sens.*, vol. 6, no. 1, pp. 25–32, 2016.
- [2] V. Podgorelec, "Analyzing EEG Signals with Machine Learning for Diagnosing Alzheimer ' s Disease," no. X, pp. 61–64, 2012.
- [3] C. Neuper, G. . Müller, A. Kübler, N. Birbaumer, and G. Pfurtscheller, "Clinical application of an EEG-based brain–computer interface: a case study in a patient with severe motor impairment," *Clin. Neurophysiol.*, vol. 114, no. 3, pp. 399–409, Mar. 2003.
- [4] J. Kaur and A. Kaur, "A review on analysis of EEG signals," in *2015 International Conference on Advances in Computer Engineering and Applications*, 2015, pp. 957–960.
- [5] D. Cireşan, U. Meier, and J. Schmidhuber, "Multi-column Deep Neural Networks for Image Classification," Feb. 2012.
- [6] K. Kamnitsas, C. Ledig, V. F. J. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, D. Rueckert, and B. Glocker, "Efficient Multi-Scale 3D CNN with Fully Connected CRF for Accurate Brain Lesion Segmentation," Mar. 2016.
- [7] M. Akin, "Comparison of wavelet transform and FFT methods in the analysis of EEG signals," *J. Med. Syst.*, vol. 26, no. 3, pp. 241–247, 2002.
- [8] L. Deng and D. Yu, *Deep Learning: Methods and Applications*. 2014.
- [9] N. Buduma and N. Locascio, *Fundamentals of Deep Learning*, First Edit. O'Reilly Media, Inc., 2017.
- [10] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi, "A survey of deep neural network architectures and their applications," *Neurocomputing*, vol. 234, pp. 11–26, 2017.
- [11] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "DeepFace: Closing the Gap to Human-Level Performance in Face Verification," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1701–1708.
- [12] D. C. Cireşan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, "Mitosis Detection in Breast Cancer Histology Images with Deep Neural Networks," Springer, Berlin, Heidelberg, 2013, pp. 411–418.
- [13] Y. LeCun and Y. Bengio, "Convolution Networks for Images, Speech, and Time-Series," in *The handbook of brain theory and neural networks*, M. A. Arbib, Ed. MIT Press, 1995.
- [14] Colorado State University, "EEG dataset from BCI at CSU." [Online]. Available: http://www.cs.colostate.edu/eeg/main/data/2011-12_BCI_at_CSU. [Accessed: 31-Jan-2018].
- [15] E. M. Forney, C. W. Anderson, W. J. Gavin, P. L. Davies, M. C. Roll, and B. K. Taylor, "Echo State Networks for Modeling and Classification of EEG Signals in

- Mental-Task Brain-Computer Interfaces,” 2015.
- [16] Y. Le Cun, L. D. Jackel, B. Boser, J. S. Denker, H. P. Graf, I. Guyon, D. Henderson, R. E. Howard, and W. Hubbard, “Handwritten digit recognition: applications of neural network chips and automatic learning,” *IEEE Commun. Mag.*, vol. 27, no. 11, pp. 41–46, 1989.
 - [17] B. Xu, N. Wang, T. Chen, and M. Li, “Empirical Evaluation of Rectified Activations in Convolution Network,” *ICML Deep Learn. Work.*, pp. 1–5, 2015.
 - [18] K. G. Pasi and S. R. Naik, “Effect of parameter variations on accuracy of Convolutional Neural Network,” in *2016 International Conference on Computing, Analytics and Security Trends (CAST)*, 2016, pp. 398–403.

PREPOZNAVA TEKAČEV NA FOTOGRAFIJAH

Miha Grgič Jelen, dr. Borut Batagelj

Fakulteta za računalništvo in informatiko

Univerza v Ljubljani

E-pošta: miha.grgic.jelen@gmail.com

POVZETEK V članku je predstavljen postopek, ki smo ga razvili v sklopu diplomske naloge na fakulteti za računalništvo in informatiko, kot rešitev za problem identifikacije tekačev na fotografijah športnih prireditev. Postopek najprej izolira posamezne osebe, nato za vsako osebo definira tarčno regijo v kateri se najverjetneje nahaja štartna številka. V naslednjem koraku izvede segmentacijo tarčne regije in izolira posamezne številke. Nato vsak izsek v katerem se potencialno nahaja številka, poskušamo prepoznati z optičnem branjem znakov, z uporabo orodja *TesseractOCR*. Na koncu prepoznane številke dodatno filtriramo in združimo v štartno številko. Razvit je bil tudi uporabniški vmesnik, ki nam močno olajša pregledovanje fotografij in vizualizacijo rezultatov. Na koncu predstavimo nerešene probleme ter nekaj idej za nadaljnji razvoj.

1. UVOD

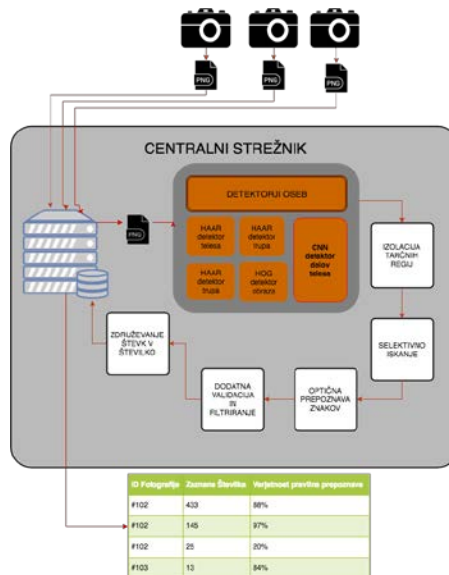
Kot študent sem več let delal kot fotograf tekaških prireditev v Sežani. Nastale fotografije so končale v obsežni spletni galeriji. Skoraj vsak udeleženec po prihodu domov pregleda svoje rezultate ter galerijo, v upanju da bo našel kakšno svojo fotografijo. V tem trenutku, pa je njegova vztrajnost na resnem preizkusu, saj je v spletni galeriji zelo veliko števil fotografij. V večini primerov z iskanjem odneha po nekaj desetih fotografijah.

Najprej smo pregledali, če takšna tehnološka rešitev že obstaja. Izkazalo se je, da je na internetu dostopnih ogromno aplikacij ter rešitev, namenjenih prepoznavi teksta ter klasifikaciji fotografij, kot so *TesseractOCR* [1] in *Google Vision* [3], vendar nobena od njih ni specializirana za prepoznavo štartnih števil. Še najbližje rešitve so tiste, ki služijo za programsko branje registrskih tablic avtomobilov, npr. knjižnica *OpenALPR* [2]. Vendar, ko smo jih poskusili prirediti in uporabiti na našem primeru, se je izkazalo, da je naš problem kompleksnejši in nismo prišli do uporabnih rezultatov. Tako je nastala ideja o razvoju prototipa za računalniško identifikacijo oseb na fotografijah, z namenom da bi udeleženci prireditev lahko enostavno dostopali do svojih fotografij.

Shema na sliki 1 prikazuje celoten proces prepoznave štartne številke. Fotografiji, ki so delovali na terenu, fotografije najprej pretočijo na centralni strežnik, kjer se shranijo v primerno strukturo. Nato na vsaki fotografiji iz zbirke izvedemo postopke, ki iz slike prepoznajo štartne številke.

Začnemo z izolacijo tarčnih regij (podrobneje opisano v sekciji 2.1), v katerih pričakujemo, da se nahajajo štartne številke. Za doseg tega najprej uporabimo detektorje različnih delov telesa. Na podlagi dobljenih rezultatov lahko nato izračunamo in izoliramo tarčne regije.

V naslednjem koraku se ukvarjamo z detekcijo posameznih števk, ki se nahajajo v tarčni regiji (podrobneje opisano v sekciji 2.2). Na vsaki izmed regij izvedemo algoritem selektivnega iskanja. Tako dobimo nove izseke slike, na kateri se potencialno nahaja posamezna številka. Na vsakem od njih izvedemo postopek optične prepoznave znakov. Na pridobljenih rezultatih izvedemo dodatno validacijo in filtriranje glede na verjetnost pravilne optične prepoznave, razmerja, lokacije in velikosti znaka. Na koncu rezultate združimo v številko.



Slika 1: Shema celotnega postopka detekcije štartnih številke

2. POSTOPEK DETEKCIJE ŠTARTNE ŠTEVILKE

2.1 Izolacija tarčne regije

Cilj v prvem koraku postopka je, da izoliramo regije, ki potencialno vsebujejo štartno številko. To lahko storimo na več načinov oz. uporabimo lahko različne detektorje. Nekateri od njih detektirajo obraz, drugi trup, tretji pa posamezne dele telesa. Na podlagi rezultatov omenjenih detektorjev, lahko izračunamo tarčno regijo, kjer je največja verjetnost, da se štartna številka nahaja. Na njej bomo nato izvajali nadaljnje operacije, ki nas bodo pripeljale do končnih rezultatov. Med preizkušenimi detektorji se je najbolje izkazal detektor, imenovan *Realtime Multi-Person 2D Pose Estimation using Part*

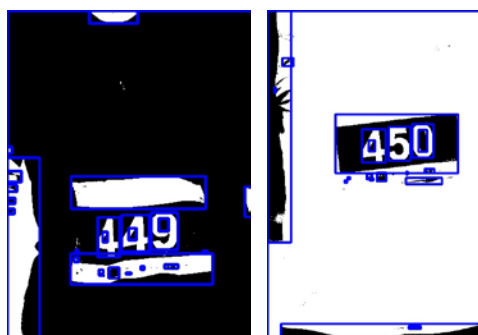
Affinity Fields [4]. Posebnost tega detektorja je, da v realnem času (ob uporabi GPU enote) detektira različne dele telesa (obraz, rame, boki, kolena, stopala...). V primerjavi z drugimi detektorji, kot je npr. detektor obraza, je pri uporabljenem detektorju tarčna regija veliko bolj natančno definirana in ne vsebuje delov drugih tekačev (slika 2).



Slika 2: Primer detekcije oseb in delov telesa z uporabo detektorja *Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields*[4]. Skelet in točke, ki predstavljajo dele telesa so označene z rdečo in rumeno barvo, izolirana tarčna regija pa z zeleno.

2.2 Detekcija števk

V tem delu postopka, znotraj že definiranih tarčnih regij, najprej poiščemo t.i. kandidate, na katerih bomo kasneje izvedli algoritem, ki bo detektiral ali gre za števkico ali ne. Na binariziranem delu fotografije smo uporabili algoritem selektivnega iskanja [6]. Uporabili smo metodo *find_candidate_object_locations*, ki je del knjižnice *Dlib* [5]. Med rezultati ostanejo nekatere nepravilnosti. Lahko se bi jih znebili z uporabo *strožjega filtriranja*, vendar bi s tem tudi izpustili kakšno števkico. Na tej točki jih ignoriramo, saj bomo zanje poskrbeli kasneje v postopku. Tako so rezultati v veliki večini primerov res izolirane števkice (slika 3).



Slika 3: Primer rezultatov z uporabo algoritma *Selective Search for Object Recognition* [6]



Slika 4: Primer filtriranih rezultatov z uporabo algoritma *Selective Search for Object Recognition* [6]

Na predstavljenem primeru (slika 2) so na fotografiji številke bele barve, ozadje pa rdeče. Tako po izvedbi postopka binarizacije dobimo slike števk, ki imajo črno ozadje, sama številka pa je bele barve (slika 4). Na sliki, ki je rezultat binarizacije izvedemo inverz in dobimo posamezne številke zapisane črne na belem ozadju.

Na vsakem rezultatu algoritma *selektivno iskanje*, izvedemo optično prepoznavo znakov. To storimo z orodjem *TesseractOCR* [1]. Rezultati vsebujejo zaznano številko ter verjetnost pravilne detekcije.

Tabela 1: Način oblikovanja posameznih delov prispevka

Detektirane številke	Verjetnost pravilne detekcije
1	71%
0	93%
5	94%
4	90%

Na tej točki, imamo za vsako detekcijo dovolj podatkov, da lahko izvedemo *nekoliko pametnejše* filtriranje rezultatov. To storimo predvsem z razlikovanjem rezultatov, ki imajo visoko verjetnost pravilne detekcije (v nadaljevanju VVPD), v kombinaciji s preostalimi lastnostmi. V prvem koraku definiramo številke z VVPD (tabela 1). Med njimi so tudi anomalije, ki jih bomo poskušali odstraniti s pomočjo dodatnega filtriranja glede na razliko v višini, horizontalno oddaljenost od ostalih števk VVPD in vertikalno oddaljenost od števk VVPD.

2.3 Združevanje števk v številko

Na koncu preostale številke (rezultate, ki so prestali validacijo) le še združimo v celotno štartno številko, na podlagi horizontalnega položaja posameznega rezultata.

3. REZULTATI

Postopek smo preizkusili na različnih zbirkah fotografij, naključno izbranih iz različnih tekaških dogodkov v Sežani (*Mali kraški maraton*, *Tekaški pozdrav jeseni*). Na večini

izmed njih so 2 ali 3 osebe, na nekaterih je tudi 10 oseb, na drugih pa samo 1. Večina je posneta v ležeči kompoziciji, nekatere izmed njih pa v pokončni. V prvi zbirki (bele številke na rdečem ozadju), se na fotografijah pojavi 145 štartnih števil, metoda jih je pravilno prepoznala 85 kar pomeni 58,6% uspešnost (*eng. true positive rate*). V 53 primerih je metoda detektirala napačno številko (*eng. false positive*), izpustila pa je 60 štartnih števil, to je 41,4% (*eng. false negative*).

V drugi zbirki (črne številke na belem ozadju), se na fotografijah pojavi 78 štartnih števil, metoda jih je pravilno prepoznala 34, kar pomeni 43,5% uspešnost. V 41 primerih je metoda detektirala napačno številko, izpustila pa je 44 štartnih števil, to je 56,4%.

V trenutni fazi rezultati niso zadovoljivi. Tipične napake so:

- Detektor oseb ne detektira posamezne osebe, torej se za dano osebo tarčna regija sploh ne izračuna ter optična prepoznavna znakov sploh ne izvede,.
- Zaradi različnih barv tekaških majic pride do nepravilne binarizacije. Posledica tega je, da algoritem selektivnega iskanja ne izolira pravilno posameznih števk za kasnejšo prepoznavo.
- V primerih, ko je na sliki oseba delno zakrita (pred njo stoji druga oseba) niso vidne vse številke, ki sestavljajo štartno številko. Tako sistem prepozna samo nekaj izmed njih.

Za rešitev problema oseb, ki jih nismo našli, bi lahko uporabili kombinacijo več različnih detektorjev za definicijo tarčne regije. Detektor obraza, bi v primeru delno prekrita telesa še vedno zaznal osebo, v kolikor je na fotografiji viden obraz. Tako bi zmanjšali število manjkajočih detekcij.

Težave z binarizacijo bi lahko odpravili, če bi tarčno regijo zmanjšati na pravokotnik v katerem se nahaja štartna številka. Binarizacija po metodi *Otsui* [7] bi tako zelo uspešno binarizirala sliko in ločila številke od ozadja. Poskusili bi lahko tudi z drugimi algoritmi za izolacijo teksta iz ozadja [8] in naravnih scen [9].

4.1 NEREŠEN PROBLEM PREKRITIH ŠTEVK

Eden izmed težjih nerešenih problemov, so prekrite številke. Kot lahko vidimo na sliki 5 se velikokrat zgodi, da eno izmed števk, ki sestavlja štartno številko, prekriva drugi tekač ali del telesa. Štartna številka na sliki 5, bi bila prepoznana kot 37. V danem primeru je rezultat napačen, saj se sistem ne zaveda skrite številke, ki se skriva pod tekačevo roko.

Ena od možnih rešitev bi bili markerji. Vsak štartni listek bi vseboval 4 markerje - vsakega v enem od kotov štartne številke. Markerji so grafike, ki jih enostavno prepoznamo s programskim orodjem. V kolikor na prepoznani številki ne moremo prešteti vseh 4 markerjev, rezultat označimo kot neveljaven. Problem tega pristopa je, da bi najverjetneje zavrgli tudi veliko pravih rezultatov. Hkrati bi morali z organizatorjem uskladiti obliko štartne številke, da bi vsebovala markerje.

Naslednja možna rešitev bi bilo preverjanje razmerja stranic pravokotnika, ki zaobjema številke. V kolikor pravokotnik ne ustreza pravemu razmerju, rezultat zavržemo, saj je velika verjetnost da je vsaj ena izmed števk prekrita. Tak pristop daje dobre rezultate samo v primeru, da so številke na unikatnem ozadju (kot na primer rdeče ozadnje na sliki 5). V kolikor številke nimajo enakomernega oz. kontrastnega ozadja, pristop ne deluje več. Tudi v tem primeru bi bilo potrebno usklajevanje z organizatorjem glede oblike štartne številke.



Slika 5: Primer fotografije z delno prekrito štartno številko

5. ZAKLJUČEK

Predstavili smo problem identifikacije tekačev na fotografijah ter motivacijo za implementacijo rešitve. Ugotovili smo, da je najprimernejša metoda identifikacije osebe, da prepoznamo štartno številko, ki jo vsak tekač nosi na dresu. To smo dosegli tako, da smo najprej s pomočjo različnih detektorjev objektov iz fotografije izolirali posamezne osebe. Ugotovili smo, da se je najbolj izkazal detektor posameznih delov telesa, ki uporablja konvolucijske nevronske mreže. Na podlagi dobljenih rezultatov, smo zlahka definirali tarčno regijo v kateri se nahaja štartna številka. Nato smo na njej izvedli algoritem selektivnega iskanja, ki nam je definiral izseke slik, na katerih se morda nahaja posamezna številka. Na vsakem izseku smo s pomočjo orodja *TesseractOCR* izvedli optično detekcijo znakov. Vse rezultate smo dodatno validirali in filtrirali. Dobljene rezultate smo na koncu združili v štartno številko.

Razvili smo tudi sodoben uporabniški vmesnik, ki služi pregledovanju fotografij in vizualizaciji rezultatov. Z njegovo pomočjo smo algoritme in različne korake postopka veliko lažje prilagajali, saj smo lahko na pregleden način hitro opazili razlike ter shranjevali vmesne rezultate.

Problem identifikacije tekačev, nam je predstavljala velik izziv. Vseskozi smo imeli v mislih kako bi razvili funkcionalen sistem in ne zgolj prototipa. Tekom razvoja smo se

veliko naučili, še ogromno pa ostaja odprtega. Želja je da bi sistem dodatno izpopolniti in preizkusiti v produkcijskem okolju.

6. LITERATURA

1. Tesseract Open Source OCR Engine (main repository), dostopno na: <https://github.com/tesseract-ocr/tesseract>.
2. OpenALPR Technology, Inc, Automatic License Plate Recognition library, dostopno na: <https://github.com/openalpr/openalpr>.
3. Google, Inc, Cloud Vision API, dostopno na: <https://cloud.google.com/vision/>.
4. Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime multi-person 2d pose estimation using part affinity fields," 2017 IEEE Conference on Computer Vision and Pattern Recognition, str. 1302–1310, 2017.
5. Davisking, A toolkit for making real world machine learning and data analysis applications in C++, dostopno na: <https://github.com/davisking/dlib>.
6. J. R. R. Uijlings, K. E. A. van de Sande, T. Gevers, and A. W. M. Smeulders, "Selective search for object recognition," International Journal of Computer Vision, vol. 104, št. 2, str. 154–171, 2013.
7. Chen Yu, Chen Dian-ren, Li Yang and Chen Lei, "Otsu's thresholding method based on gray level-gradient two-dimensional histogram," 2010 2nd International Asia Conference on Informatics in Control, Automation and Robotics (CAR 2010), Wuhan, 2010, str. 282-285.
8. Neumann L., Matas J.: Real-Time Scene Text Localization and Recognition, CVPR 2012. dostopno na: <http://cmp.felk.cvut.cz/~neumalu1/neumann-cvpr2012.pdf>.
9. B. Epshtein, E. Ofek and Y. Wexler, "Detecting text in natural scenes with stroke width transform," 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, 2010, str. 2963-2970.



SLOVENSKI PODJETNIŠKI SKLAD

Javni sklad Republike Slovenije za podjetništvo (skrajšano Slovenski podjetniški sklad)
Ulica kneza Koclja 22
2000 Maribor
T: +386 (0)2 234 12 60
E: info@podjetniskisklad.si
www.podjetniskisklad.si

Smo javna finančna institucija, ki zagotavlja boljši dostop mikro, malim in srednje velikim podjetjem do ugodnih finančnih virov na trgu.

V letu 2018 bo podjetjem na razpolago 106 mio EUR različnih finančnih spodbud za večje investicije in za obratna sredstva. Tako bomo preko 1.000 podjetjem zagotovili nadaljnjo rast oz. razvoj na trgu.

Finančne spodbude za MSP-je:

- 1. Garancije za zavarovanje bančnih kreditov s subvencijo obrestne mere** omogočajo hitrejšo, lažje in cenejše pridobivanje bančnih kreditov za obratna sredstva ali večje investicije.
- 2. Mikrokrediti** so neposredni krediti manjših vrednosti, namenjeni za manjše investicije in obratna sredstva, s tem pa omogočijo nemoteno tekoče poslovanje.
- 3. Nepovratna sredstva za zagon podjetij** bodo tudi letos na voljo različnim start up podjetjem oziroma tistim, ki svojo podjetniško pot šele začinjajo.

Podpiramo tiste razvojno – poslovne projekte, ki zagotavljajo povečanje rasti podjetja, ustvarjajo višjo dodano vrednost na zaposlenega in zagotavljajo uspešen prenos razvojnih idej v tržno usmerjene uspešne podjeme.

