

Storitev označevanja nogometne tekme z informacijami o igralcih

Danilo Zimšek¹, Luka Banfi²

Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko,
Koroška cesta 46, 2000 Maribor

E-pošta: danilo.zimsek@um.si, luka.banfi@student.um.si

Football match annotation service with player information

Our paper describes an implementation of a football match annotation service with player information. Use case for the service is discussed with its overall implementation.

In this paper the colors of jerseys are analysed and limit color values for process of binarization are determined for 14 jerseys of randomly selected teams. Then the process of localisation and extraction of numbers is described. Special consideration is put on cases with same color of upper and lower part of jerseys. Other problematic frames are discussed, where there is colour similarity between jersey and background. Afterwards the algorithm for number recognition is described. The use of described approach is explained and annotation process described. The results of the process of localisation and recognition of numbers are also given.

In conclusion, future work and possible improvements of the process are discussed.

1 Uvod

Nogomet je že od nekdaj eden od najbolj popularnih športov predvsem v Evropi in Južni Ameriki, svojo popularnost pa pridobiva tudi v regijah, kjer je slabše poznan. Zanimanje za ta šport skozi zgodovino raste. Zaradi velikega zanimanja je za ponudnike telekomunikacijskih storitev ta šport posebej zanimiv za ponujanje novih storitev ogleda vsebine. Storitvi kot sta označevanje vsebine in skupinski ogled vsebine z videokonferenco lahko omogočita uporabnikom interaktivno spremljanje tekme v mikro socialnem omrežju, kjer lahko uporabniki v izbranem časovnem trenutku določenemu igralcu pridajo izbran čustveni simbol in tako na ta način izrazijo svoje mnenje o potezi igralca.

Takšna storitev, zasnovana na okolju IP televizije, ponudi uporabniku dodatno možnost socializacije in je iz tega vidika posebej pomembna za starejše in invalide, saj jim lahko uporaba tehnologij omogoči izkušnjo, ki je blizu izkušnji skupinskega spremljanja nogometne tekme, ki je dokaj pogost pojav.

Storitev skupinskega spremljanja nogometne tekme je zgrajena iz naslednjih modulov:

- storitev video konference;
- sistem povabljanja uporabnikov;

- storitev označevanja nogometne tekme.

Sistem povabljanja uporabniku omogoča, da drugega uporabnika povabi k ogledu nogometne tekme. Drug uporabnik lahko povabilo sprejme ali zavrne. V primeru, da vabilo sprejme vsaj en uporabnik, storitev video konference ob času začetka predvajanja nogometne tekme tvori konferenčno sobo v katero vključi uporabnike, ki so sprejeli vabilo. Video konferenca omogoča, da uporabniki v manjših delih zaslona spremljajo video portrete drugih udeležencev in se med seboj pogovarjajo. Hkrati se uporabnikom predvaja obogaten video tok nogometne tekme, ki ga tvori storitev označevanja nogometne tekme. Interakcija z vsebino je omogočena preko razširjene funkcionalnosti storitve video konference.

Eden glavnih gradnikov pri implementaciji zgoraj opisanih storitev je označevanje vsebine z informacijo o trenutno prikazanem igralcu. V nadaljevanju bomo opisali pri tem uporabljene pristope.

Zaznava in razpoznavna igralca nogometne tekme je pogosto obravnavana tematika med postopki označevanja video vsebin za razpršeno oddajo. Pristop v [2] uporabi v procesu zaznave pri empiričnem določanju mejnih vrednosti za izločanje dresov igralcev na podlagi barv metodo odvisno od svetlosti, kar povzroči, da se pojavljajo težave pri video okvirjih, ki so delno ali v celoti senčni. Za zmanjšanje vpliva osvetljenosti na algoritem zaznave dresa igralca, smo uporabili barvni model HSV. Ta se je po testiranju 5 primerov video okvirjev v barvnih prostorih RGB, BGR, HSV in YCbCr izkazal kot najprimernejši, saj smo empirično ocenili najboljše ločevanje med dresom in ozadjem okvirja.

Poleg odvisnosti od osvetljenosti, se v procesu zaznave pojavljajo omejitve vezane na podobnost med barvami zastopanimi na sceni in dresom igralca ter težave pri popačitvah predvsem zaradi umazanih dresov. Posebno težavo predstavljajo za do sedaj znane algoritme dresi, ki so v celoti v eni barvi, ki je enaka barvi nogavic. Te situacije so v zadnjem času pogostejše.

Naš cilj je z uporabljenim pristopom zmanjšati vpliv naštetih omejitev. V literaturi se za razpoznavo igralca uporabljajo različne kombinacije naslednjih tehnik: razpoznavna številke in razpoznavna obraza. V našem pristopu se omejimo na razpoznavo številke. Omejitev našega pristopa je delovanje na bližnjih posnetkih, saj lahko le tako zagotovimo dovolj natančno razpoznavo številke, kar pa



Slika 1: Vhodni video okvir.



Slika 2: Binariziran video okvir.

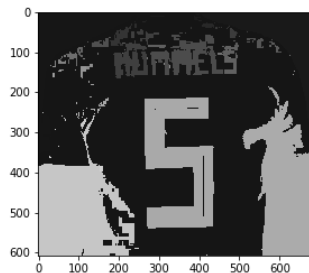
je smiselna omejitev, saj se bližnji posnetki navadno pojavljajo po pomembnem dogodku, katerega glavni akter je prikazan igralec.

V naslednjem poglavju opišemo postopek lokalizacije dresa, način določanja mejnih vrednosti barv za ta namen in zaznavo številke, pri čemer izpostavimo zahtevnejše primere. V tretjem poglavju obravnavamo postopek razpoznave številke pri čemer uporabimo deskriptor in mero sorodnosti z referenčno sliko številke. V četrtem poglavju opišemo označevanje vsebine. V petem predstavimo in komentiramo rezultate zaznave in razpoznave ter v zaključku podamo predloge za nadaljnje delo in možnosti za izboljšave.

2 Lokalizacija dresa

Prvi korak v postopku označevanja nogometne tekme z informacijami o igralcih predstavlja lokalizacija dresa, to je določanje pikslov v okvirju, ki pripadajo igralcu na igrišču. Predpostavljeno je bilo, da so določene barve povezane z določenimi nogometnimi ekipami. Za pravilno delovanje metode lokalizacije dresa, ta potrebuje podatek o barvnem razponu, ki predstavlja veliko večino barv zastopanih v dresih posamezne ekipe. Razpon vrednosti smo za vsako ekipo določili eksperimentalno v barvnem prostoru YUV. Proces vzpostavljanja primernih razponov vrednosti znotraj določenega barvnega modela prikazujejo enačbe 1, 4 in 3.

$$\Delta Y = \max_n Y_n - \min_n Y_n \quad (1)$$



Slika 3: Pridobljeni objekti izločenega območja video okvirja.

$$\Delta U = \max_n U_n - \min_n U_n \quad (2)$$

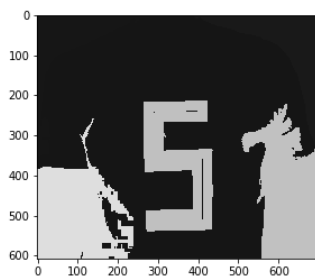
$$\Delta V = \max_n V_n - \min_n V_n \quad (3)$$

pri čemer je $n \in \{1, \dots, \text{št.okvirjev}\}$ in Δ označuje uporabljen barvni razpon, n pa indeks analiziranega okvirja. Skupno število analiziranih okvirjev za posamezno barvo dresa je bilo odvisno od razpršenosti trenutno pridobljenih podatkov. Večja kot je bila razpršenost, več okvirjev smo analizirali.

Proces lokalizacije dresa sestoji iz binarizacije slike, izločanja povezanih komponent in odstranjevanja madežev. Binarizacijo smo izvedli z uporabo empirično določenih mejnih vrednosti za vsako od komponent barvnega prostora. Slika 1 prikazuje primer vhodne, slika 2 pa primer binarizirane slike.

V nadaljevanju znotraj binarizirane poiščemo največje povezano področje bele barve. Izbrano področje uokvirimo in v nadaljevanju procesiramo samo to območje. Znotraj izločenega področja označimo vsa povezana področja (slika 3). Dobimo predstavitev večjih objektov, med katerimi je tudi številka.

Opazimo lahko, da so nekateri objekti zelo majhni, zato odstranimo objekte, ki so manjši od 5% celotne površine izbranega območja. S to omejitvijo izločimo napise imen igralcev na dresih in druge nepravilnosti v video okvirju. Dobljeno prikazuje slika 4. Ker je v nekaterih primerih barva majice enaka barvi hlač igralca, barvi nogavic ter ozadju (golovi vratnici), dobimo več objektov, iz katerih želimo izločiti samo številko. Do sedaj znani pristopi[1] za ta namen uporabljajo omejitve glede minimalne širine objekta ali omejitve glede razmerja stranic. Minimalna širina objekta v takem zahtevnejšem primeru ni zadosten pogoj, zato uporabimo omejitve glede razmerja stranic. Za večino pisav je razmerje med višino in širino števk med $2 : 1$ in $1,5 : 1$, zato smo uporabili ta pogoj. Poleg tega smo dodali pogoj o maksimalni višini objekta glede na višino izločenega območja in o maksimalni površini objekta glede na površino izločenega območja. Višina objekta ne sme presegati dveh tretjin višine območja. Površina objekta pa ne sme presegati ene polovice izbranega območja. Omejitve so bile določene



Slika 4: Problematični objekti video okvirja po odstranitvi majhnih objektov.

empirično glede na dobljene objekte 10 naključno izbranih testnih okvirjev. Če upoštevamo vse našete pogoje se število napačno izločenih objektov zmanjša na zamejerno število, tudi pri težavnih primerih. Do napačnih zaznav pride v manj kot 5% primerov.

3 Razpoznava številke

Eden od pristopov razpoznave števil vključujejo deskriptorje oblike. Eden od bolj primernih, zlasti zaradi neznanosti velikosti in usmerjenosti številke, je generični Fourierjev deskriptor (angl. Generic Fourier Descriptor), v nadaljevanju GFD [3].

Deskriptor obliko opiše tako, da se koordinate pikselov, ki pripadajo področju slike z obliko, transformirajo v polarne koordinate in prenesejo v pravokotno kartezijsko sliko. Nad podatki je uporabljena dvodimenzionalna Fourierjeva transformacija. Med dobljenima predstavitevama se izračuna Evklidova razdalja, ki predstavlja mero podobnosti med referenčnim objektom in kandidatom. V postopek GFD damo binarno sliko, ki ima en objekt v središču. Deskriptor izračuna stopnjo ujemanja vhodnega objekta z referenčnim objektom. V našem primeru so referenčni objekti slike števil pridobljene iz števil zapisanih v pisavi Arial v krepkem načinu. Algoritem izračuna s pomočjo GFD podobnost za vsak izločen objekt do vseh referenčnih objektov. Najvišjo mero podobnosti primerja z eksperimentalno določeno mejo. Če je ta višja od empirično določene meje, gre za ujemanje in objekt na sliki predstavlja obravnavano številko. V tem primeru steče proces označevanja razpoznavne na video tok.

Eksperimentalno določanje meje je potekalo na podlagi naključnega izbora dveh vzorcev video okvirjev vsake ekipe. Vzeli smo povprečno vrednost rezultata GFD iz vsake kategorije, izračunali povprečno vrednost nad njimi in zmanjšali za 20%. Opisan postopek prikazuje enačba 4.

$$m = \frac{\sum_{k=1}^{\text{št. kategorij}} f_k}{\text{št. kategorij}} \times 0.8$$

(4)

pri čemer m označuje mejno vrednost in f_k povprečje mejnih vrednosti video okvirjev posamezne kategorije.

4 Označevanje vsebine

Za potrebe opisane storitve je potrebno uporabniku posredovati informacijo o prepoznanem igralcu. Slednje smo izvedli z uporabo informacij iz elektronskega programskega vodiča (ang. electronic program guide) in z dodajanjem delno transparentne vsebine na video sled predvajanega video streama.

Pridobivanje informacij iz EPG je izvedeno s skripto, ki pregleda EPG v zapisu XML in na podlagi zapisa naslova vsebine izloči vse nogometne tekme in informacije o ekipah. Na podlagi teh informacij pridobi podatke o sestavi ekipe in preslikavami med imeni igralcev in številko dresa. Ti podatki se uporabijo v procesu označevanja video toka.

Označevanje video vsebine izvedemo z uporabo programske knjižnice ffmpeg [6]. Ta ponuja funkcionalnosti označevanja vsebin v realnem času, kar omogoča prikaz poljubne tudi delno transparentne vsebine na vsakem video okvirju. Po izvedenem označevanju se obogatena vsebina struja do uporabniškega klienta.

Storitev smo izvedli z uporabo programskega jezika Python in uporabo knjižnic scikit-image, opencv in scipy. Storitev smo poganjali na štiri jedrnem sodobnem procesorju intel. Pri tem je sistem za delovanje uporabljal dve jedri, saj vzporedno tečeta dva procesa, za vsako od dveh ekip nogometne takme za isti vhodni video okvir. Ob predpostavki, da ima video tok 25 sličic na sekundo, smo uspeli storitev na navedeni strojni poganjati za vsak deseti video okvir, torej je celoten postopek potreboval manj kot 0.4 sekunde.

5 Rezultati

Za evalvacijo opisane metode je bilo najprej potrebno zgraditi bazo. Baza video okvirjev nogometnih tekem sestoji iz 14 kategorij. V vsaki kategoriji je od 10 do 15 video okvirjev iz bližnjih posnetkov z igralci istih ekip. Zgradbo baze in uspešnost uporabljenega pristopa prikazuje tabela 1.

Iz dobljenih rezultatov je razvidno, da je omenjen pristop z uporabo barvnega modela YUV in določanje mejnih vrednosti za binarizacijo iz izbranega modela uspešnejši od pristopa opisanega v [1]. Tak rezultat smo pripisali predvsem določanju mejnih vrednosti za vsako ekipo posebej v barvnem prostoru YUV. Ker pa rezultati niso testirani na isti bazi in ker je pridobljena baza relativno majhna, je težko z gotovostjo trditi, da je to edini faktor, ki vpliva na izboljšanje.

K nepravilnim razpoznavam pa so prispevali predvsem okvirji, pri katerih so igralci z dresi in nogavicami v isti barvi kot ozadje ter okvirji slabe kakovosti.

Trenutni pristop je mogoče uporabiti za implementacijo opisanih storitev, a se pri tem pojavlja težava pri razpoznavi igralcev s številko dresa sestavljeno iz dveh števil. To težavo rešujemo na način, da na podlagi koordinat skrajnih vrednosti mejnega polja številke in njegove oddaljenosti od druge zaznane številke določimo, ali

Tabela 1: Uspešnost uporabljenega pristopa po kategorijah video okvirjev.

Razred	št. okvirjev s števili	št. pravih razpoznav	uspešnost [%]
1.	15	13	87
2.	12	10	83
3.	11	10	91
4.	14	13	93
5.	13	12	92
6.	10	8	80
7.	12	10	83
8.	13	12	92
9.	11	9	82
10.	15	13	87
11.	12	10	83
12.	10	9	90
13.	12	11	92
14.	11	10	91
skupaj			88

Literatura

- [1] Frejlichowski, Dariusz: Identification of Football Players based on Generic Fourier Descriptor Applied for the Recognition of Numbers, Image Processing & Communications. 2016
- [2] Frejlichowski, Dariusz: A Method for Data Extraction from Video Sequences for Automatic Identification of Football Players Based on Their Numbers, Image Analysis and Processing – ICIAP, 356–364, 2011
- [3] Dengsheng Zhang, Guojun Lu: Generic Fourier descriptor for shape-based image retrieval, IEEE International Conference on Multimedia and Expo, 425–428, 2002
- [4] Yin, Wu, Zheng Ncfm: Accurate handwritten digits recognition using Convolutional Neural Networks, International Joint Conference on Neural Networks (IJCNN), 525–531, 2016
- [5] Dengsheng Zhang, Guojun Lu: Enhanced Generic Fourier Descriptors for object-based image retrieval, IEEE International Conference on Acoustics, Speech, and Signal Processing, 3668–3671, 2002
- [6] Ffmpeg. <https://www.ffmpeg.org/>. Dostopno 13.07.2018.

Števkata sestavljata število. Če je njuna medsebojna oddaljenost manjša od polovice širine števkata, predvidevamo, da števkata pripadata istemu številu.

6 Zaključek

V sklopu nadaljnega dela načrtujemo vključitev opisane storitve označevanja vsebin v storitev videokonference. Za doseg tega cilja je potrebno avtomatizirati procese znotraj celotne storitve, ki trenutno potrebujejo za uspešno delovanje vnos informacij s strani človeka. Tak je predvsem proces pridobivanja podatkov o igralcih vključenih v nogometno tekmo. V nadaljevanju bomo procese pridobivanja poskušali izvesti na enega od naslednjih načinov:

- razpoznavanje imen igralcev in števil dresov iz začetnih zaslonov prenosa tekme;
- pridobivanje podatkov iz dostopnih aplikacijskih vmesnikov.

Nadalje bomo omogočili interaktivno rokovanje z vsebino s strani uporabnika na način, da bo uporabnik s funkcijskimi tipkami daljinskega upravljalnika imel možnost dodajanja čustvenih simbolov ob času razpoznavanja igralca in določeno časovno obdobje po koncu bližnjega posnetka.

Želeli bi doseči tudi izboljšanje natančnosti razpoznavanja števil. Za ta namen bi lahko generični Fourier deskriptor nadomestili z drugim [5] ali bi za ta namen uporabili pristop z nevronskimi mrežami [4]. Celoten postopek je kritičen predvsem zaradi zahteve po izvajanju tekom predvajanja vsebine, zaradi česar bomo pred menjavo pristopa preučili časovno zahtevnost posameznih alternativnih deskriptorjev in pristopov z nevronskimi mrežami.