

*Vizualne tehnike analize človeškega gibanja in njihova
uporaba v biometriji*

Jure Kovač

DOKTORSKA DISERTACIJA

PREDANA

FAKULTETI ZA RAČUNALNIŠTVO IN INFORMATIKO

KOT DEL IZPOLNJEVANJA POGOJEV ZA PRIDOBITEV NAZIVA

DOKTOR ZNANOSTI

S PODROČJA

RAČUNALNIŠTVA IN INFORMATIKE



Ljubljana, 2014

*Vizualne tehnike analize človeškega gibanja in njihova
uporaba v biometriji*

Jure Kovač

DOKTORSKA DISERTACIJA

PREDANA

FAKULTETI ZA RAČUNALNIŠTVO IN INFORMATIKO

KOT DEL IZPOLNJEVANJA POGOJEV ZA PRIDOBITEV NAZIVA

DOKTOR ZNANOSTI

S PODROČJA

RAČUNALNIŠTVA IN INFORMATIKE



Ljubljana, 2014

IZJAVA

Izjavljam, da sem avtor dela in da to ne vsebuje materiala, ki bi ga kdorkoli predhodno že objavil ali oddal v obravnavo za pridobitev naziva na univerzi ali na drugem visokošolskem zavodu, razen v primerih, kjer so navedeni viri.

— Jure Kovač —
september 2014

ODDAJO SO ODOBRILI

dr. Peter Peer

docent za računalništvo in informatiko

MENTOR IN ČLAN OCENJEVALNE KOMISIJE

dr. Franc Solina

profesor za računalništvo in informatiko

PRESEDNIK OCENJEVALNE KOMISIJE

dr. Božidar Potočnik

izredni profesor za računalništvo in informatiko

ZUNANJI ČLAN OCENJEVALNE KOMISIJE

Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko

PREDHODNA OBJAVA

Izjavljam, da so bili rezultati obravnavane raziskave predhodno objavljeni/sprejeti za objavo v recenzirani reviji ali javno predstavljeni v naslednjih primerih:

- [1] J. Kovač and P. Peer. Human Skeleton Model Based Dynamic Features for Walking Speed Invariant Gait Recognition. *Mathematical Problems in Engineering*, 2014(1):1–15, 2014. doi: [10.1155/2014/484320](https://doi.org/10.1155/2014/484320)
- [2] J. Kovač and P. Peer. Transformation Based Walking Speed Normalization for Gait Recognition. *KSI Transactions on Internet and Information Systems*, 7(x11):2690–2701, 2013. doi: [10.3837/tiis.2013.11.008](https://doi.org/10.3837/tiis.2013.11.008)

Potrjujem, da sem pridobil pisna dovoljenja vseh lastnikov avtorskih pravic, ki mi dovoljujejo vključitev zgoraj navedenega materiala v pričujočo disertacijo. Potrjujem, da zgoraj navedeni material opisuje rezultate raziskav, izvedenih v času mojega podiplomskega študija na Univerzi v Ljubljani.

POVZETEK

Ljudje smo sposobni prepoznati druge ljudi, ki jih dobro poznamo, samo po tem, kako se ti premikajo. Ta sposobnost predstavlja osnovno motivacijo za uporabo načina gibanja kot sredstva za biometrično identifikacijo. Takšno biometriko lahko zajamemo na daljavo, na javnih mestih, brez posameznikovega sodelovanja, zavedanja ali celo privoljenja. Kljub temu, da današnji pristopi dajejo vzpodbudne rezultate, smo še vedno daleč od učinkovite rabe v vsakdanjih aplikacijah. V splošnem je treba upoštevati različne omejitve, da lahko metode zaobidejo številne dejavnike vpliva, kot so spremembe hitrosti hoje, pogleda, obleke, obutve in nošenja predmetov, ki imajo negativen vpliv na učinkovitost razpoznavanja.

V predloženem delu predlagamo sistem za razpoznavanje ljudi na osnovi modela skeleta človeškega telesa, pri čemer se osredotočamo predvsem na modeliranje dinamike gibanja in izločanje vpliva posameznikovega videza na razpoznavanje. Predlagani metodi za klasifikacijo s fuzijo značilk in klasifikacijo na osnovi okvirjev demonstrirata primerljivost razvitega sistema z drugimi sorodnimi metodami, ki sicer temeljijo na videzu, celo v takšnih okoljih, kjer imajo slednje prednost. Dodatno naslovimo problem spremembe hitrosti hoje in s predlagamo transformacijo prostora značilk omilimo njegov negativen vpliv na učinkovitost razpoznavanja. Ta daje v kombinaciji s klasifikacijo na osnovi okvirjev celo boljše rezultate razpoznavanja kot druge – najboljše (angl. *state of the art*) – metode.

S podrobno evalvacijo prikažemo učinkovitost in robustnost delovanja predlaganih metod tako pri nespremenjenih pogojih kot pri spremembah hitrosti hoje ter pri prisotnosti različnih ovir značilnih za širše področje računalniškega vida (npr. šum, nizka ločljivost ipd.), ki v splošnem ovirajo učinkovitost delovanja aplikacij s področja računalniškega vida.

Ključne besede: model skeleta, neodvisnost od hitrosti hoje, razpoznavanje ljudi, računalniški vid, biometrija

ABSTRACT

Humans are able to recognize small number of people they know well by the way they walk. This ability represents basic motivation for using human gait as the means for biometric identification. Such biometric can be captured at public places from a distance without subject's collaboration, awareness and even consent. Although current approaches give encouraging results, we are still far from effective use in real-life applications. In general, methods set various constraints to circumvent the influence of covariate factors like changes of walking speed, view, clothing, footwear, object carrying, that have negative impact on recognition performance.

In this thesis we propose a skeleton model based gait recognition system focusing on modeling gait dynamics and eliminating the influence of subject appearance on recognition. We propose feature fusion scheme for classification and frame based classification which both demonstrate how such system is comparable to other state of the art appearance based methods even in the environments where these have clear advantage. Furthermore, we address the problem of walking speed variation and propose feature space transformation technique that mitigates its negative influence on recognition performance. Together with frame based classification, the method achieves state of the art classification results and outperforms other similar methods.

With the extensive evaluation we demonstrate state of the art performance and robustness of proposed methods under unchanged conditions, varying walking speeds and even under presence of different computer vision related obstacles (e.g. noise, low resolution etc.) that obstruct the performance of real-life computer vision applications in general.

Key words: skeleton model, walking speed invariance, gait recognition, computer vision, biometrics

ZAHVALA

Iskreno se zahvaljujem vsem, ki so mi pomagali pri nastajanju doktorske disertacije. Najprej se zahvaljujem mentorju, doc. dr. Petru Peeru, in mag. Mateju Meži, ki sta mi v času študija stala ob strani in mi pomagala s številnimi idejami ter nasveti. Zahvaljujem se tudi članom Laboratorija za računalniški vid, in sicer za konstruktivne diskusije in nasvete. Prav tako se zahvaljujem članoma komisije, prof. dr. Francu Solini in prof. dr. Božidarju Potočniku, za koristne pripombe in nasvete.

Na koncu se zahvaljujem tudi družini, partnerki Barbari ter hčerama Neji in Brini, ki so mi ves čas študija stali ob strani, me podpirali, bodrili in predvsem navdihovali z inspiracijo, potrebno za vztrajanje na načrtani poti. Navsezadnje je ravno njihova podpora tista, ki je omogočila uspešen zaključek naloge.

— Jure Kovač, Ljubljana, september 2014.

KAZALO

<i>Povzetek</i>	<i>i</i>
<i>Abstract</i>	<i>iii</i>
<i>Zahvala</i>	<i>v</i>
1 <i>Uvod</i>	1
1.1 Motivacija	2
1.2 Hipoteza in cilji disertacije	4
1.3 Pregled vsebine	5
2 <i>Pregled področja</i>	7
2.1 Uvod	8
2.2 Pristopi brez modela	8
2.2.1 Osnovni pristop in slika povprečne silhuete	9
2.2.2 Verjetnostni prostorsko-časovni model	10
2.3 Pristopi na osnovi modela	13
2.3.1 Skeletni model človeškega telesa	13
2.4 Komponentni pristopi	16
2.5 Biološki pristopi	17
2.6 Faktorji vpliva	18
2.6.1 Sprememba hitrosti hoje	19
2.6.2 Sprememba pogleda	20
2.6.3 Drugi faktorji	20
2.7 Podatkovne zbirke za evalvacijo	21

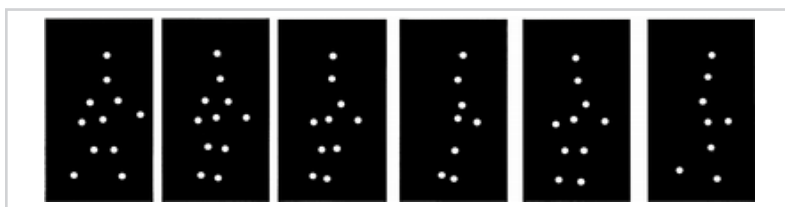
3	<i>Predlagana metoda</i>	25
3.1	Arhitektura predlagane rešitve	26
3.2	Segmentacija gibanja	26
3.3	Segmentacija skeleta človeškega telesa	27
3.4	Iskanje gibalnih značilnk	28
3.4.1	Parametri gibanja na osnovi slikovnih momentov	30
3.4.2	Detekcija ciklov gibanja	31
3.4.3	Določanje faze dvojne opore	32
3.4.4	Detekcija križanja nog	33
3.4.5	Poobdelava signalov	34
3.5	Klasifikacija	38
3.5.1	Fazna poravnava značilnk v ciklih	38
3.5.2	Opisnik gibanja	39
3.5.3	Celostni pristop klasifikacije	41
3.5.4	Klasifikacija na osnovi fuzije značilnk	41
3.5.5	Klasifikacija na osnovi okvirjev	43
3.5.6	Dekompozicija signalov z valčki	46
4	<i>Sprememba hitrosti hoje</i>	49
4.1	Uvod	50
4.1.1	Časovna normalizacija cikla gibanja	50
4.2	Normalizacija hitrosti hoje	50
4.3	Transformacija prostora značilnk	53
5	<i>Rezultati in analiza</i>	57
5.1	Evalvacija normalizacije hitrosti hoje z geometrijskimi transformacijami	58
5.2	Evalvacija predlagane metode na osnovi skeleta	60
5.2.1	Rezultati eksperimentov brez transformacije prostora značilnk	61
5.2.2	Rezultati eksperimentov s transformacijo prostora značilnk . .	62
5.2.3	Povzetek rezultatov medhitrostnih eksperimentov	63
5.2.4	Primerjava s sorodnimi deli	69
5.3	Vpliv velikosti galerije na učinkovitost razpoznavanja	71
5.4	Vpliv valčkov na učinkovitost razpoznavanja	72
5.5	Vpliv faznega zamika ciklov na učinkovitost razpoznavanja	74

5.6	Vpliv šuma, ločljivosti in netipičnih meritev	76
5.6.1	Vpliv ločljivosti silhuet na učinkovitost razpoznavanja	77
5.6.2	Vpliv šuma v silhuetah na učinkovitost razpoznavanja	79
5.6.3	Vpliv šuma v skeletu na učinkovitost razpoznavanja	80
6	<i>Zaključek</i>	85
6.1	Prispevki k znanosti	87
6.2	Nadaljnje delo	87
	<i>Literatura</i>	89

Uvod

1.1 Motivacija

Psihološke študije so pokazale, da imamo ljudje majhno, a pomembno sposobnost prepoznavanja ljudi, ki jih dobro poznamo, s pomočjo opazovanja njihovega načina hoje in značilnosti gibanja. Ta sposobnost je vzpodbudila raziskave v smeri uporabe značilnosti hoje kot sredstva za biometrično identifikacijo. Zgodnje študije [1] s pomočjo točkovnih prikazov (angl. *Point Light Displays – PLD*), ki omogočajo izolirano preučevanje gibanja brez motnje drugih kontekstov, so potrdile, da lahko ljudje v manj kot sekundi ugotovimo, ali gre za človeško gibanje (slika 1.1). Še več, nadaljnje so ugotovili, da so ljudje zmožni opravljati tudi klasifikacijo spola in celo razpoznavanje [2].



Slika 1.1

Točkovni prikaz (PLD)
človekove hoje

Običajne biometrike, ki temeljijo na prstnih odtisih, obrazih, šarenici ipd., imajo dve očitni pomanjkljivosti:

1. Slabo se obnesejo pri slikah slabše kakovosti (npr. pri nizki resoluciji).
2. Potrebujemo aktivno sodelovanje uporabnika.

Nasprotno lahko človeško hojo zajamemo s preprosto opremo, kot so običajne varnostne kamere brez posameznikovega zavedanja ali celo privoljenja. Glavni pomanjkljivosti te biometrike pa sta nepoznavanje možnosti razločevanja med posamezniki, torej nivo unikatnosti človekove hoje, ter številni faktorji vpliva, ki spreminjajo značilnosti posameznikove hoje. Faktorji vpliva so lahko zunanji, npr. spremembe pogleda, smeri in hitrosti gibanja, pogojev osvetlitve, vremena, obleke, obutve in terena, ali notranji, npr. spremembe hoje zaradi bolezni, poškodb, staranja, nosečnosti ipd. Probleme lahko povzročajo tudi nezanesljive meritve, prekrivanja, uporaba neinvazivnih metod zajemanja (brez senzorjev ali dodatnih oznak – markerjev). Vsi ti faktorji negativno vplivajo na rezultate razpoznavanja v realnem okolju, ki so trenutno še vedno preslabi za učinkovito uporabo v biometriji.

Področje, kjer lahko aplikativno uporabimo predstavljeno tematiko, je zelo široko in nudi perspektivne možnosti za razvoj tako s stališča raziskovalnega kot gospodarskega vidika. Opisano tematiko srečamo na področjih varnosti in nadzora [3], kjer gre običajno za sledenje in nadzor oseb, v medicini za diagnozo in nadzor zdravljenja bolezni, povezanih z gibanjem, kot so Parkinsonova bolezen, multipla skleroza, kap in detekcija padcev [4], v športu za izboljšanje atletskih zmogljivosti športnika, evalvacijo in spremljanje treninga [5], na področju zajemanja gibanja (angl. *motion capture*) [6], pri nadzoru prometa [7] itd.

Čeprav številne današnje metode za razpoznavanje ljudi po načinu hoje ponujajo impresivne rezultate v kontroliranem laboratorijskem okolju [8–14], je uporaba tovrstne biometrike v vsakdanjih aplikacijah še vedno omejena. Predvsem zaradi faktorjev vpliva, ki spreminjajo posameznikovo gibanje in s tem otežujejo nalogo. Kljub temu obstajajo primeri učinkovite rabe v vsakdanjih aplikacijah. Avtorji v [15, 16] so demonstrirali, kako lahko s pomočjo nadzora parametrov gibanja varovancev doma za ostarele zaznamo anomalije v gibanju varovancev [15] in celo razločujemo varovance od obiskovalcev [16]. Takšen nadzor lahko uporabimo za ocenjevanje tveganja padcev, zaznavanje zdravstvenih težav in nadzor rehabilitacije pacientov po kirurških posegih. Čeprav avtorji direktno ne obravnavajo faktorjev vpliva, predlagajo učinkovit način eliminacije tistih vzorcev hoje, ki ne ustrezajo omejitvam samih metod za analizo.

Pogosto pa je v vsakdanjih okoliščinah na voljo le omejena količina posameznikovih vzorcev hoje (npr. iz varnostne kamere), ki so lahko poleg tega še kratki, vsebujejo le nekaj korakov in so podvrženi raznim prej omenjenim faktorjem vpliva. V takšnih okoliščinah je treba faktorje vpliva posebej obravnavati in zmanjšati njihov vpliv, da dobimo zadostno število vzorcev za uspešno analizo. V našem delu se osredotočamo predvsem na vpliv spremembe hitrosti hoje, ki predstavlja enega najbolj problematičnih faktorjev vpliva z negativnim učinkom na rezultate razpoznavanja, je skoraj vedno prisoten in zato potrebuje posebno obravnavo. V literaturi obstajajo številni pristopi k reševanju tega problema, najbolj izstopajoče povzemamo v nadaljevanju. V nasprotju z našim delom pa nobeden od njih ne sodi v skupino pristopov na osnovi modela, pri razpoznavanju si dodatno pomagajo s posameznikovim videzom, kar ima negativen učinek predvsem pri prisotnosti faktorjev vpliva, hkrati pa s tem zapostavljajo dinamično gibanje, ki igra ključno vlogo pri razpoznavanju ljudi po načinu hoje in je zato tudi glavni fokus našega dela.

1.2 Hipoteza in cilji disertacije

Glavni namen doktorske disertacije je proučiti trenutno stanje na področju razpoznavanja ljudi po načinu hoje, identificirati prednosti in slabosti obstoječih metod ter predlagati ustrezne izboljšave in rešitve najbolj perečih problemov področja. Kljub temu, da je področje relativno mlado, je literatura precej bogata. V doktorski nalogi zato izpostavljam le tiste primere literature, ki izstopajo oz. so tako ali drugače relevantni za naše konkretne izzive.

Poglaviten cilj naših raziskav je izboljšava najboljših (angl. *state of the art*) metod področja, primarno z vidika zmanjšanja negativnega vpliva faktorjev, ki spreminjajo gibanje ter s tem negativno vplivajo na učinkovitost razpoznavanja. Pri tem posebno pozornost posvečamo značilkam, ki so že v osnovi manj občutljive na te dejavnike in postavljamo naslednjo hipotezo:

Kljub temu, da imajo t. i. dinamične značilke, ki opisujejo zgolj dinamiko gibanja, manj razločevalne moči kot statične značilke, še vedno vsebujejo dovolj razločevalne informacije za doseganje primerljivih rezultatov razpoznavanja z metodami, ki temeljijo na videzu in so hkrati bolj odporne ter bolj prilagodljive na faktorje vpliva, ki so pogosto prisotni v scenarijih razpoznavanja.

S pomočjo razvitih metod bomo zasnovali in izdelali prototip modela arhitekture človeškega vida za področje razpoznavanja značilnosti gibanja, ki bo omogočal avtomatski zajem parametrov gibanja iz videoposnetkov in vrednotenje teh v smislu biometrične razpoznavne. Sistem bomo s pomočjo standardnih javno dostopnih zbirk podatkov, ki odražajo opisane dejavnike vpliva, primerjali z obstoječimi podobnimi rešitvami in ovrednotili posledice, ki ga imajo ti dejavniki na razpoznavanje ter uporabo v biometriji.

Pričakujemo naslednja neposredna prispevka k znanosti:

1. Razvoj metode za razpoznavanje gibanja, ki bo omogočala zajem značilk, ki so že v osnovi manj občutljive na dejavnike vpliva.
2. Razvoj prijemov za dodatno blaženje negativnih učinkov faktorjev vpliva s posebno pozornostjo na spremembi hitrosti hoje.

1.3 Pregled vsebine

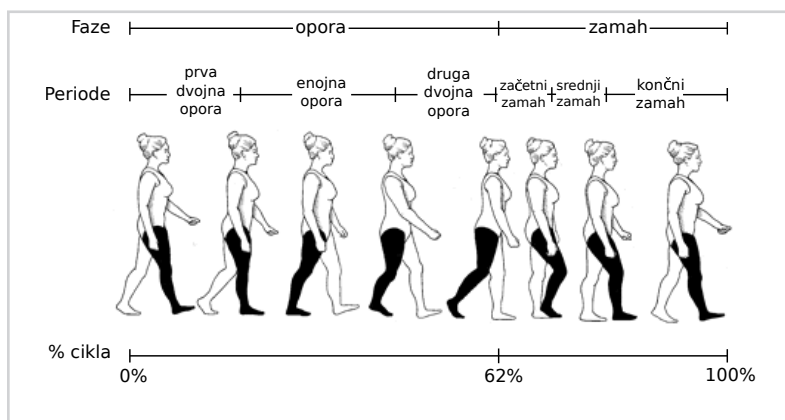
Vsebina v naslednjih poglavjih je organizirana na naslednji način: v poglavju 2 je podan pregled področja razpoznavanja ljudi po načinu hoje z opisi metod najvidnejših predstavnikov, s pregledom izzivov področja ter z opisom evalvacijskih zbirk. V poglavju 3 je podan natančen opis vseh segmentov predlagane metode na osnovi modela skeleta človeškega telesa. V poglavju 4 je podrobneje predstavljen faktor vpliva spremembe hitrosti hoje na gibanje skupaj s predlaganimi rešitvami, ki omilijo njegov negativni učinek na rezultate razpoznavanja. V poglavju 5 navajamo rezultate evalvacije predlaganih postopkov skupaj z diskusijo in analizo vpliva faktorjev, kot sta šum in ločljivost videoposnetka na opisane postopke. V poglavju 6 povzemamo sklepne ugotovitve, prispevke k znanosti in nekatere smernice za nadaljnje delo.



Pregled područja

2.1 Uvod

Značilnost gibanja (angl. *gait*) definiramo kot koordinirano ciklično kombinacijo posameznikovih kretenj, ki tvorijo gibanje (npr. hoja, tek ipd.) [17]. Osnovni element, na katerem temelji večina metod, je gibalni cikel. Cikel sestavljata faza opore (angl. *stance*) in faza zamaha (angl. *swing*). Gre torej za pot, ki jo opravi ena noga, da se spet vrne v začetni položaj (slika 2.1). Dodatno definiramo hojo kot gibanje, pri katerem je ena noga vedno v stiku s podlago, za razliko npr. od teka, kjer to ne velja.



Slika 2.1

Faze gibalnega cikla normalne hoje odraslega človeka [18].

Natančnejši pregled področja in klasifikacijo pristopov najdemo v [19–23], v tem delu pa se osredotočamo predvsem na pristope, ki so kakorkoli neposredno povezani z našim delom. V prvi meri lahko pristope razdelimo v dve glavni skupini, podrobneje razdelani v nadaljevanju.

2.2 Pristopi brez modela

Pristopi brez modela (angl. *Model-free*) [11–13, 24] pridobijo parametre gibanja z izvajanjem meritev direktno na 2D-slikah, brez uporabe nekega specifičnega modela človeškega telesa ali njegovega gibanja. Korespondenco značilk v zaporednih slikah dosežejo s predvidevanjem hitrosti, oblike, teksture in barve. Večinoma uporabljajo geometrične predstavitve, kot so npr. silhuete (angl. *silhouettes*), optični tok (angl. *optical flow*), trajektorije sklepov, zgodovina gibanja ipd. Metode se ne zanašajo na

dinamiko gibanja, ampak predvsem na meritve posameznika med gibanjem, s čimer za namene razpoznavanje neposredno ali posredno upoštevajo tudi posameznikov videz. Metode te skupine so zato manj občutljive na faktorje vpliva, ki spreminjajo dinamiko gibanja (npr. staranje, bolezn, sprememba hitrosti hoje ipd.), a nekoliko slabše pri faktorjih vpliva, ki spreminjajo videz (npr. sprememba obleke, debelosti, pričeske ipd.), pogled in smer gibanja.

2.2.1 Osnovni pristop in slika povprečne silhuete

Osnovni pristop, predstavljen v [12], temelji na neposredni primerjavi silhuet videosekvence testnega primera s silhuetami videosekvenc primerov iz galerije. Mera podobnosti za primerjavo je definirana kot:

$$\text{frsim}(\mathbf{s}_p(i), \mathbf{s}_g(j)) = \frac{\text{num}(\mathbf{s}_p(i) \cap \mathbf{s}_g(j))}{\text{num}(\mathbf{s}_p(i) \cup \mathbf{s}_g(j))}, \quad (2.1)$$

kjer $\text{frsim}(\mathbf{s}_p(i), \mathbf{s}_g(j))$ določa podobnost med dvema silhuetama $\mathbf{s}_p(i)$ in $\mathbf{s}_g(j)$, $\mathbf{s}_p(i)$ pripada okvirjem testne videosekvence, $\mathbf{s}_g(j)$ okvirjem videosekvence primera iz galerije, funkcija num pa vrne število neničelnih elementov v binarni sliki. Gre torej za razmerje med presekom in unijo dveh silhuet, pri čemer silhueto predstavljajo neničelni elementi binarne slike.

Podobnost med dvema sekvencama določimo s pomočjo korelacije:

$$\text{corr}(\mathbf{s}_{p_k}, \mathbf{s}_g)(l) = \sum_{j=1}^N \text{frsim}(\mathbf{s}_p(k+j), \mathbf{s}_g(l+j)), \quad (2.2)$$

kjer postopamo tako, da celotno testno sekvenco razbijemo na k podsekvenc \mathbf{s}_{p_k} z dolžino N , vsako od teh podsekvenc pa nato koreliramo s celotno dolžino sekvence primera iz galerije \mathbf{s}_g . $\text{corr}(\mathbf{s}_{p_k}, \mathbf{s}_g)$ predstavlja torej vrednosti korelacije na vseh okvirjih l sekvence iz galerije \mathbf{s}_g s posamezno testno podsekvenco \mathbf{s}_{p_k} . Takšna razdelitev na podsekvence omili vpliv segmentacijskih napak, ki so včasih prisotne v posameznih ali več zaporednih okvirjih videosekvenc. Končno vrednost podobnosti predstavlja mediana korelacij vseh podsekvenc \mathbf{s}_{p_k} s sekvenco galerije \mathbf{s}_g :

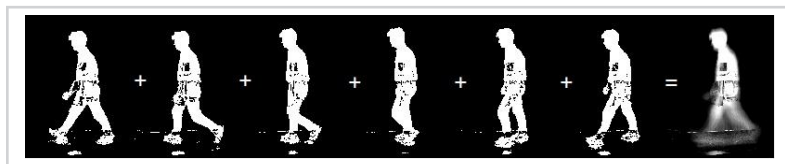
$$\text{sim} = (\mathbf{s}_p, \mathbf{s}_g) = \text{median}_k \left(\max_l \text{corr}(\mathbf{s}_{p_k}, \mathbf{s}_g)(l) \right). \quad (2.3)$$

V nespremenjenih pogojih dosega metoda presenetljivo dobre rezultate, zaradi česar se uporablja kot osnova (angl. *baseline*) za ocenjevanje novih metod.

V nekoliko naprednejšem pristopu [11] avtorji uvedejo sliko povprečne silhuete (angl. *averaged silhouette*) (slika 2.2), ki predstavlja vsoto obrisov enega gibalnega cikla, model gibanja pa predstavljajo povprečni obrisi vseh ciklov gibanja. Za takšen pristop je videosekvenco treba razdeliti na cikle gibanja, kar dosežemo s pomočjo iskanja periode gibanja. Periodo ocenimo s pomočjo števila neničelnih elementov N_f v posameznih okvirjih t videosekvence. Teh je največ, ko sta nogi silhuete najbolj narazen, in najmanj, ko sta nogi skupaj. Poiščemo torej lokalne maksimume $N_f(t)$ in s pomočjo razdalj med njimi ocenimo periodo N_{gait} .

V nadaljevanju sekvenco \mathbf{s} razbijemo na podsekvence $\mathbf{s}_{p_k} = \{\mathbf{s}(k), \dots, \mathbf{s}(k + N_{gait})\}$, ki predstavljajo podsekvence silhuet posameznih ciklov celotne sekvence, in za vsako od teh podsekvenc izračunamo povprečno silhueto:

$$\mathbf{s}_a(i) = \frac{1}{N_{gait}} \sum_{k=iN_{gait}}^{(i+1)N_{gait}-1} \mathbf{s}(k). \quad (2.4)$$



Slika 2.2

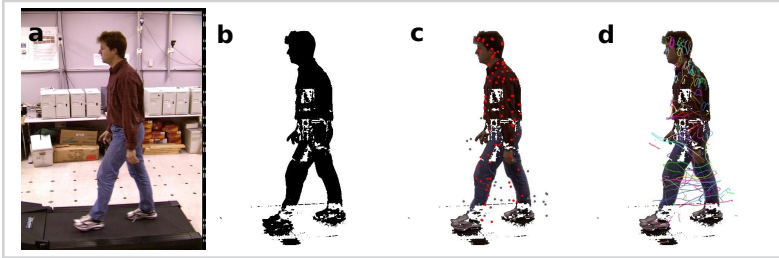
Slika povprečne silhuete.

Pristop je neobčutljiv na razlike v dolžini ciklov in od izbire začetne faze gibalnega cikla, a po drugi strani zahteva detekcijo ciklov z iskanjem periode, ki je občutljiva na šum in napake v silhuetah.

2.2.2 Verjetnostni prostorsko-časovni model

Metoda temelji na trajektorijah sledi naključno izbranih točk [14], s pomočjo katerih zgradimo model gibanja človeškega telesa. Kljub temu, da metoda zgradi model gibanja, ne sodi v skupino pristopov na osnovi modela, saj je v tem primeru model rezultat metode. Gre torej za modeliranje značilk, pridobljenih z 2D-slik, kot je značilno za pristope brez modela, in ne gre za model gibanja, iz katerega pridobivamo značilke.

Najprej s pomočjo standardnih algoritmov za segmentacijo gibanja [25] izločimo ozadje. Rezultat tega koraka so maske ospredja za vsak okvir videosekvenca, kar v primeru ljudi pomeni obliko človeške silhete (slika 2.3 b). Nato čez celotno video-



Slika 2.3

Slika prikazuje korake metode na osnovi trajektorij sledenih točk: a. vhodna slika, b. izločena maska ospredja, c. naključno izbrane sledene točke, d. trajektorije, ki jih tvorijo sledene točke v času.

sekvenco sledimo določenemu številu naključno izbranih točk na silhuetah (slika 2.3 c). Te točke v času tvorijo trajektorije, ki jih opišemo s t. i. prostorsko-časovnimi krivuljami (slika 2.3 d):

$$c(t) = (c_x(t), c_y(t)). \quad (2.5)$$

Ker je signal trajektorij, ki ga tvorijo točke, periodičen, lahko s pomočjo avtokorelacije in glasovanja poiščemo periodo (slika 2.4 a, b). Najdena perioda ustreza dolžini gibalnega cikla. Vsako od krivulj nato razrežemo na dolžine cikla, jo fazno poravnamo z ostalimi krivuljami in interpoliramo po časovni komponenti na neko fiksno dolžino N . Tako pridobimo množico cikličnih poravnanih krivulj iste dolžine, ki predstavljajo osnovo za izgradnjo modela gibanja (slika 2.4 c). Krivuljo opišemo s centroidom (enačba 2.6) in vektorjem (enačba 2.7) odstopanj od centroida, ki opisuje obliko krivulje:

$$\mathbf{o} = (o_x, o_y) = \frac{1}{N} \sum_{i=1}^N (\mathbf{x}_i), \quad (2.6)$$

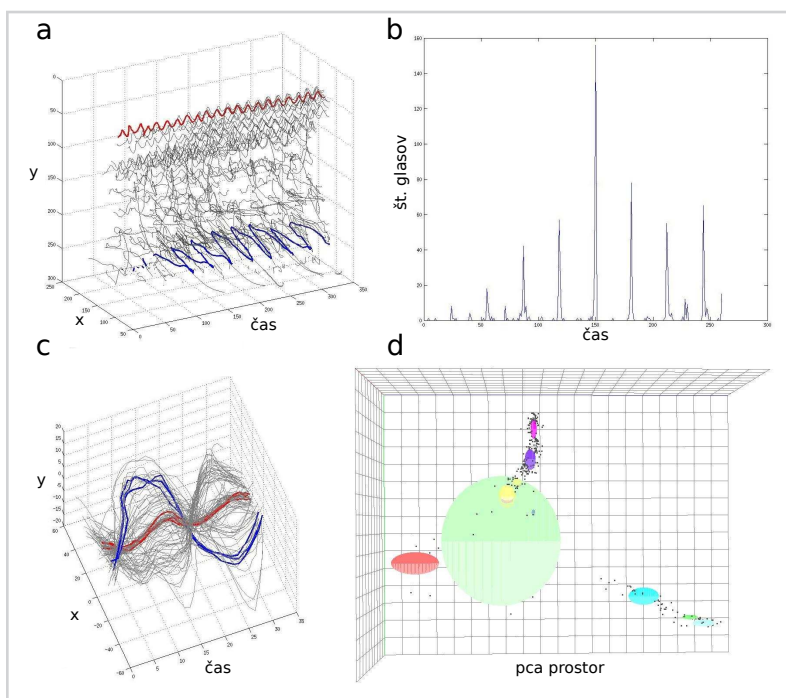
$$\mathbf{x} = \{x_1 - o_x, y_1 - o_y, \dots, x_N - o_x, y_N - o_y\}. \quad (2.7)$$

Zaradi velikega števila tako nastalih krivulj in še večjega števila parametrov zmanjšamo dimenzionalnost z uporabo algoritma analize osnovnih komponent (angl. *Principal Component Analysis – PCA*). PCA preslika vsako od cikličnih krivulj, pridobljenih v modulu iskanja značilik, v točko v nižjedimenzionalnem prostoru (slika 2.4 d). Vse te točke v novem prostoru predstavljajo porazdelitev, ki jo nadaljnje opišemo z mešanico

Gaussov (angl. *Gaussian mixture model – GMM*). Vsaka krivulja z neko verjetnostjo, ki jo lahko izračunamo z uporabo Bayesove formule, pripada določeni mešanici Gaussov, postopek EM (angl. *Expectation maximization*) pa poišče takšno mešanico, ki maksimizira to verjetnost za vse krivulje istega modela. Vhod v EM-algoritem predstavljajo torej šestdimenzionalni opisniki (angl. *descriptors*) krivulj, sestavljeni iz centroida krivulje in štirih koordinat točke v PCA-prostoru, ki opisujejo obliko krivulje:

$$y = \{o_x, o_y, p_1, p_2, p_3, p_4\}. \quad (2.8)$$

Končni model gibanja je opisan z mešanico 15 Gaussov, od katerih vsakega določa povprečna vrednost in kovariančna matrika, ki opisuje varianco.



Slika 2.4

Slika prikazuje korake metode na osnovi trajektorij sledenih točk: a. prikazuje trajektorije vseh sledenih točk v sekvenci, b. rezultate glasovanja najboljšega kandidata za dolžino periode, c. krivulje razrezane na posamezne cikle, d. točke preslikanih krivulj v PCA prostor (prve tri najtežje PCA komponente na oseh), elipsoidi pa predstavljajo mešanico Gaussov.

Zgrajene modele hranimo v galeriji modelov, razpoznavanje pa izvajamo tako, da na podoben način poiščemo opisnike značilk testnega primera, ki jih preslikamo v

PCA-prostor vsakega od modelov v galeriji. Model, ki ustreza maksimalni posteriorni verjetnosti pri prisotnosti opisnikov krivulj testnega primera (enačba 2.9), je izbran kot zmagovalni:

$$p(g_i|\mathbf{Y}) \propto \sum_{n=0}^N \log p(\mathbf{y}_n|g_i), \quad (2.9)$$

kjer g_i predstavlja model iz galerije, \mathbf{Y} matriko opisnikov krivulj testnega primera, kjer so v vrsticah posamezni opisniki \mathbf{y}_n . N ustreza številu opisnikov krivulj testnega primera.

Metoda se obnaša podobno, kot je značilno za pristope brez modela. Zaradi odvisnosti od videza, prikazuje relativno dobre rezultate pri nespremenjenih pogojih, a je hkrati izredno občutljiva na različne faktorje vpliva. V primerjavi z metodo baseline pa je zaradi kompleksnosti računsko bolj zahtevna in zato tudi nekoliko počasnejša.

2.3 Pristopi na osnovi modela

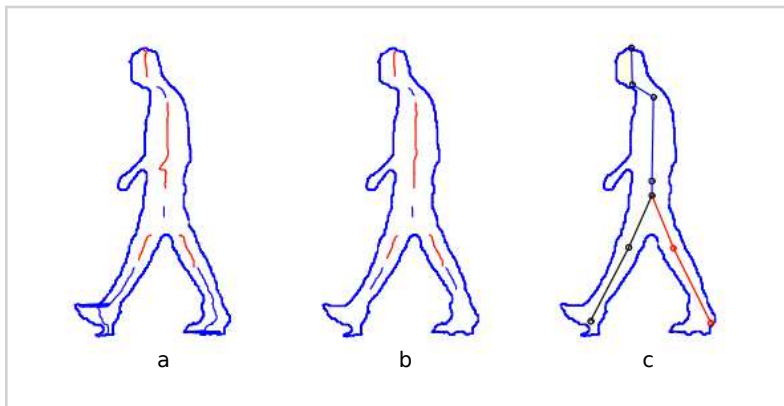
Pristopi na osnovi modela (angl. *Model-based*) [8–10, 26–29]: zgradijo model človeškega telesa ali njegovega gibanja v 3D in pridobijo parametre gibanja iz tega modela (npr. dimenzijo korakov, kadenco, človeško okostje, dimenzijo telesa, pozicijo in orientacijo delov telesa, sklepno kinematiko ipd.). Metode te skupine se osredotočajo predvsem na dinamiko gibanja in manj na posameznikov videz, zaradi česar so tudi bolj odporne proti spremembam pogleda in neodvisne od merila (angl. *scale*). V splošnem pa ne dosegajo tako dobrih rezultatov kot metode, ki pri razpoznavanju uporabljajo videz. Dodatno so te metode računsko bolj zahtevne in posebej občutljive na probleme zaradi prekrivanja.

2.3.1 Skeletni model človeškega telesa

Skeletni model človeškega telesa se je v preteklosti izkazal kot učinkovito orodje za opisovanje človekovega gibanja in se zaradi tega uporablja v številnih pristopih za razpoznavanje gibanja, ki parametre gibanja pridobivajo iz modela [9, 29].

Avtorji v [9] pridobijo biometrični podpis gibanja iz silhuet s pomočjo segmentacije 2D-modela človeškega skeleta in nato uporabijo ta model za nadaljnje pridobivanje parametrov gibanja. Segmentacija skeleta poteka tako, da najprej poiščejo obrobe silhete in nato sredinske točke silhete vzdolž osi y (slika 2.5 a, b), s pomočjo katerih

nato poiščemo premice, ki predstavljajo kosti segmentov telesa (slika 2.5 c). Dolžine segmentov telesa pa pridobijo iz povprečnih anatomskih značilnosti človeškega telesa.



Slika 2.5

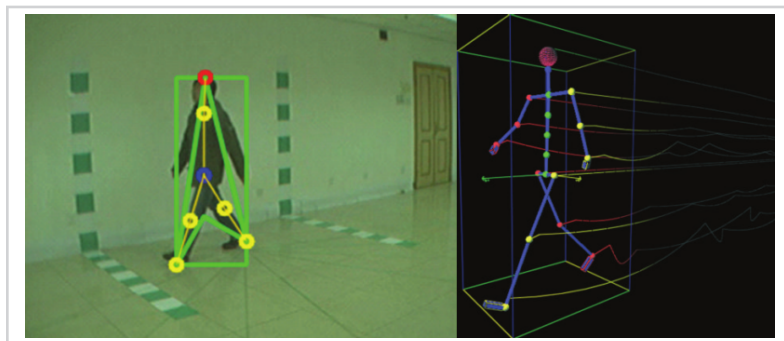
Segmentacija skeleta [9]:
a. obrobe človeške silhete z najdenimi središčnimi točkami, b. odstranjen šum iz središčnih točk, c. poenostavljen model skeleta človeškega telesa.

Iz najdenega skeleta človeškega telesa skozi celotno videosekvenco nato avtorji pridobijo naslednje značilke parametrov gibanja: višino telesa, čas cikla, dolžino koraka, hitrost, povprečne kote sklepov, varianco kota v kolkih, koeficient korelacije med kotoma desne in leve noge ter koordinate centra kolčno-kolenskega ciklograma, ki tvorijo opisnike gibanja, s pomočjo katerih nato izvajajo klasifikacijo s pomočjo algoritma *K*-najbližjih sosedov (angl. *K nearest neighbours* – *KNN*).

Kljub temu, da gre za pristop na osnovi modela, avtorji uporabljajo značilke, ki opisujejo predvsem statične parametre gibanja, kot so višina, čas cikla, dolžina koraka in hitrost, ki imajo sicer visoko moč razločevanja pri nespremenjenih pogojih zajemanja, torej v primeru normalne hoje (tj. naravna, običajna, udobna hitrost hoje, ki je lastna vsakemu posamezniku) s konstantno hitrostjo, a so po drugi strani zelo občutljive na faktorje vpliva, kot sta perspektivna deformacija zaradi spremembe pogleda in sprememba hitrosti hoje. V prvi meri torej metoda predpostavi, da subjekti hodijo normalno s konstantno hitrostjo. Dodatno, kljub temu, da uporabljajo poleg statičnih značilk še parametre gibanja, ki opisujejo dinamiko (koti sklepov), ti niso dobro izkoriščeni za namene razpoznavanja. Uporabijo namreč le povprečja in variance kotov skozi celotno sekvenco, ki so podobno občutljivi na faktorje vpliva kot statične značilke. Avtorji v tej metodi torej ne obravnavajo nobenih faktorjev vpliva.

Avtorji v [29] najprej pridobijo posameznikove značilnosti gibanja s pomočjo prin-

cipa ujemanja spremenljivih predlog (angl. *deformable template matching*) in nato s pomočjo principa dekompozicije pogleda ter z integracijo omejitev, ki izhajajo iz splošnega znanja o človeškem telesu, poiščejo najboljše ujemanje s splošnim 3D-modelom človeškega skeleta. Tako preslikajo zajeto gibanje iz 2D-videosekvenca v gibanje 3D-modela skeleta človeškega telesa (slika 2.6).



Slika 2.6

Prileganje 3D-modela skeleta [29]: Na levi: 2D-videosekvenca hoje subjekta, na desni: 3D-model gibanja skeleta.

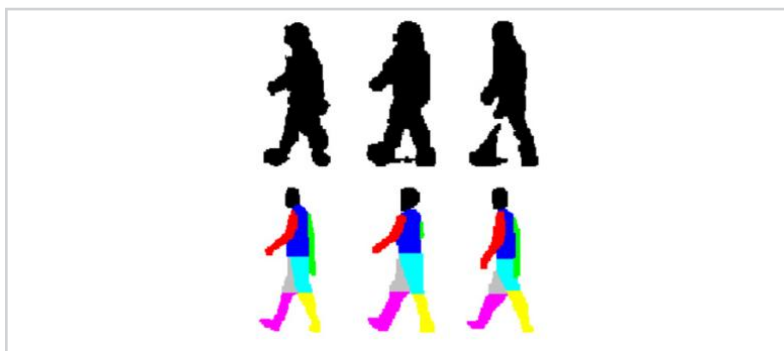
V nadaljevanju avtorji iz zgrajenega 3D-modela človeškega skeleta pridobijo 13 značilk. Gre za časovno-prostorske krivulje, ki jih tvorijo lokacije sklepov gležnjev, kolen, kolkov, zapestij, komolcev in ramen ter lokacija glave. Krivuljam zmanjšajo dimenzionalnost s pomočjo algoritma vdelahe ohranjanja sosednosti (angl. *Neighborhood preserving embedding – NPE*), ki za razliko od PCA lokalno ohranja strukturo v sosesčinah podatkov in je zato manj občutljiva na netipične meritve (angl. *outliers*) [30], ter izvedejo klasifikacijo z integriranjem prikritih pogojnih naključnih polj (angl. *Hidden conditional random fields – HCRF*) [31] v klasifikator podpornih vektorjev (angl. *Support vector machine – SVM*). HCRF omogoča učenje množice skritih spremenljivk, pogojenih z lokalnimi značilkami, in hkrati ne predpostavlja neodvisnosti vzorcev, ki se poleg tega lahko prekrivajo v času in prostoru, SVM pa maksimizira razlike med mejami končnih odločitev klasifikacije v višjedimenzionalnem prostoru značilk.

Kljub temu, da avtorji dosegajo določeno stopnjo neodvisnosti od pogleda s pomočjo rekonstrukcije 3D-modela skeleta, podobno kot avtorji v [9] predpostavljajo konstantno hitrost hoje in uporabljajo le lokacije sklepov skeleta kot značilke za opis gibanja, te pa so prav tako občutljive na različne faktorje vpliva, kot je npr. sprememba hitrosti hoje.

2.4 Komponentni pristopi

Pristope razločujemo tudi na osnovi tega, kako obravnavajo opazovani subjekt. Obravnavamo ga lahko kot celoto (celostni pristopi) ali kot kombinacijo več delov (komponentni pristopi) [32–34]. Ideja izhaja iz medicinskih študij, ki so pokazale, da je človeško gibanje sestavljeno iz sočasnega gibanja več delov telesa, kar nakazuje, da lahko komponentni pristopi izboljšajo rezultate razpoznavanja. Naštete raziskave kažejo, da komponentni pristopi omilijo prisotnost faktorjev, kot so spremembe smeri pogleda, prekrivanja, spremembe v pogojih osvetlitve itd., ki sicer negativno vplivajo na kakovost razpoznavanja.

Pristopi na osnovi modela so predvsem komponentni, medtem ko so pristopi brez modela praviloma celostni. Kljub temu obstajajo primeri iz druge skupine, ki se problema lotijo na komponenten način. Avtorji v [32] uporabljajo ročno označene dele silhuet (slika 2.7) in raziskujejo vpliv posameznih delov na uspešnost razpoznavanja.



Slika 2.7

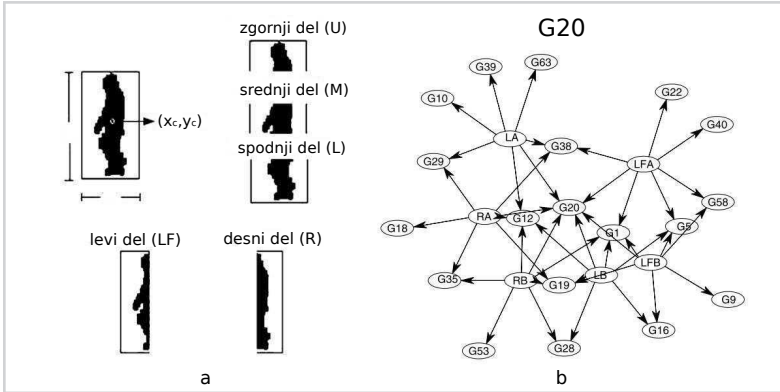
Gibalni podvzorci [32]:
primer ročno označenih
delov telesa silhete.

Avtorji v [34] namesto ročnega označevanja silhuet uporabijo fiksno razdelitev silhete na tri dele po višini in dva dela po širini. Gibanje teh posameznih delov telesa imenujejo gibalni podvzorci, celotno gibanje pa modelirajo z Bayesovimi mrežami (slika 2.8), ki opisujejo verjetnost gibanja subjekta iz galerije z opazovanjem verjetnosti, ki jih pogojujejo posamezni podvzorci. Učenje Bayesovih mrež izvajajo s pomočjo ocenje maksimalne verjetnosti (angl. *Maximum Likelihood Estimation – MLE*), končno razvrščanje modelov iz galerije pa določijo z verjetnostno porazdelitvijo vseh modelov

galerije:

$$P(G_m) = \sum_{s_i} P(s_i)P(G_m|s_i), \tag{2.10}$$

kjer je $P(G_m)$ verjetnost hoje modela m iz galerije, $P(s_i)$ je apriori verjetnost določenega podvzorca, $P(G_m|s_i)$ pa predstavlja verjetnost hoje subjekta m pri prisotnosti vpliva določenih podvzorcev s_i . Model iz galerije z največjo verjetnostjo predstavlja torej najboljše ujemanje.



Slika 2.8

Gibalni podvzorci [34]: a. razdelitev silhuete na več delov imenovanih gibalni podvzorci, b. primer tipične Bayesove mreže, ki prikazuje mapiranje podobnosti med podvzorci in vzorci hoje iz galerije, na katere vplivajo.

Metoda demonstrira robustno delovanje pri prisotnosti nekaterih faktorjev vpliva, kot je npr. manjša variacija pogleda, in dobro odpornost na napake v silhuetah npr. v primeru manjkajočih delov silhuet. Grafična narava Bayesovih mrež pa omogoča intuitivno vizualizacijo relacij med posameznimi podvzorci in prikazuje, kako ti kolektivno prispevajo k procesu razpoznavanja.

2.5 Biološki pristopi

Obstajajo tudi pristopi, ki temeljijo na bioloških spoznanjih o človeškem zaznavanju gibanja [35] in vsebujejo nekatere osnovne koncepte za posnemanje človeškega korteksa v povezavi z razpoznavanjem objektov. Pridobljena spoznanja so se že izkazala za uspešna tako na področju razpoznavanja akcij [36, 37] kot na področju razpoznavanja gibanja [38]. S t. i. biološko značilko (angl. *biologically inspired feature – BIF*) lahko predstavimo gibanje. Njena prednost je predvsem v neobčutljivosti na spremembe v

skali in poziciji, slabosti pa ostajata problem občutljivosti na spremembo pogleda in visoka dimenzionalnost podatkov.

2.6 Faktorji vpliva

Na različne načine se pristopa tudi k reševanju problema varianc, ki jih povzročajo različni vplivi na gibanje. Pristope lahko ločimo glede na to, kako obravnavajo te faktorje, čeprav gre v večini konkretnih primerov reševanja problematike za kombinacijo spodnjih pristopov:

Na osnovi izbora značilke [39, 40]:

- Osredotočajo se na iskanje in izbiranje najboljših značilke gibanja, ki so invariantne na faktorje vpliva.
- Temeljijo na fizikalnih oz. fizioloških predpostavkah, ki izhajajo iz domenskega predznanja o okoliščinah faktorjev vpliva, ki jih obravnavajo.
- Obravnavajo le posamezne tipe faktorjev vpliva in običajno ne postavljajo splošnih rešitev (npr. spremembo hitrosti ali spremembo pogleda).
- Ne delujejo dobro pri ekstremnih spremembah pogojev (npr. večje spremembe pogleda).

Na osnovi transformacij [24, 41–44]:

- Delujejo bolje tudi pri ekstremnejših spremembah.
- Obravnavajo posamezne tipe faktorjev vpliva.

Na osnovi učenja [45, 46]:

- Poskušajo maksimizirati medrazredne razdalje in istočasno minimizirati znotrajrazredno varianco.
- Uporabimo jih lahko kot dodatno izboljšavo skupaj s preostalima dvema načinoma.
- Predvidevajo prisotnost istih razredov (subjektov) in istih pogojev faktorjev vpliva, tako v učni kot v testni množici, ter hkrati predpostavljajo dovoljšno zastopnost v številu vzorcev za vse pogoje, ki jih obravnavajo.

2.6.1 Sprememba hitrosti hoje

Vpliv spremembe hitrosti hoje je eden najbolj problematičnih faktorjev vpliva, saj spreminja dinamiko gibanja in je skoraj vedno prisoten v sistemih za razpoznavanje ljudi po načinu hoje. Posebno obravnavo mu posvečajo avtorji spodaj opisanih del.

Avtorji v [24] raziskujejo vpliv spremembe hitrosti hoje na učinkovitost razpoznavanja pri uporabi kadence in dolžine koraka ter predlagajo izboljšavo s pomočjo normalizacije silhuet. Predlagajo normalizacijo koraka silhuet dvojne opore na osnovi statističnih relacij med hitrostjo hoje in dolžino koraka. Za demonstracijo razpoznavanja uporabijo osnovni algoritem [12] na petih silhuetah gibalnega cikla (dve iz faze enojne opore in tri iz faze dvojne opore). Izbor petih ključnih faz cikla sicer poenostavi postopek, a se pri tem lahko izgubijo informacije o gibanju iz preostalih silhuet, ki bi lahko dodatno prispevale k boljšim rezultatom.

Avtorji v [47] predlagajo iskanje značilk gibanja na osnovi časovne normalizacije s pomočjo skritih markovih modelov (angl. *Hidden markov models – HMM*), katerih učinke demonstrirajo na hitri in počasni hoji. Metoda sicer nevtralizira vpliv časovnih sprememb (npr. trajanje cikla), a pri tem ne obravnava prostorskih sprememb, ki jih povzroči sprememba hitrosti (npr. dolžina koraka).

Avtorji v [48] predstavijo gibanje kot variacijo oblike na osnovi izometričnih vzorcev (angl. *Shape Variation Based Frieze Patterns – SVBFP*), ki zajame gibanje v času in predstavlja normalizirano razliko med okvirji znotraj cikla gibanja. Nato zgradijo časovno karto simetrijskih gibalnih vzorcev in s kombinacijo z vertikalnimi ter horizontalnimi izometričnimi vzorci določijo mero za oceno podobnosti med videosekvencami hoje.

Avtorji v [49] predlagajo pristop na osnovi dinamičnega ukrivljanja časa (angl. *Dynamic Time Warping – DTW*), ki uporablja množico DTW-funkcij za predstavitev distribucije vzorcev gibanja z uporabo enakomernih (angl. *uniform*) in ukrivljenih Gaussovih (angl. *wrapped Gaussian*) porazdelitev.

Avtorji v [50] predlagajo 3D-metodo (x, y in čas) avtokorelacije, ki učinkovito izloči časovno-prostorske lokalne geometrijske značilke, ki karakterizirajo gibanje. Imenujejo jo kubična lokalna avtokorelacija višjega reda (angl. *Cubic Higher-order Local Auto-Correlation – CHLAC*). Je relativno odporna proti manjšim variacijam hitrosti hoje, saj uporablja le vsote lokalnih značilk skozi celotno sekvenco gibanja in kot taka ne uporablja informacije o fazah gibanja eksplicitno. Metoda je torej učinkovita le za

manjše variacije hitrosti znotraj posameznih sekvenc hoje.

Avtorji v [44] predlagajo transformacijo prostornine dinamike silhuete (angl. *Gait Silhouette Volume Transformation – GSVT*). S pomočjo iskanja ujemanja z modelom človeka izločijo statične in dinamične značilke iz silhuet ter za dinamične značilke s pomočjo učnih primerov različnih subjektov in različnih hitrosti hoje generirajo transformacijski model na osnovi faktorizacije. Model omogoča transformacijo dinamičnih značilk iz referenčne hitrosti hoje v poljubno hitrost.

Avtorji v [51] predlagajo nov opisnik značilke, imenovan konfiguracija oblik višjega reda (angl. *High-order Shape Configuration – HSC*), ki zna zgraditi robustne značilke, odporne proti spremembam oblike telesa, ki jih povzroča sprememba hitrosti hoje. Pri tem uporabijo prokrustno analizo oblik (angl. *procrustes shape analysis*) za podpis gibanja, iz katerega HSC pridobi razločevalno informacijo, pri tem pa še vedno tolerira spremembe hitrosti hoje. Metodo nadgradijo z uporabo modela razlik v kompoziciji (angl. *Differential Composition Model – DCM*) [52], ki razlikuje med različnimi vplivi, ki jih ima sprememba hitrosti hoje na razne dele človeškega telesa.

Nobeden od zgornjih primerov ne sodi v skupino pristopov na osnovi modela, kar kaže na splošno pomanjkanje zanimanja za reševanje problema spremembe hitrosti pri pristopih te skupine. S predlagano metodo na osnovi modela predstavljeno v nadaljevanju tega dela želimo to vsaj do neke mere popraviti.

2.6.2 Sprememba pogleda

Drugi problematičen faktor predstavljajo spremembe pogleda, ki so v realnih sistemih prav tako zelo pogoste in jih je zato treba posebej obravnavati za uspešno delovanje. Obstajajo poskusi uporabe prijemov iz geometrije več pogledov [53, 54], vendar je tu dodatna težava, da ima lahko isto gibanje drastično spremenjen videz v sliki, ko ga opazujemo iz različnih pogledov. Zaradi tega klasifikatorjev na osnovi pogleda ne moremo dobro posplošiti za uporabo v sistemih z več kamerami ali več pogledi. V splošnem je potreben ločen klasifikator za vsak pogled. Avtorji v [41–43, 55] se osredotočajo na posebno obravnavo spremembe pogleda pri zajemanju gibanja.

2.6.3 Drugi faktorji

Med pogosteje obravnavane faktorje vpliva sodijo še sprememba obleke [40, 56] in nekateri manj zastopani faktorji, kot so sprememba obutve, terena, nošenje predmetov ter daljši časovni razmiki v času zajemanja [45].

2.7 Podatkovne zbirke za evalvacijo

Dostopnost standardnih javnih evalvacijskih zbirk je vedno ključnega pomena za vsako nalogo razpoznavanja vzorcev, saj omogočajo študijo potenciala metode za določeno nalogo s konsistentno evalvacijo in primerjavo učinkovitosti. Razpoznavanje ljudi po načinu hoje tu ni izjema.

Evalvacijske zbirke za razpoznavanje ljudi obravnavajo dva vidika, ki vključujeta unikatnost in praktičnost. Prvi običajno razkrije število in raznolikost subjektov v zbirki, drugi pa pogoje zajemanja (kakovost slik). Če je namen zbirke prikazati zmožnost reševanja problema, potem je zbirka lahko zajeta v laboratorijskem okolju, če pa je namen prikazati zmožnost, da nalogo opravi računalniški vid, pa je treba vključiti tudi posnetke iz zunanjega okolja in realnih scen. Zbirke težijo k temu, da vključujejo dovolj vzorcev različnih subjektov za oceno znotraj in medrazredne variance. Trenutne zbirke imajo večinoma manjše število subjektov kot zbirke za evalvacijo drugih biometrik, kot so prstni odtisi in obrazi, a se pri tem bolj osredotočajo na vključevanje variacij pogojev hoje (faktorji vpliva), ki so za razpoznavanje ljudi v realnem okolju ključnega pomena. Ti so pomembni za evalvacijo robustnosti metod, saj se pogoji hoje običajno razlikujejo med primeri iz galerije in testnimi primeri, medtem ko je številko subjektov v zbirki pomembno za zagotavljanje statistične zanesljivosti rezultatov evalvacije. Poleg tega se lahko zbirke uporabljajo za evalvacijo t. i. mehkih biometričnih metod (razločevanje spola, starosti) in takrat igra raznolikost subjektov z vidika teh dveh dejavnikov ključno vlogo za zanesljivost rezultatov.

Večina današnjih zbirk obravnava predvsem variacije pogojev hoje, medtem ko sta število in raznolikost subjektov, razen zelo redkih primerov [57], še vedno premalo obravnavana za zanesljivejše biometrične eksperimente. Pregled evalvacijskih zbirk z osnovnimi značilnostmi podajamo v tabeli 2.1, v nadaljevanju pa podajamo kratke opise najvidnejših predstavnikov.

Zbirka CMU Mobo [58] je nastala na univerzi Carnegie Mellon. Vsebuje 25 oseb, ki hodijo po tekoči stezi (angl. *threadmill*) v notranjosti laboratorija. Vsebuje štiri različne načine hoje, to so hitra hoja, počasna hoja, hora v klanec in hoja z žogo, ter šest različnih pogledov posnetih istočasno s šestimi enakomerno porazdeljenimi kamerami.

Zbirka USF [12] je nastala na univerzi Južne Floride (angl. *University of South Florida*). Vsebuje 122 oseb, vsaka oseba hodi okrog eliptične poti pred kamerami, posnetki pa so nastali zunaj, pred zgradbo univerze. Zbirka obravnava do pet faktorjev vpliva

za vsako osebo (dva različna tipa čevljev, nošenje kovčka, travnata oz. betonska podlaga, levi in desni kot pogleda zajemanja ter dve različni časovni obdobji pridobivanja posnetkov z razmakom pol leta).

Zbirka Sotton Human ID [59] je nastala na univerzi v Southamptonu. Obstajata dve različici, manjša (12 oseb) in večja (115 oseb). Večja je bila posneta v zunanosti in v laboratoriju (poligon, tekoča steza) z ozadjem enakomerne zelene barve za natančnejšo segmentacijo gibanja. Dodatno obravnava šest različnih pogledov zajemanja.

Zbirka CASIA [60, 61] je nastala na inštitutu za avtomatizacijo kitajske akademije znanosti (angl. *The Institute of Automation Chinese Academy of Sciences*). Vsebuje tri zbirke: A, B in C. Najpogosteje uporabljena je zbirka B, ki zajema 124 oseb in 11 različnih pogledov zajemanja, zato je najprimernejša za raziskave vplivov spremembe pogleda. Poleg tega vsebuje posnetke za normalno hojo, nošenje torbe ter spremembo obleke. Zbirka C zajema 153 oseb, tri hitrosti hoje in nošenje predmetov. Zajeta je z IR-kamero za natančnejšo segmentacijo gibanja.

Zbirka OU-ISIR [57, 62] je nastajala na univerzi v Osaki od leta 2007 dalje. Prvi del predstavlja zbirka na tekoči stezi, katere glavna značilnost so predobdelava in ročno pripravljene silhuete. Obravnava spremembo hitrosti hoje (oznaka A v zbirki zajema 34 oseb), spremembo obleke (oznaka B v zbirki zajema 68 oseb) in spremembo pogleda (oznaka C v zbirki zajema 200 oseb). Leta 2010 pa so zajeli tudi največjo zbirko za razpoznavanje ljudi z vidika števila in raznolikosti subjektov [57], ki je danes edina tovrstna zbirka. Žal tudi ta velikost še vedno ni primerljiva s sodobnimi zbirkami drugih biometrik, kot so prstni odtisi in obrazi.

Zbirka TUM [63] je nastala na tehnični univerzi v Münchnu. Vsebuje 35 oseb, ki hodijo v različnih konfiguracijah hoje: normalna hoja, hoja z rokami v žepu, hoja z nahrbtnikom, hoja v halji, hoja z dinamičnim prekrivanjem (dve mimoidoči osebi) in hoja s statičnim prekrivanjem (dve stoječi osebi). Zbirka je edinstvena v pogledu obravnavanja prekrivanj, saj jih posebej ne obravnava nobena druga zbirka.

Tabela 2.1

Pregled podatkovnih zbirk za evalvacijo metod za razpoznavanje ljudi po načinu hoje.

Zbirka	Osebe	Posnetki	Leto	Ločljivost	Okvirji (FPS)	Okolje	Faktorji vpliva
CMU Mobo [58]	25	600	2001	640×480	30	notranjost, tekoča steza	6 pogledov, 2 hitrosti, predmeti, klanec
Soton [59]	115	2128	2001	2240×1680	25	notranjost, zunanost, tekoča steza zeleno ozadje	6 pogledov
USF [12]	122	1870	2001	720×480	30	zunanost	2 pogleda, čas, podlaga, čevlji, predmeti
CASIA (A) [60]	20	240	2001	352×240	25	zunanost	3 pogledi
CASIA (B) [60]	124	13600	2005	320×240	25	notranjost	11 pogledov, obleka, predmeti
CASIA (C) [61]	153	1530	2005	320×240	25	zunanost, noč, IR-kamera	hitrost, predmeti
OU-ISIR (A) [62]	34	612	2007	640×480	60	notranjost, tekoča steza	9 hitrosti
OU-ISIR (B) [62]	68	2746	2007	640×480	60	notranjost, tekoča steza	obleka
OU-ISIR (C) [62]	200	5000	2007	640×480	60	notranjost, tekoča steza	25 pogledov
OU-ISIR velika [57]	3835	7670	2012	640×480	30	notranjost, zeleno ozadje	5 pogledov
TUM [63]	35	840	2010	640×480	30	notranjost, hodnik	čas, predmeti, prekrivanja

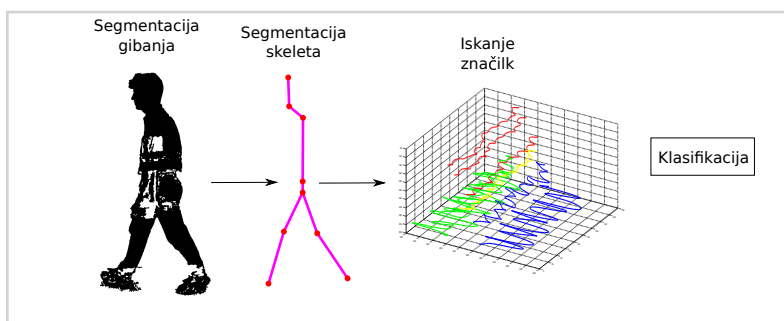


Predlagana metoda

3.1 Arhitektura predlagane rešitve

Arhitekturno metodo gradijo štirje večji gradniki (slika 3.1), ki jih bomo podrobneje obdelali v nadaljevanju:

- segmentacija gibanja (angl. *motion segmentation*),
- segmentacija skeleta človeškega telesa,
- iskanje in obdelava značilk ter
- klasifikacija.



Slika 3.1

Arhitektura sistema za razpoznavanje ljudi po načinu hoje s pomočjo skeleta človeškega telesa.

3.2 Segmentacija gibanja

Segmentacija gibanja je prvi in nujni korak vseh metod, ki kakorkoli analizirajo gibanje v videoposnetku. Omogoča ločitev ozadja videoposnetka od ospredja, pri čemer ozadje predstavljajo statični, ospredje pa dinamični elementi videoposnetka. V našem primeru se v videoposnetku gibljejo ljudje, torej gre za iskanje silhuet oseb, ki se v videoposnetku premikajo. Pri tem uporabimo standardni postopek za segmentacijo gibanja na osnovi Gaussovih mešanic [25]. Ker večina podatkovnih zbirk za evalvacijo metod razpoznavanja po načinu hoje že vsebuje izločene silhuete, temu delu, ki je sicer aktivno področje raziskovanja (glej [64] za pregled področja), v nadaljevanju ne posvečamo posebne pozornosti. Kot vhod metode, opisane v nadaljevanju, predpostavimo, da so silhuete na voljo ali kot del podatkovne zbirke ali kot rezultat omenjenega

postopka za segmentacijo gibanja. Zaradi tega je smiselno predpostaviti, da so silhuete bolj ali manj podvržene šumu, kar je pri izgradnji metode v nadaljevanju treba ustrezno obravnavati.

V prvi vrsti uporabimo za izboljšanje kakovosti silhuet morfološke operacije, s čimer zmanjšamo vpliv nekaterih tipov šuma, kot so npr. luknje, beli šum, sence ipd. Z morfološkim zapiranjem (angl. *close*) zapolnimo manjše luknje v silhueti. Morfološko odpiranje (angl. *open*) odstrani beli šum in osamljene elemente ospredja. Z morfološko dilacijo (angl. *dilate*) naknadno zapremo večje luknje, nazadnje pa odstranimo še večje osamljene elemente, ki so premajhni, da bi predstavljali del telesa oz. silhuete.

Preostali postopki, ki preprečujejo vpliv šuma v silhuetah na nadaljnje korake metode, so podrobneje obravnavani v podpoglavju 5.6.

3.3 Segmentacija skeleta človeškega telesa

Za namen segmentacije skeleta človeškega telesa smo prilagodili postopek iz [9], ki opisuje relativno preprosto metodo za pridobitev skeleta iz enega pogleda 2D-videosekvence, a kljub temu ohranja dovolj informacije o posameznikovem gibanju za potrebe identifikacije. Postopek segmentacije skeleta vsebuje naslednje korake:

Najprej poiščemo obris (angl. *contour*) silhuete iz vsake sličice videosekvence. Nato izračunamo višino silhuete in na osnovi te razdelimo telo na posamezne segmente glede na povprečne anatomske značilnosti človeškega telesa [65]. Za neko višino H podamo začetne ocene za vertikalne pozicije vratu, rame, pasu, medenice, kolen in gležnjev kot $0,870H$, $0,818H$, $0,530H$, $0,480H$, $0,285H$ in $0,039H$ (horizontalne črtkane linije na sliki 3.2).

Nato razdelimo silhuete na posamezne dele telesa (glava, vrat, trup, pas, zgornji del obeh nog, spodnja leva in spodnja desna noga, levo ter desno stopalo) glede na ocenjene dolžine posameznih delov telesa (slika 3.2 a). Za vsak segment poiščemo točke levega in desnega robu ter izračunamo njuno sredino (slika 3.2 b). Nadaljnje poiščemo premice skozi te točke s pomočjo Houghove transformacije in izberemo tisto z največjo vrednostjo v akumulacijskem polju (slika 3.2 c). Koti teh premic predstavljajo kote posameznih delov telesa, na osnovi katerih skupaj z ocenjenimi dolžinami segmentov določimo lego posameznih sklepov v telesu, pri čemer lokacije sklepov določamo po vrsti od glave do gležnjev. Lokacije sklepov dokončno definirajo skelet človeškega telesa, kot ga prikazuje slika 3.2 d.

Medtem ko je segmentacija zgornjih delov telesa relativno preprosta, predstavljajo

spodnji deli (obe stegenici in golenici) precejšen problem, kar zahteva posebno obravnavo. Najprej poiščemo golenici, s pomočjo katerih potem določimo še stegenici. V fazi enojne opore, ko se nogi prekrivata, najde zgoraj opisani algoritem le en segment v spodnjem delu. V tem primeru uporabimo kot oceni za kota leve in desne golenice kar točke levega in desnega robu celotnega spodnjega segmenta, medtem ko sta v primeru faze dvojne opore obe golenici določeni z ločenima segmentoma telesa, kar omogoča uporabo sredinskih točk segmentov, podobno kot ocenjevanje kotov v zgornjem delu telesa. Kota obeh stegenic nato določimo s pomočjo sredinske točke pasu (sklep v kolkih) in začetnih točk (kolenska sklepa) leve ter desne golenice. Nazadnje tako s pomočjo začetnih točk, ocenjenih kotov in dolžin segmentov izračunamo še končne točke vseh kosti v telesu (slika 3.2 d).

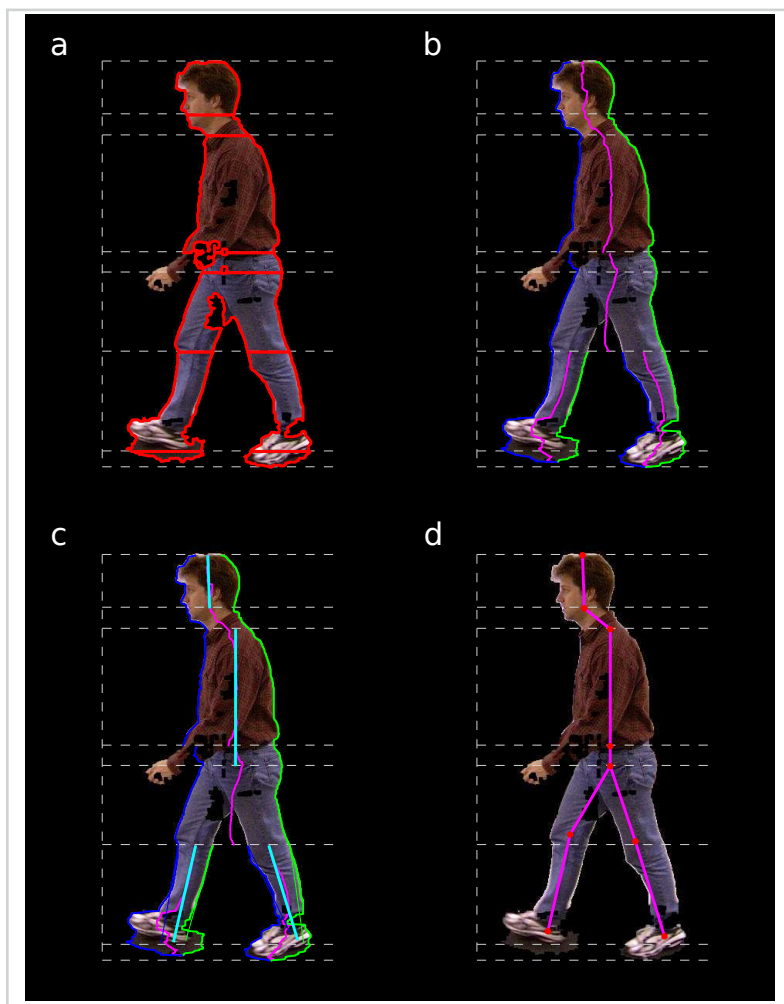
Rezultat opisane segmentacije je torej poenostavljena oblika skeleta (brez rok, ramen, kolkov) človeškega telesa za vsako posamezno silhueto videosekvence, ki predstavlja biometričen podpis posameznikove hoje. Takšen podpis omogoča neposredno pridobivanje značilk, kot so npr. koti kosti in sklepov, ki tvorijo v času signale časovnih vrst (angl. *time-series signals*), na osnovi katerih lahko izvajamo identifikacijo.

3.4 Iskanje gibalnih značilk

Iz pridobljenih biometričnih podpisov nato izluščimo množico signalov v obliki časovnih vrst. Pri sami izbiri se opiramo na delo pionirjev področja analize gibanja z medicinske domene [66], ki so predlagali nekaj več kot 20 značilnosti gibanja, ki najbolj specifično definirajo posameznikovo hojo:

- koti delov telesa: glava, vrat, trup, leva in desna stegenica, leva in desna golenica,
- lokacije sklepov: levo in desno koleno ter gležnja,
- mase in centri mas (angl. *Center of mass – COM*) za celotno silhueto in posamezne dele telesa ter
- širina in višina silhuete.

Vsi ti t. i. surovi signali predstavljajo osnovo za izvedene signale, ki se na koncu dejansko uporabljajo pri razpoznavanju. V medicinski domeni so koti sklepov definirani relativno glede na dva sosednja sklepa. Relativne kote izračunamo iz pridobljenih



Slika 3.2

Posamezni koraki algoritma za segmentacijo skeleta človeškega telesa.

vrednosti surovih kotov. Pri normalni hoji je trup skoraj vertikalni, zato sta relativna kota bokov kar enaka pridobljeni surovi vrednosti kota stegenice, medtem ko relativ-

ne kote kolenskih sklepov izračunamo kot:

$$\theta_{\text{kolenno}} = \theta_{\text{stegnenica}} - \theta_{\text{golenica}}. \quad (3.1)$$

Fazne razlike v kotih obeh stegnenic oz. golenic, ki opisujejo korelacijo med levo in desno stegnenico oz. golenico in je lahko za posameznika zelo specifična, izračunamo po formuli:

$$\theta_{\Delta} = \theta_{\text{leva}} - \theta_{\text{desna}}. \quad (3.2)$$

Razdalje med centri mas posameznih delov telesa in centrom mase celotnega telesa prav tako določajo specifične posameznikove hoje:

$$D_{c_i} = \sqrt{(c_{i,x} - c_x)^2 + (c_{i,y} - c_y)^2}, \quad (3.3)$$

kjer sta (c_x, c_y) koordinati centra mase celotnega telesa, $(c_{i,x}, c_{i,y})$ pa koordinati centra mase i -tega dela telesa.

Skupno tako pridobimo vsega 64 signalov značilik, ki jih kasneje uporabimo pri klasifikaciji.

3.4.1 Parametri gibanja na osnovi slikovnih momentov

Slikovni momenti (angl. *image moments*) binarnih slik in njihovi derivati, kot je npr. center mase, so se izkazali kot zelo uporabni za ocenjevanje parametrov gibanja. Na primer za registracijo silhuet [48, 67], kot značilka v povezavi z razdaljo točk od centra mase [68], za določanje frekvence gibanja [9] ipd. Še več, specifičnost centra gravitacije, katerega geometrijska analogija je ravno center mase, v relaciji s centrom vrtenja človeškega telesa analizirajo že zgodnje psihološke študije na temo razpoznavanja ljudi po načinu hoje [69]. V nadaljevanju so podrobneje opisani uporaba slikovnih momentov v naši metodi za ocenjevanje frekvence gibanja, iskanje začetkov gibalnih ciklov in posameznih faz cikla hoje (faza enojne, dvojne opore).

Slikovne momente definiramo kot:

$$M_{ij} = \sum_x \sum_y x^i y^j I(x, y), \quad (3.4)$$

kjer ij predstavlja red momenta in $I(x, y)$ vrednost slike v točki (x, y) . Ker binarne slike silhuet vsebujejo le vrednosti 0 in 1, lahko izračunamo center mase kot:

$$\mathbf{C} = \left[\frac{M_{10}}{M_{00}}, \frac{M_{01}}{M_{00}} \right], \quad (3.5)$$

kjer sta M_{10} in M_{01} vsoti vseh koordinat x in y vseh neničelnih elementov v sliki, M_{00} pa predstavlja število vseh neničelnih elementov. Temu pravimo tudi masa objekta oz. silhuete, saj predstavlja analogijo s fizikalno maso objekta na sliki. Oba opisana momenta, tako maso M_{00} kot center mase C , uporabimo tudi kot značilko za razpoznavanje.

3.4.2 Detekcija ciklov gibanja

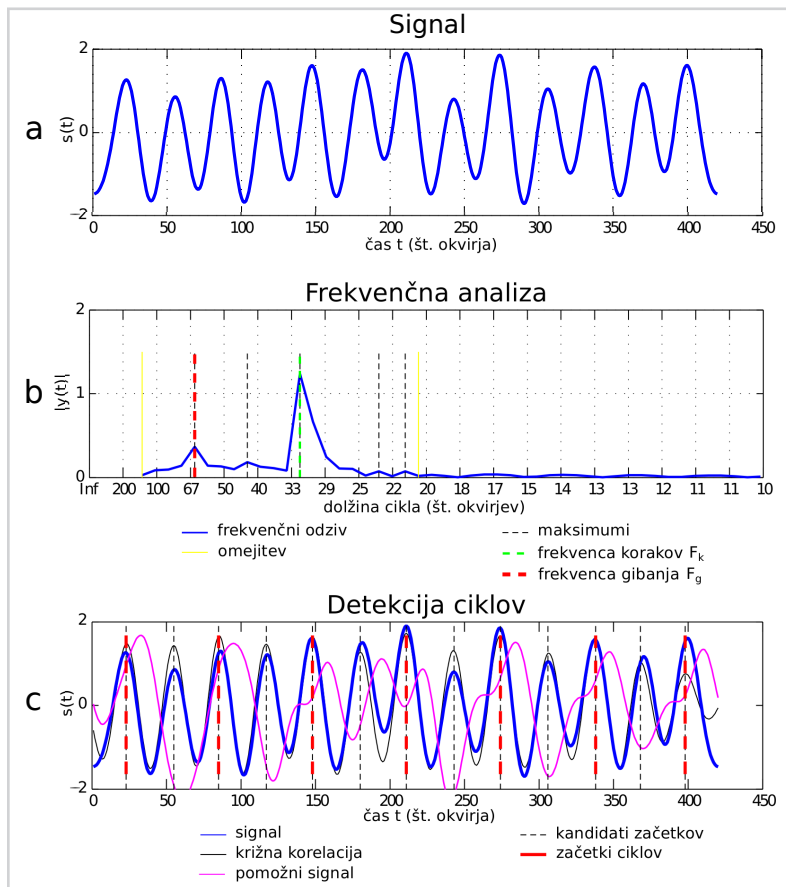
Videosekvence hoje različnih osebkov v splošnem med seboj niso sinhronizirane in se lahko začnejo na katerikoli fazi gibalnega cikla. Da olajšamo kasnejšo primerjavo ciklov med seboj, moramo definirati, v kateri fazi se naj gibalni cikel začne. Kot najbolj primerna se tu izkaže faza dvojne opore, ki jo je najlažje poiskati.

Najprej določimo frekvenco gibanja na osnovi frekvenčne analize pridobljenih signalov gibanja. Za ta namen so najprimernejši signali vertikalno gibanje (koordinata y) centra mase, masa celotnega telesa in širina silhuete. Skupna značilnost vseh treh je, da imajo signali maksimume pri fazah dvojne opore in minimume pri fazah enojne opore. Ko opazujemo osebo iz stranskega pogleda, je center mase na najnižji točki (najvišja vrednost koordinate y), masa telesa je na najvišji točki, ker so roke in noge maksimalno raztegnjene ter zavzamejo več prostora v sliki kot pri fazi enojne opore. Isto velja tudi za širino silhuete, ki je še najbolj občutljiva na šum v silhuetah, ker nanjo močno vplivajo premiki rok. Širina je zato najmanj zanesljiva in manj primerna za detekcijo ciklov. Po drugi strani je center mase najbolj odporen na šum, a nekoliko manj stabilen v smislu izražanja maksimumov, ki nakazujejo fazo dvojne opore, kot masa celotnega telesa. Ta je zato najbolj primerna za iskanje ciklov, kadar imamo opravka s kakovostnimi silhuetami, ob prisotnosti večje količine šuma pa je primernejša izbira center mase.

V našem primeru smo za nadaljnjo analizo z diskretno Fourierjevo transformacijo (DFT) izbrali maso. V pridobljenem frekvenčnem spektru signala izstopata dve frekvenci z največjima amplitudama (slika 3.3 b). Frekvenca koraka F_k predstavlja eno polovico cikla hoje in ima pri stranskem pogledu običajno največjo amplitudo, saj ustreza korakom hoje, ki so najbolj zastopani v signalu. Frekvenca gibanja F_g ima nekoliko nižjo amplitudo in predstavlja osnovni približek frekvence gibalnih ciklov, ki so sestavljeni iz dveh korakov. Zato velja:

$$F_g \approx 2F_k. \quad (3.6)$$

Iskanje frekvenc je dodatno omejeno na interval $[F_v/3, 2F_v]$, ki ga določimo s pomočjo frekvence vzorčenja F_v , podane v okvirjih na sekundo (angl. *frames per second – FPS*), da preprečimo ekstremne vrednosti v primeru napak.



Slika 3.3

Analiza frekvence gibanja: slika a. prikazuje normaliziran signal mase z maksimumi v fazah dvojne opore (polovica cikla gibanja), slika b. prikazuje spekter DFT z izstopajočima frekvencama: frekvenco korakov in frekvenco gibanja. Enote osi x so podane s številom okvirjev dolžine cikla, ki ustreza pripadajoči frekvenci f : $x = F_v/f$, kjer je F_v frekvenca vzorčenja v FPS. Slika c. prikazuje najdene začetke ciklov gibanja. Faze dvojne opore, ki se začnejo z levo nogo, določimo s pomočjo vrednosti pomožnega signala (center mase stopala) ob maksimumih glavnega signala.

3.4.3 Določanje faze dvojne opore

Vsak cikel gibanja vsebuje dve fazi dvojne opore. Prvo, ko je spredaj leva noga, in drugo, ko je spredaj desna noga. Da lahko kasneje primerjamo signale med seboj, je

nujno določiti začetke ciklov tako, da se vedno začnejo v isti fazi dvojne opore. Torej, ali tiste z levo nogo (pogled od strani, hoja od desne proti levi) ali tiste z desno nogo (pogled od strani, hoja od leve proti desni). To storimo tako, da poiščemo vse faze dvojne opore, ki se začnejo z isto nogo, s čimer omogočimo vsaj grobo poravnavo ciklov hoje različnih subjektov, ki jih primerjamo. Takšna poravnava je potrebna, ker se del cikla hoje, ki se začne z levo nogo, razlikuje od cikla hoje, ki se začne z desno nogo, delno zaradi samih značilnosti hoje, ki so lahko pri posamezniku različne za levo in desno nogo, delno pa zaradi vpliva perspektivne projekcije, preko katere zajemamo gibanje. Razlike med levo in desno nogo se odražajo na značilkah celotnega telesa in niso omejene le na obe nogi, zato zgolj menjava leve in desne noge ne pride v poštev.

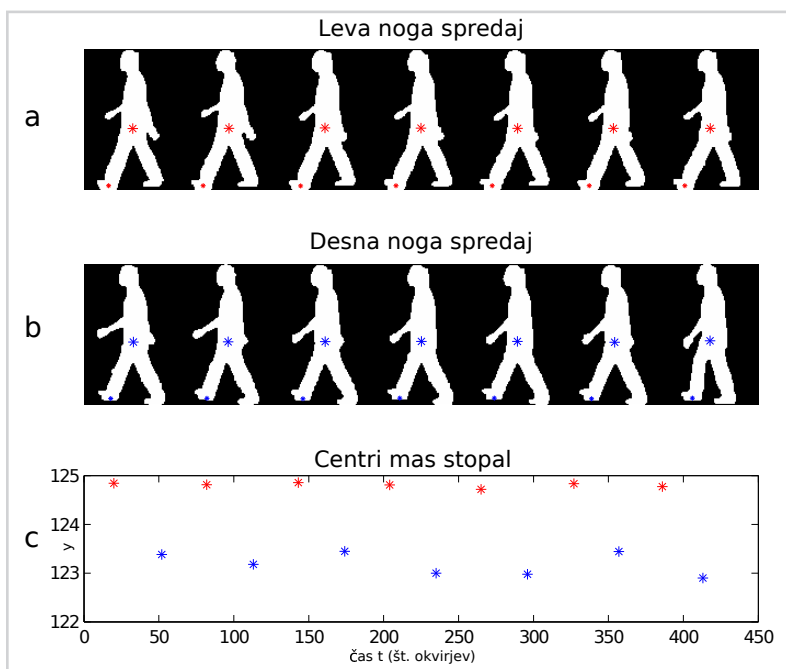
Problem rešujemo tako, da izkoriščamo hibo perspektivne projekcije, ki se pri stranskem pogledu na hojo odraža tako, da se noga, ki je bližje kameri, pojavlja na posnetku nekoliko nižje kot noga, ki je bolj oddaljena od kamere (slika 3.4). To lastnost uporabimo preko analize pomožnega signala nog (koordinata y sprednjega stopala) ob fazah dvojne opore. Ko je noga, ki je bližje kameri, spredaj, je vrednost COM_y stopala višja kot vrednost COM_y stopala noge, ki je bolj oddaljena od kamere.

Na koncu izračunamo začetke ciklov na osnovi ocenjene frekvence gibanja in faz dvojne opore, ki se začnejo z isto nogo. Te uporabimo za izračun začetne ocene dolžine cikla in začetka centralnega cikla, ki je osnova za avtokorelacijo. Tako pridobljeni del signala predstavlja torej en cikel v celotnem posnetku hoje, ki ga avtokoreliramo čez celoten signal. Maksimumi avtokorelacije določajo začetke preostalih ciklov gibanja v dotičnem posnetku hoje (slika 3.3 c).

3.4.4 Detekcija križanja nog

Enega od problemov postopka za tako pridobljen biometrični podpis predstavljajo križanja nog. Če opazujemo hojo iz stranskega pogleda, se nogi križata v fazi enojne opore. Prej opisani postopek ne more zaznati takšnega križanja, ampak vedno tretira sprednjo nogo kot prvo in zadnjo kot drugo. Ko se nogi prekrížata, pa se prednja in zadnja noga zamenjata.

Avtorji v [9] rešujejo ta problem z izkoriščanjem fizikalnih predpostavk, kot je npr. dejstvo, da se noga, ki je v stiku s tlemi, ne premika, torej ne spreminja koordinat v sliki. Izkoriščanje tovrstnih predpostavk je lahko problematično npr. pri hoji po tekoči stezi (angl. *threadmill*), kjer se steza premika skupaj z nogo med njunim stikom. V razviti metodi zato uporabimo vertikalno gibanje centra mase, ki odpravi omenjeno



Slika 3.4

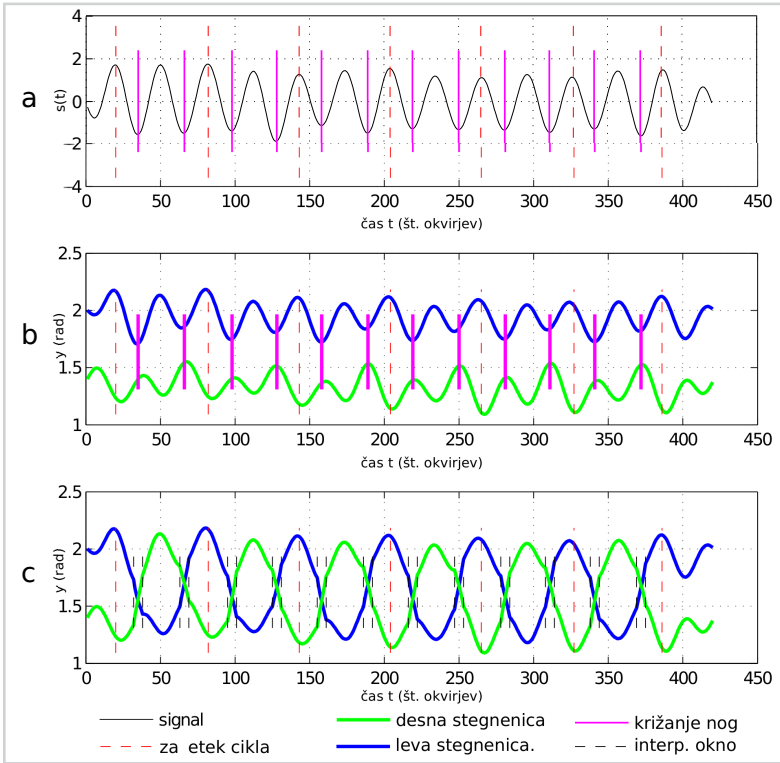
Iskanje faze dvojne opore: slika a. prikazuje fazo dvojne opore, ko je spredaj leva noga, slika b. prikazuje fazo dvojne opore, ko je spredaj desna noga. Rdeče in modre oznake prikazujejo lokacijo centrov mas telesa in stopal. Slika c. prikazuje razmik vrednosti COMy med levo in desno fazo dvojne opore. Os x označuje čas v okvirjih videoposnetka, os y pa označuje koordinato y COM.

pomanjkljivost in je torej bolj primerno za detekcijo točke (časa) križanja nog.

Opazimo lahko, da se lokacije (časovne) faze enojne opore ujemajo z lokalnimi minimumi signalov COMy oz. mase. Postopamo podobno kot v primeru iskanja faz dvojne opore (glej podglavje 3.4.2). Signale nog prilagodimo s t . i. mehko zamenjavo signalov kotov leve in desne stegnenice na najdenih lokacijah križanja. Mehko zamenjavo dosežemo z interpolacijo skozi manjše časovno okno v okolici točke križanja. Signale kotov golenic in lokacije kolenskih sklepov ter sklepov gležnjev naknadno prilagodimo glede na spremenjene vrednosti kotov stegnenic (slika 3.5).

3.4.5 Poobdelava signalov

Signali v pridobljenem biometričnem podpisu gibanja so občutljivi na raznovrstne napake v detekciji, ki se manifestirajo kot izstopajoče konice (angl. *spikes*) v signalu. Te konice je možno odkriti s pomočjo pomičnega medianinega filtra (angl. *moving medi-*



Slika 3.5

Križanje nog: slika a. prikazuje normaliziran signal mase (črna črta), začetke ciklov (črtkane rdeče vertikalne črte) in lokacije križanja nog (vijolične navpične črte), slika b. prikazuje sprednji (modra) in zadnji (zeleni) signal kotov stegenic, kot so bili pridobljeni iz skeleta, slika c. prikazuje signala kotov stegenic, transformirana z interpolacijo v predelu interpolacijskega okna (navpične črne črtkane črte) v signala kotov leve (modra) in desne (zeleni) stegenice.

an filter), ki ga skozi pomično okno apliciramo čez celoten signal. Pomični medianin filter torej omogoča lokalno detekcijo izstopajočih konic.

Najprej izračunamo razliko $d(t)$ med signalom $s(t)$ in pomično mediano $s_m(t)$. Nato poiščemo konice po naslednjem postopku:

$$d(t) = \|s(t) - s_m(t)\|$$

$$s(t) = \begin{cases} \text{NaN} & d(t) > \alpha * 1,4826 * s_{m_\sigma}(t) \\ s(t) & \text{sicer,} \end{cases} \quad (3.7)$$

kjer je $s(t)$ vrednost signala v času t , α je parameter, ki določa velikost konice, s_{m_σ} je medianina deviacija, ki je analogna standardni deviaciji, le da je v primeru mediane

izračunana na osnovi mediane namesto povprečja. Povečanje medianine deviacije za faktor 1,4826 privede do absolutne medianine deviacije, ki je v povprečju enaka standardni deviaciji v normalnih – Gaussovih porazdelitvah. Pri tako najdenih konicah postavimo signal na nedefinirano (NaN) vrednost, ki jo v nadaljevanju nadomestimo z interpoliranimi vrednostmi.

Na primer: v primeru signalov kotov kosti poiščemo konice za obe stegenici in obe golenici, nato manjkajoče nedefinirane vrednosti nadomestimo z interpoliranimi vrednostmi B-zlepkov. Uporabimo torej interpolacijo z B-zlepki, ki dobro deluje za sinusne signale, kot so naši. Po takšni spremembi je treba ustrezno pripraviti še druge strukture skeleta (kote in lokacije sklepov), ki so neposredno odvisni od kotov stegenic in golenic (npr. sprememba kota leve stegenice vpliva tudi na položaj leve golenice).

Dodatno signale filtriramo še s pomočjo trigonometričnih polinomov v fazi izgradnje opisnika značilik, kot je to opisano v podpoglavju 3.5.2.

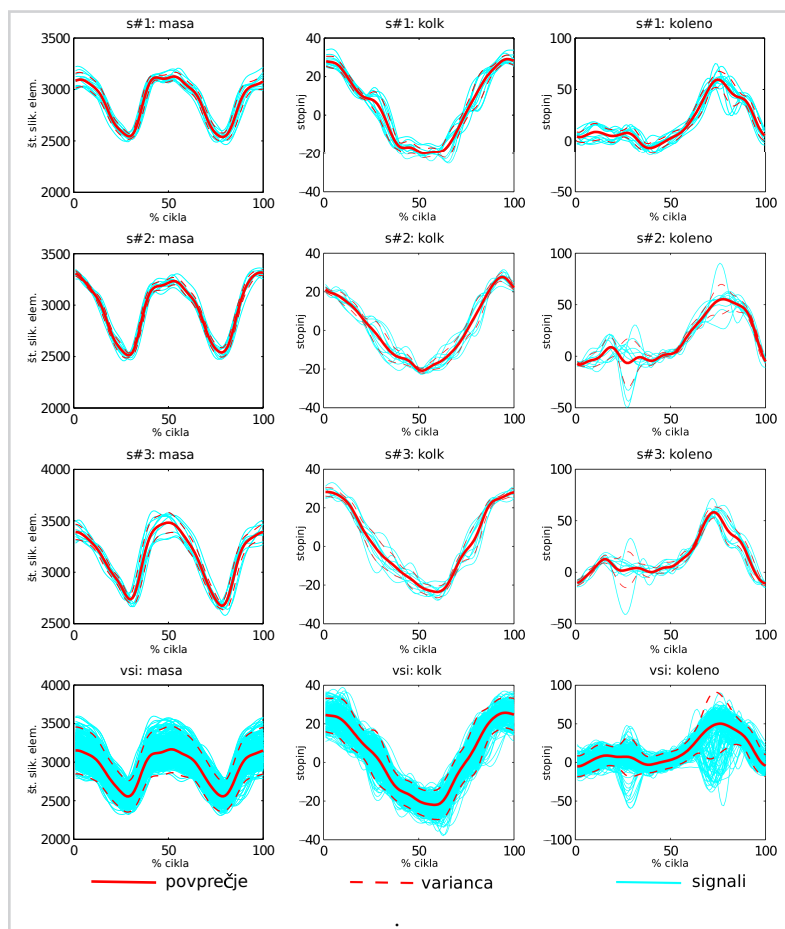
Ker je človeško gibanje ciklične narave, nato vse signale razrežemo na manjše dele glede na najdene začetke in dolžine ciklov. Cikličnost hoje implicira tudi predpostavko, da so cikli hoje istega subjekta med seboj podobni pri nespremenjenih pogojih opazovanja. Za pridobljene signale to pomeni, da se mora naslednji cikel začeti približno na istih vrednostih, kot se je končal prejšnji [14]. Zaradi prisotnosti šuma na različnih nivojih segmentacije pa ni vedno tako. Signale zato dodatno prilagodimo, da ustrezajo tej predpostavki. Naredimo jih ciklične tako, da poiščemo razliko med začetno in končno vrednostjo signala v ciklu ter interpoliramo vse vmesne vrednosti, da izničimo izračunano razliko:

$$S_i = S_i - \frac{2i - (L - 1)}{2(L - 1)} * D, \quad (3.8)$$

kjer je S del signala, ki ustreza določenemu ciklu hoje, S_i je vrednost v točki i -tega cikla, L je dolžina cikla in D je razlika med začetno ter končno vrednostjo signala v ciklu.

Nadalje, različni ljudje hodimo z različnimi hitrostmi, kar vpliva na frekvenco gibanja in s tem na trajanje ter dolžino ciklov hoje. Če je hoja zajeta z isto frekvenco vzorčenja, se dolžine ciklov razlikujejo v številu vzorcev, ki jih en cikel zajema. Da omogočimo primerjavo med cikli hoje različnih ljudi oz. istih ljudi, ki hodijo z različnimi hitrostmi, oz. zajemamo iste ljudi z različnimi frekvencami vzorčenja, je treba signale ciklov gibanja prevzorčiti na neko skupno, konstantno dolžino cikla. To dosežemo z interpolacijo in decimacijo signalov v ciklih.

Vsak cikel na koncu vsebuje le tiste dele originalnih signalov, ki natanko ustrezajo najdenim ciklom gibanja. Te dele v nadaljevanju imenujemo značilke (glej sliko 3.6), ki predstavljajo osnovo za nadaljnjo primerjavo subjektov. Vsaka značilka torej ustreza enemu ciklu določenega signala. Slika 3.6 dodatno demonstrira razliko med znotrajrazredno in medrazredno varianco. Opazimo lahko, da je prva občutno manjša od druge, kar je dober indikator zmožnosti razločevanja subjektov.



Slika 3.6

Primeri značilk: slika prikazuje značilke mase, kota kolka in kolena (v stolpcih) za več ciklov gibanja treh subjektov (vrstice od 1 do 3), označenih s s#1, s#2 and s#3. Modre linije ponazarjajo signale, rdeče odebeljene linije ponazarjajo povprečni signal vseh ciklov in rdeče črtkane linije ponazarjajo znotrajrazredno varianco. Spodnja vrstica prikazuje povprečje in varianco (medrazredno) vseh signalov vseh subjektov (vsi).

3.5 Klasifikacija

Klasifikacijo izvajamo na tri različne načine, podrobneje opisane v nadaljnjih podpodglavjih, bistvena razlika med njimi pa je v pojmovanju osnovne enota vzorca za klasifikacijo:

- Celostni pristop: vzorec za klasifikacijo je en cikel gibanja z opisom vseh značilk, ki so na voljo za ta cikel.
- Klasifikacija na osnovi fuzije značilk: vzorec za klasifikacijo je posamezna značilka oz. njen opis v določenem ciklu gibanja.
- Klasifikacija na osnovi okvirjev (angl. *frame*) videoposnetka: vzorec za klasifikacijo je en sam okvir gibanja oz. vrednost vseh značilk določenega cikla hoje v okvirju na določenem delu cikla.

Klasifikacijo izvedemo tako, da primerjamo testne primere z modeli, shranjenimi v galeriji podpisov, pri čemer značilke modelov in testnih primerov predhodno opišemo z opisniki gibanja. Primerjava pri celostnem pristopu in fuziji značilk temelji na evklidski razdalji med posameznimi značilkami oz. cikli gibanja, za kar je nujno treba zagotoviti, da so primerjani cikli med seboj popolnoma poravnani, saj že manjša odstopanja močno vplivajo na velikost razdalje. Delno dosežemo grobo poravnavo že v fazi detekcije ciklov, a je treba kljub temu zagotoviti natančnejše ujemanje začetka ciklov pred klasifikacijo. V primeru klasifikacije na osnovi okvirjev pa to omejitev zaobidemo, s čimer dosežemo neodvisnost klasifikacije od začetne točke cikla.

3.5.1 Fazna poravnava značilk v ciklih

Fazno poravnavo značilk v ciklih gibanja izvedemo tako, da poiščemo odmik dveh ciklov primerjanih subjektov na osnovi neke podmnožice izbranih značilk (COMy, masa in širina). Cikle hoje enega subjekta premikamo za neko omejeno število časovnih enot (okvirjev videoposnetka) levo in desno ter računamo razdalje med izbranimi značilkami za vse te premike. Izberemo tisti premik, ki privede do najmanjše razdalje,

nato pa vse značilke ciklov testnega primera ustrezno zamaknemo:

$$o = \arg \min_{i=-10}^{10} d(\vec{F}_1, \vec{F}_2) \quad (3.9)$$

$$\vec{F}_1 = [\vec{f}_{c_y,1}, \vec{f}_{m,1}, \vec{f}_{w,1}] \quad (3.10)$$

$$\vec{F}_2 = [\vec{f}_{c_y,2}, \vec{f}_{m,2}, \vec{f}_{w,2}] \quad (3.11)$$

kjer je o premik, d razdalja in \vec{F}_j vektor značilk subjekta j , sestavljen iz vektorjev $\vec{f}_{i,j}$ izbranih značilk. c_y, m, w označujejo COMy, maso in širino. Iskanje optimalnega premika je treba omejiti na majhen obseg (npr. $[-10..10]$), saj lahko večji premiki privedejo do pomika cikla hoje v naslednjo fazo dvojne opore, ki se začne z napačno nogo.

Pred klasifikacijo tako poravnamo vsak testni primer z modelom v galeriji, s katerim ga trenutno primerjamo.

3.5.2 Opisnik gibanja

Opisnike (angl. *descriptor*) gibanja zgradimo tako, da najprej vse signale ciklov gibanja aproksimiramo s trigonometričnimi polinomi, kar ima dvojni učinek. Na ta način signale filtriramo z nizko prepustnim filtrom, ki s signala odstrani visoke frekvence, ki običajno predstavljajo šum, hkrati pa dosežemo tudi zmanjšanje dimenzionalnosti vzorcev za klasifikacijo. Periodične signale lahko aproksimiramo kot:

$$S(x) = a_0 + \sum_{n=1}^N a_n \cos(nx) + i \sum_{n=1}^N b_n \sin(nx), \quad (3.12)$$

kjer so a_n in b_n koeficienti trigonometričnega polinoma, pridobljeni s pomočjo DFT. V našem primeru je dolžina signalov v enem ciklu 100, z $N = 10$ dobimo tako 21 koeficientov, ki jih zapišemo v t. i. opisnike značilk:

$$\vec{d}_i = [a_{0_i}, \dots, a_{N_i}, b_{1_i}, \dots, b_{N_i}]. \quad (3.13)$$

Opisnike značilk nadalje združimo v t. i. opisnik cikla, ki vsebuje zapis celotnega cikla hoje:

$$\vec{D}_{c_j} = [\vec{d}_1, \vec{d}_2, \dots, \vec{d}_{N_z}], \quad (3.14)$$

kjer je \vec{D}_{c_j} opisnik c_j -tega cikla hoje določenega subjekta, $\vec{d}_1 \dots \vec{d}_{N_z}$ pa so opisniki vseh N_z različnih značilik, kar je v našem primeru 64.

Tako pridobljeni opisniki ciklov so še vedno preveliki za klasifikacijo. V našem primeru s 64 značilkami, kjer vsaka vsebuje 21 koeficientov, je dolžina takšnega opisnika enaka 1344. Dimenzionalnost zato dodatno zmanjšamo s pomočjo analize glavnih komponent (angl. *Principal Component Analysis – PCA*) tako, da ohranimo približno 99 % variance, kar dosežemo s približno 100 koeficienti PCA.

Dodatno obdelamo še vse opisnike z linearno razločevalno analizo (angl. *Linear Discriminant Analysis – LDA*), da povečamo razdalje med različnimi modeli v galeriji, s čimer omogočimo boljše razločevanje.

Opisniki ciklov so osnova za celostni pristop klasifikacije. V primeru fuzije značilik je vzorec za klasifikacijo ena značilka znotraj nekega cikla, zato v tem primeru uporabimo opisnike značilik (21 koeficientov), ki jih obdelamo le z LDA, saj dodatno zmanjšanje dimenzionalnosti s PCA v tem primeru ni potrebno.

V primeru klasifikacije na osnovi okvirjev sestavimo opisnik iz vrednosti vseh značilik v določenem okvirju cikla gibanja. Tako dobimo t. i. opisnik okvirja:

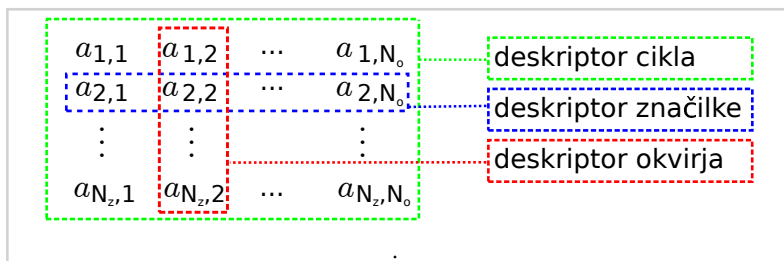
$$\vec{d}_o = [a_{1_0}, a_{2_0}, \dots, a_{N_z_0}], \quad (3.15)$$

kjer o ponazarja določen okvir v ciklu, $a_i, i \in 1..N_z$ pa vrednosti posameznih značilik v istem okvirju ($N_z = 64$). Opisnike naknadno obdelamo z LDA, s čimer zmanjšamo dimenzionalnost in povečamo razdalje med vzorci različnih subjektov.

Slika 3.7 ponazarja vse tri tipe opisnikov za tri različne načine klasifikacije, ki so podrobneje razloženi v nadaljevanju.

Slika 3.7

Opisniki gibanja: slika prikazuje tri tipe opisnikov gibanja, ki jih uporabimo za tri različne načine klasifikacije: opisnik cikla (zelena), opisnik značilke (modra) in opisnik okvirja (rdeča). Matrika vsebuje v vrsticah različne značilke enega cikla gibanja, stolpci pa ponazarjajo vrednosti teh značilik v posameznih okvirjih cikla.



3.5.3 Celostni pristop klasifikacije

V nadaljevanju izračunamo razdalje med opisniki gibanja testnih primerov in opisniki gibanja vseh modelov v galeriji. Razdaljo med dvema subjektoma definiramo kot:

$$D(p, g_m) = f_c \left(d(\vec{p}_i, \vec{g}_{m_j}) \right), \quad (3.16)$$

$i \in |P|, j \in |G_m|$

kjer p in g_m določata testni primer in model galerije, \vec{p}_i in \vec{g}_{m_j} pa določata opisnika iz množic opisnikov P in G_m posameznih subjektov. $d(\cdot)$ označuje funkcijo evklidske razdalje med dvema opisnikoma. f_c je kriterijska funkcija, ki določa, kako izberemo končno razdaljo med subjektoma izmed vseh razdalj, ki jih pridobimo z razdaljami med posameznimi cikli. Najpogosteje se za to uporabljajo povprečje, najmanjša razdalja in mediana. V našem primeru smo izbrali mediano, ki do neke mere odstrani vpliv najbolj deviantnih razdalj (izstopajočih vrednosti (angl. *outliers*)), ki se lahko pojavijo kot rezultat napak v segmentaciji. Nato izračunamo $D(p, g_m)$ za vse modele v galeriji in tako dobimo vektor razdalj:

$$\vec{D}_p = [D(p, g_1), \dots, D(p, g_M)], \quad (3.17)$$

kjer je M število modelov v galeriji. Končno izberemo model iz galerije z minimalno razdaljo.

3.5.4 Klasifikacija na osnovi fuzije značilk

Različne značilke imajo različno moč razločevanja, zato lahko predpostavimo, da bi lahko s kontrolo vpliva, koliko naj neka značilka doprinese h končnemu rezultatu klasifikacije, izboljšali rezultate klasifikacije. V ta namen vpeljemo fuzijo značilk, kjer vsaka značilka predstavlja ločen klasifikator, končni rezultat pa izračunamo tako, da s fuzijo združimo rezultate posameznih klasifikatorjev [27]. Takšen postopek ima v prvi meri pozitiven vpliv zaradi velikega števila klasifikatorjev, ki jih združujemo, dodatno pa pozitivno vpliva v primeru manjkajočih oz. napačno izmerjenih značilk. V primeru prekrivanj je na primer poškodovanih le nekaj značilk (manjkajoči deli nog ipd.), negativne posledice pa lahko zmanjšamo s klasifikacijo na osnovi preostalih nepoškodovanih značilk, ki v končnem rezultatu preglasujejo problematične.

Postopamo tako, da najprej izvedemo klasifikacijo na osnovi razdalje za vsako značilko posebej, torej na osnovi opisnikov značilk, namesto opisnikov ciklov, kot smo

postopali v primeru celostnega pristopa klasifikacije. Zmanjšanje dimenzionalnosti s PCA v tem primeru ni potrebno, saj že trigonometrični polinomi zagotavljajo dovolj majhne opisnike za uporabo LDA. Vsaka značilka kot rezultat klasifikacije priskrbi vektor razdalj:

$$\vec{D}_{p,z} = [D_z(p, g_1), \dots, D(p, g_M)], \quad (3.18)$$

kjer $D_z(p, g_m)$ označuje razdaljo med primerom p in modelom g_m glede na značilko z . Vektorje razdalj vseh značilk $\vec{D}_{p,z}$ najprej normaliziramo na interval $[0, 1]$, da odstranimo pristranskost posameznih značilk zaradi razlik v enotah merjenja (razdalje med koti so drugačne kot razdalje med masami ali koordinatami) s pomočjo t. i. normalizacije min-max, nato pa s fuzijo združimo vektorje razdalj vseh značilk v skupni vektor:

$$\vec{S}_p = \frac{1}{Z} \sum_{z=1}^Z (w_z \|\vec{D}_{p,z}\|_{\min-\max}), \quad (3.19)$$

kjer je Z število značilk, w_z pa so uteži posameznih značilk. $\|\vec{x}\|_{\min-\max}$ označuje normalizacijo vektorja \vec{x} na interval $[0, 1]$, katerega elemente izračunamo kot:

$$x'_i = \frac{x_i - \min(\vec{x})}{\max(\vec{x}) - \min(\vec{x})}. \quad (3.20)$$

Končno lahko izmed vseh razdalj v vektorju \vec{S}_p poiščemo minimalno, ki določa ujemanje z modelom v galeriji.

V tem delu smo pridobili uteži eksperimentalno z oceno razločevalne sposobnosti posameznih značilk, ki smo jo ocenili na osnovi razmerja med znotrajrazredno in medrazredno varianco. Razred je v tem primeru subjekt. Vse signale \vec{s}_i značilke z , ki ustrezajo ciklom subjekta s , zapišemo v matriko S_z :

$$S_z = \begin{bmatrix} \vec{s}_1 \\ \vec{s}_2 \\ \vdots \\ \vec{s}_c \end{bmatrix}, \quad (3.21)$$

kjer je c število ciklov subjekta s . Znotrajrazredna varianca $\sigma_{s,z}^2(S_z)$ je varianca vseh signalov (vrstice matrike S_z) določene značilke z določenega subjekta s . Medrazredna varianca značilk σ_z^2 pa je izmerjena na vseh signalih vseh subjektov za isto značilko z .

Primeri varianc so prikazani na sliki 3.6. Razmerje med povprečno znotrajrazredno varianco značilnik in medrazredno varianco:

$$r_{\sigma} = \frac{\bar{\sigma}_{s,z}^2}{\sigma_z^2} \quad (3.22)$$

je indikator razločevalne moči značilke: manjše kot je to razmerje, večja je razločevalna moč značilke.

Nedvomno bi bilo možno določiti te uteži na kakšen bolj sofisticiran način z uporabo strojnega učenja, a se je izkazalo, da so rezultati klasifikacije relativno neobčutljivi na manjše variacije v utežeh, zato je izbrana metoda določanja uteži dovolj dobra za demonstracijo vpliva posameznih značilnik na rezultate razpoznavanja. Za demonstracijo učinka večjega števila klasifikatorjev na rezultate razpoznavanja pa je povsem legitimno določiti uteži enakomerno, tako da so vse značilke enakovredne pri postopku fuzije.

3.5.5 Klasifikacija na osnovi okvirjev

Enega večjih problemov zgoraj opisanega postopka predstavlja primerjava signalov značilnik na osnovi razdalje. Signali morajo biti popolno poravnani glede na gibalni cikel, da je lahko takšna primerjava uspešna. Poravnavo dosegamo na različnih nivojih metode. Prvič naredimo grobo poravnavo že v fazi iskanja ciklov (glej podpoglavje 3.4.3), kjer pazimo predvsem, da se cikli različnih posnetkov hoje začnejo z isto nogo. Drugič naredimo fino poravnavo faze gibalnih ciklov tik pred primerjavo (glej podpoglavje 3.5.1). Prva poravnava zagotavlja, da vse cikle hoje primerjamo med seboj pravilno, ko je izbrana noga vedno na začetku cikla, saj se dela cikla z različnima nogama spredaj (korak z levo, korak z desno) med seboj razlikujeta tako zaradi perspektivne projekcije, kot zaradi samih lastnosti hoje. Napačna detekcija sprednje noge lahko poveča razdaljo med cikli hoje istega subjekta ali zmanjša razdaljo med cikli hoje različnih subjektov. Druga poravnava sicer zmanjša razdaljo med cikli hoje istih subjektov, a hkrati zmanjša tudi razdaljo med cikli različnih subjektov. Napake pri obeh primerjavah niso redkost, še posebej v primeru silhuet slabše kakovosti, in lahko negativno vplivajo na rezultate razpoznavanja.

Da zaobidemo naštetе težave s poravnavo, smo izvedli klasifikacijo na način, ki je neodvisen od trenutne faze cikla hoje. V grobem gre za to, da namesto klasifikacije na osnovi razdalje med signali značilnik izvajamo klasifikacijo na osnovi vzorcev signalov v posameznih okvirjih (angl. *frame*) videoposnetka. Vzorce za klasifikacijo v tem primeru predstavljajo trenutne vrednosti vseh signalov značilnik na določenem delu gibalnega

cikla. Označimo z $\vec{A}_i = [a_{i_1}, a_{i_2}, \dots, a_{i_n}]$ vektor, ki opisuje signal značilke i za en cikel subjekta A . n predstavlja število vzorcev v enem ciklu hoje. Signale vseh preostalih značilk zapišemo v vrstice matrice \mathbf{A} :

$$\mathbf{A} = \begin{bmatrix} \vec{A}_1 \\ \vec{A}_2 \\ \vdots \\ \vec{A}_m \end{bmatrix} = \begin{bmatrix} a_{1_1}, a_{1_2}, \dots, a_{1_n} \\ a_{2_1}, a_{2_2}, \dots, a_{2_n} \\ \vdots \\ a_{m_1}, a_{m_2}, \dots, a_{m_n} \end{bmatrix}, \quad (3.23)$$

kjer m predstavlja število značilk. Vsak posamezni stolpec takšne matrice je sestavljen iz vrednosti $[a_{1_j}, a_{2_j}, \dots, a_{m_j}]$ vseh značilk na določenem delu cikla in hkrati predstavlja vzorec za klasifikacijo. Če imamo vzorčenje signalov npr. $n = 100$ na cikel, dobimo za vsak cikel 100 vzorcev za klasifikacijo. Vzorce ciklov preostalih subjektov nanizamo v matriko \mathbf{X} :

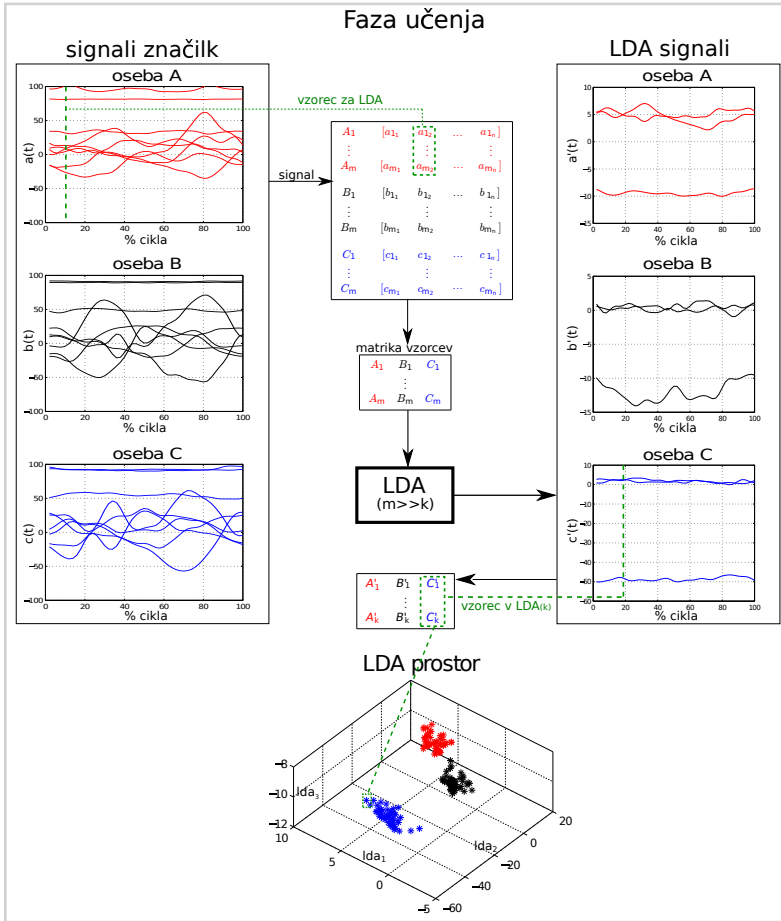
$$\mathbf{X} = \begin{bmatrix} \mathbf{A} & \mathbf{B} & \mathbf{C} & \dots \end{bmatrix} = \begin{bmatrix} \vec{A}_1 & \vec{B}_1 & \vec{C}_1 & \dots \\ \vec{A}_2 & \vec{B}_2 & \vec{C}_2 & \dots \\ \vdots & \vdots & \vdots & \dots \\ \vec{A}_m & \vec{B}_m & \vec{C}_m & \dots \end{bmatrix}, \quad (3.24)$$

kjer sta \mathbf{B} in \mathbf{C} podobno zgrajeni matriki kot \mathbf{A} , le da pripadata subjektoma B in C . Tako dobljeno matriko preslikamo z LDA, s pomočjo katerega zmanjšamo število značilk m posameznega cikla hoje na k in hkrati zagotovimo največji razmik med vzorci v novem prostoru LDA $_k$. Dobimo končno matriko \mathbf{X}' :

$$\mathbf{X}' = \begin{bmatrix} \vec{A}'_1 & \vec{B}'_1 & \vec{C}'_1 & \dots \\ \vec{A}'_2 & \vec{B}'_2 & \vec{C}'_2 & \dots \\ \vdots & \vdots & \vdots & \dots \\ \vec{A}'_k & \vec{B}'_k & \vec{C}'_k & \dots \end{bmatrix}, \quad (3.25)$$

kjer elementi predstavljajo preslikane vrednosti v manj dimenzionalnem ($k \ll m$) prostoru LDA. V našem primeru 64 značilk preslikamo v prostor LDA s 24 transformiranimi značilkami. Opisani postopek prikazuje slika 3.8.

Končno lahko izvajamo klasifikacijo s pomočjo algoritma K -najbližjih sosedov (angl. *K nearest neighbours* – *KNN*) na nevidnih vzorcih hoje testnih primerov, ki jih

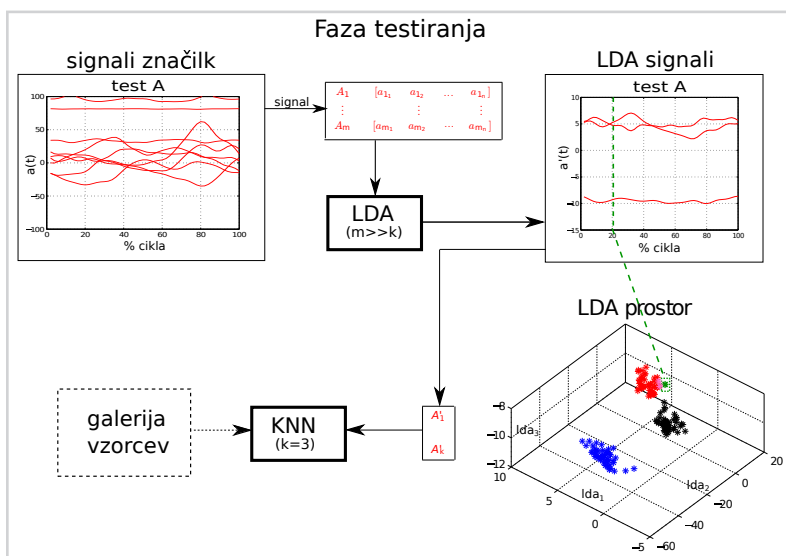


Slika 3.8

Klasifikacija na osnovi vzorcev okvirjev video-posnetka. Slika prikazuje postopek izgradnje galerije učnih primerov.

predhodno podobno obdelamo s pridobljenim modelom LDA_k (slika 3.9). Maksimalna še možna izbira za k je omejena s številom okvirjev v ciklu gibanja, vendar se v praksi izkaže, da takšna izbira potegne v oceno preveč napačnih sosedov. KNN pri manjših vrednosti k daje optimalnejše rezultate, zato smo v našem primeru uporabili $k = 3$.

Rezultati klasifikacije so torej razredi R_i , določeni vsakemu vzorcu \vec{a}_i testnega prime-



Slika 3.9

Klasifikacija na osnovi vzorcev okvirjev video-posnetka. Slika prikazuje postopek klasifikacije testnega primera.

ra, kjer i predstavlja okvir slike na določenem delu cikla. Rezultate klasifikacije vzorcev nato hierarhično združimo v višje enote. Najprej s pomočjo večinskega glasovanja določimo razrede posameznih ciklov hoje, tako da določimo razred cikla kot tisti razred, ki je bil največkrat izbran pri klasifikaciji posameznih vzorcev tega cikla:

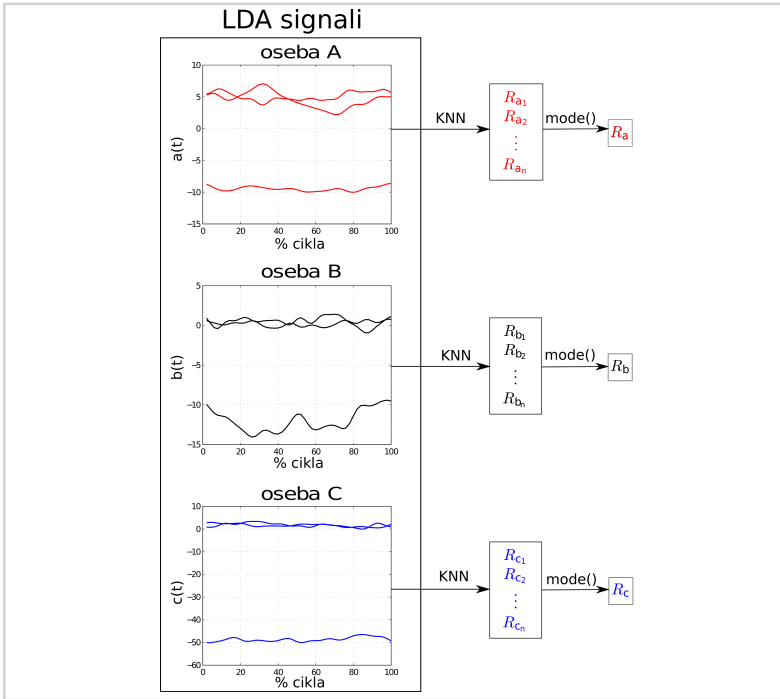
$$R_c = \text{mode}(\vec{R}), \quad (3.26)$$

kjer R_c predstavlja razred celotnega cikla, funkcija mode pa izbere vrednost, ki je v predloženem vektorju najbolj pogosta (slika 3.10). Nadalje je vsak posnetek hoje sestavljen iz več ciklov, razred posnetka hoje R_p pa določimo kot večinski razred vseh ciklov tega posnetka.

Ker klasificiramo vse okvirje posameznega cikla hoje posebej, je vseeno, s katerim okvirjem začnemo klasifikacijo. S tem dosežemo neodvisnost od začetne točke cikla hoje.

3.5.6 Dekompozicija signalov z valčki

Transformacija z valčki (angl. *wavelet transform* – *WT*) je zadnja leta postala popularna tehnika za analizo signalov in slik. Prav tako se pojavlja na področju analize gibanja



Slika 3.10

Hierarhično združevanje rezultatov klasifikacije: slika prikazuje klasifikacijo enega cikla gibanja (dolžine n okvirjev) oseb A, B in C. Rezultate klasifikacije osebe x na nivoju okvirjev predstavljajo razredi R_{x_1}, \dots, R_{x_n} , s katerimi s pomočjo večinskega glasovanja ($\text{mode}()$) določimo razred celotnega cikla R_x .

[70–72], saj podpira analizo časovno-frekvenčnih signalov na različnih ločljivostih s pomočjo lokalne frekvenčne analize, ki omogoča lokalizacijo frekvenc tako v časovnem kot v frekvenčnem prostoru, s tem pa ne predpostavlja stacionarnosti signalov [70]. Nadaljnje predstavljajo pravokotni materni valčki (angl. *mother wavelet*), matematično orodje za dekompozicijo funkcije v večločljivostno hierarhijo različnih lokaliziranih frekvenčnih kanalov, ki omogoča analizo na različnih nivojih ločljivosti, kar je primerno za strategije razpoznavanja objektov, ki delujejo po principu od grobega k finemu (angl. *coarse to fine*), ki so zelo podobne tistim, ki jih uporabljajo naši možgani in človeški vizualni sistem [71].

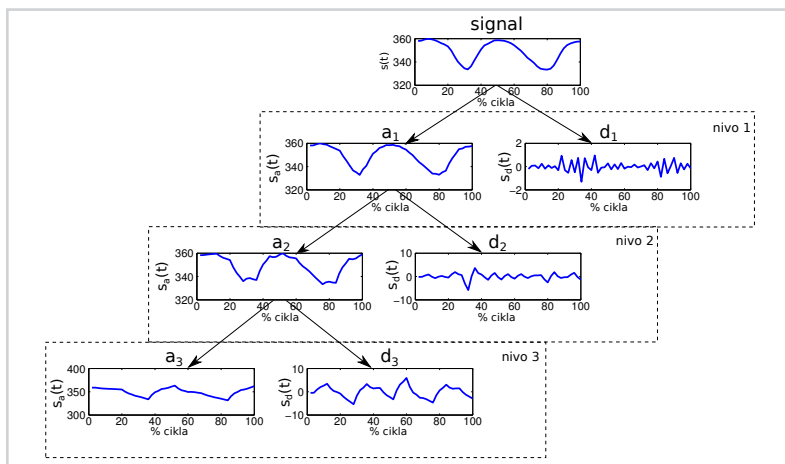
V našem primeru analiza z valčki omogoča preiskovanje signalov značilik na različnih ločljivostih frekvenčnega prostora, kar ima podoben učinek, kot če bi signal filtrirali z množico različnih pasovno prepustnih filtrov. Motivacija za uporabo valčkov izhaja iz

potrebe po učenju, ki jo zahteva transformacija prostora značilnk. Z uporabo valčkov se torej nadejamo poiskati tisti del frekvenčnega spektra, ki najbolj ohranja razločevalno moč značilnk tudi pri spremenjenih pogojih, kot je primer spremembe hitrosti hoje.

Postopamo tako, da izračunamo transformacijo z valčki za posamezne signale na nivojih L . Število nivojev dekompozicije je odvisno od dolžine signala, ker ima v našem primeru cikel 100 okvirjev, je smiselna dekompozicija do nivoja 6. Pri dekompoziciji smo uporabili *Daubechiesove* matrne valčke z dvema bežiščnima momentoma (angl. *vanishing moment*) db2, saj so se ti že izkazali kot primerni za uporabo na podobnih signalih, kot so naši [70]. Na vsakem nivoju signal razbijemo na dva dela, glede na izbrani materni valček, in tako pridobimo dvojke koeficientov. Prvi del ustreza koeficientom aproksimacije, drugi pa detajlom. Signale nato rekonstruiramo na vsakem od nivojev (glej sliko 3.11) le s pomočjo koeficientov aproksimacije in za vsak nivo posebej izvedemo poprej opisano klasifikacijo na osnovi okvirjev. S pomočjo takšne dekompozicije signalov poskušamo najti značilnosti gibanja (signalov), ki so najbolj razločevalni za posameznika in najmanj podvrženi vplivom dejavnikov, ki spreminjajo gibanje. Izkaže se (glej podpoglavje 5.4), da največjo učinkovitost filtriranja dosežemo pri nivoju 4, zato v nadaljevanju uporabljamo ta nivo aproksimacije.

Slika 3.11

Dekompozicija signala s pomočjo valčkov. Slika prikazuje dekompozicijo osnovnega signala z valčki na različnih nivojih: nivo 1, nivo 2 in nivo 3. Na vsakem nivoju poiščemo koeficiente za aproksimacijo in detajle ter izvedemo rekonstrukcijo signala s temi koeficienti. Rezultat je rekonstruiran signal nivoja l : del z aproksimacijo: a_l , del z detajli: d_l .



Sprememba hitrosti hoje

4.1 Uvod

Spremembe hitrosti hoje predstavljajo enega glavnih faktorjev vpliva z negativnim učinkom na rezultate razpoznavanja. Ko človek spremeni hitrost hoje, se spremeni sama dinamika gibanja, s tem pa tudi značilke, ki to dinamiko opisujejo (npr. dolžina koraka, koti sklepov ipd.). Statične značilke, ki niso toliko odvisne od same dinamike gibanja (npr. dolžine kosti), pa ostanejo nespremenjene.

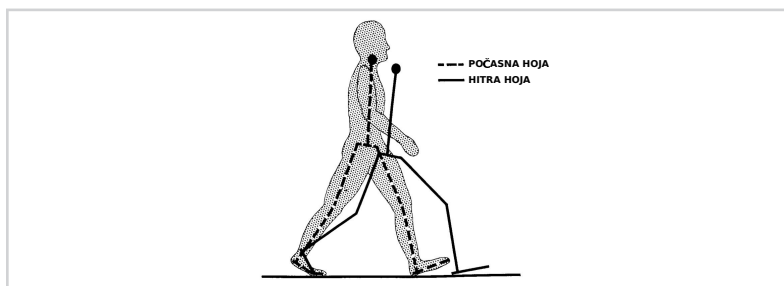
4.1.1 Časovna normalizacija cikla gibanja

Prva neposredna posledica takšne spremembe se odraža na trajanju cikla gibanja. Hitreje kot hodimo, krajši čas je potreben za izvedbo enega cikla. Metode za razpoznavanje ljudi po načinu hoje morajo ta učinek upoštevati, da omogočijo primerjavo subjektov, ki hodijo z različnimi hitrostmi. To lahko storimo implicitno z uporabo cikla hoje kot enega časovno-prostorskega vzorca, kot je to v primeru povprečnih silhuet [12], kjer združimo vse silhete celotnega cikla gibanja v eno samo povprečno silhueto. Drugo možnost predstavlja uporaba ključnih okvirjev (angl. *key frames*), kjer uporabimo za razpoznavanje le nekatere določene ključne okvirje videoposnetka, ki predstavljajo specifične faze cikla gibanja, vse ostale pa ignoriramo. Avtorji v [24, 48] so npr. uporabili le nekaj silhuet iz faze dvojne in enojne opore za predstavitev celotnega cikla hoje. Tretja možnost pa je eksplicitna prilagoditev na spremembo trajanja cikla. Avtorji v [47] uporabljajo t. i. normalizacijo dinamike gibanja, avtorji v [49] uporabljajo dinamično ukrivljanje časa (angl. *dynamic time warping*), avtorji v [14] pa uporabljajo časovno prevzorčenje (angl. *temporal resampling*) s pomočjo interpolacije trajektorij sledenih točk na isto dolžino.

4.2 Normalizacija hitrosti hoje

Večina zgoraj omenjenih metod ne upošteva antropometričnih sprememb, ki jih povzročijo spremembe hitrosti hoje. Te se v prvi meri odražajo v spremenjeni dolžini koraka [66]. V primeru hitre hoje delamo nekoliko daljše korake, v primeru počasne hoje pa nekoliko krajše (slika 4.1). Te spremembe lahko obravnavamo npr. z invariantnimi značilkami [49] ali s transformacijami skaliranja kot npr. avtorji v [24], ki so normalizirali okvirje silhuet dvojne opore v predelu nog in s tem dosegli prilagoditev na ciljno hitrost hoje.

V sklopu tega dela smo integrirali zgornjo idejo kot osnovo za normalizacijo hitrosti

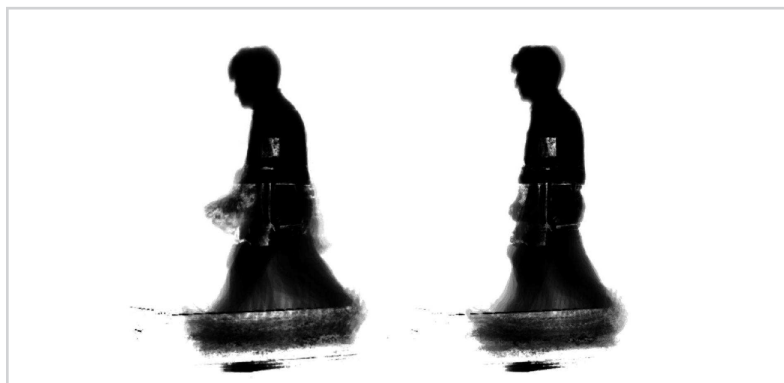


Slika 4.1

Sprememba hitrosti hoje: antropometrične razlike med hitro in počasno hojo [66].

hoje z geometrijskimi transformacijami [73] v dve metodi za razpoznavanje gibanja, ki sicer ne obravnavata spremembe hitrosti hoje. Za razliko od avtorjev v [24], ki so s pomočjo zgornje ideje transformirali pet ključnih silhuet gibalnega cikla z osnovnim (angl. *baseline*) [12] algoritmom, smo v našem delu prikazali, kako lahko podobno idejo posplošimo na celoten cikel hoje z metodo povprečnih silhuet [11] ter celo na metode, ki ne temeljijo na silhuetah, kot je verjetnostni prostorsko-časovni model na osnovi trajektorij sledenih točk [14].

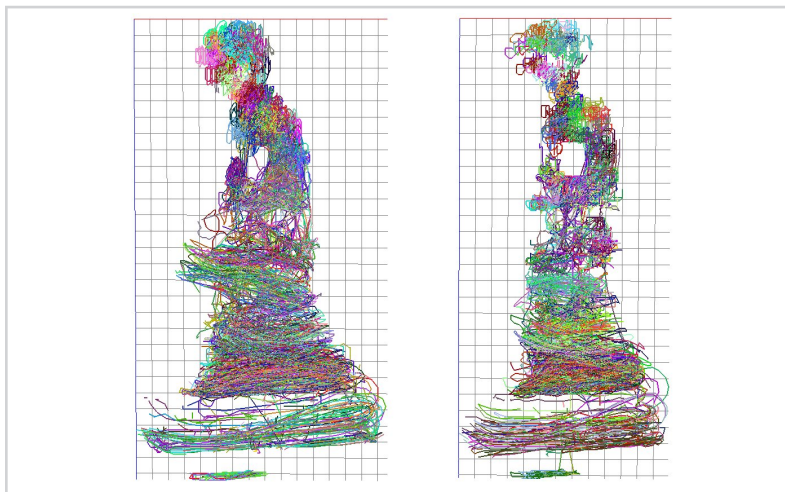
Pri metodi povprečnih silhuet je možna integracija ideje z ustreznim prilagajanjem širine silhete na ciljno hitrost. Ker je dolžina koraka sorazmerna z maksimalno širino silhete v predelu nog, lahko povprečne silhete skrčimo oz. razširimo za neki določen faktor, da jih prilagodimo na ciljno hitrost hoje (slika 4.2). Takšna normalizacija omogoča boljšo primerjavo silhuet subjektov, ki hodijo z različnimi hitrostmi.



Slika 4.2

Sprememba hitrosti hoje: primer stranskega pogleda povprečnih silhuet za hitro (levo) in počasno (desno) hojo.

Nadalje smo tovrstno normalizacijo prilagodili za verjetnostni prostorsko-časovni model, kjer smo dosegli podobni učinek s transformacijo krivulj trajektorij (slika 4.3). Postopamo tako, da obliko krivulje transformiramo v smeri hoje, ki sovpada z orienta-



Slika 4.3

Sprememba hitrosti hoje: primer stranskega pogleda trajektorij sledenih točk za hitro (levo) in počasno (desno) hojo.

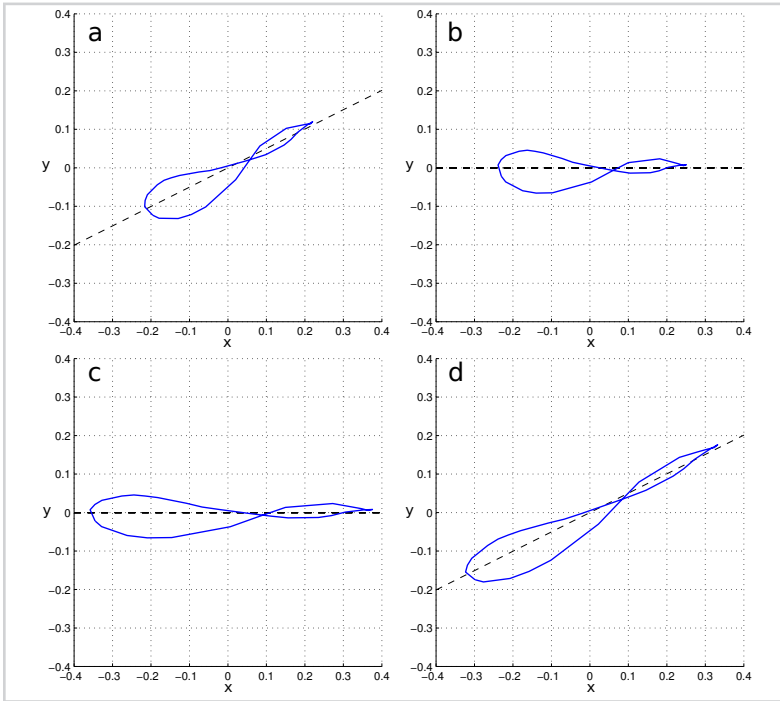
cijo glavnega lastnega vektorja analize glavnih komponent vseh trajektorij določenega cikla hoje. Ker je oblika krivulje opisana z razdaljami točk krivulje od centroida, dosežemo normalizacijo s pomočjo t . i. faktorja spremembe hitrosti, ki ga predhodno ocenimo s pomočjo razlike med maksimalnimi širinami krivulj dveh subjektov v predelu nog, ki jih primerjamo. Da lahko takšno transformacijo uporabimo za poljubno smer hoje (v kontrast stranskemu pogledu, ki ga običajno uporabljamo), je treba krivulje najprej rotirati, da sovpadajo z osjo x , šele nato povečati za faktor spremembe hitrosti in na koncu rotirati nazaj v začetno orientacijo (slika 4.4).

Posamezne točke trajektorije \vec{x}_i spremenimo z geometrično transformacijo:

$$\vec{x}'_i = \mathbf{R}^T \mathbf{S} \mathbf{R} \vec{x}_i, \quad (4.1)$$

$$\mathbf{R} = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}, \quad (4.2)$$

$$\mathbf{S} = \begin{bmatrix} s & 0 \\ 0 & 1 \end{bmatrix}, \quad (4.3)$$



Slika 4.4

Transformacija trajektorij:
 a. prikazuje trajektorijo v smeri hoje ϕ , b. prikazuje rotirano trajektorijo za kot $-\phi$, c. prikazuje skalirano trajektorijo vzdolž osi x , d. prikazuje skalirano trajektorijo, rotirano nazaj v prvotno smer hoje ϕ .

kjer je \mathbf{R} rotacija za kot ϕ , \mathbf{S} skalirna matrika in s faktor spremembe hitrosti. V primeru stranskega pogleda (slika 4.3) je smer hoje že poravnana z osjo x , zato je ϕ v tem primeru 0.

4.3 Transformacija prostora značilk

Idejo, opisano v podpoglavju 4.2, smo posplošili za uporabo na značilkah, pridobljenih s pomočjo metode na osnovi človeškega skeleta. V tem primeru predpostavimo, da če se spremembe hitrosti hoje odražajo na značilkah, kot sta dolžina koraka in sprememba kotov sklepov, se morajo te spremembe na podoben način bolj ali manj odražati tudi na vseh ostalih značilkah, pridobljenih iz skeleta telesa. Iz zgornje vrstice slike 4.5 je razvidno, kako se značilke istega subjekta razlikujejo za različne hitrosti hoje. Če opazujemo značilke skupine subjektov za hitro in skupine subjektov za počasno hoje v

drugi vrstici slike 4.5, lahko opazimo, kako so v primeru hitre hoje značilke nekoliko višje kot v primeru počasne hoje, prav tako se do neke mere spremeni tudi razpršenost teh značilk. Velikost sprememb pa ni enotna skozi celoten cikel gibanja, ampak je odvisna od faze cikla, v kateri trenutno smo. To opažanje nakazuje, da je mogoče pretvoriti značilke ene skupine v drugo z uporabo povprečij in standardnih deviacij značilk za posamezne skupine subjektov, ki hodijo z različnimi hitrostmi.

Izvedemo transformacijo prostora značilk na osnovi statičnega modeliranja prostora značilk za različne hitrosti hoje. Za takšno modeliranje potrebujemo predhodno znanje, ki ga pridobimo s pomočjo več skupin, ki vsebujejo primerke hoje različnih hitrosti. Vsaka skupina vsebuje hoje z določeno hitrostjo. Postopamo tako, da izračunamo glavne statistične momente za vse skupine z različnimi hitrostmi hoje. Izračunamo povprečje, standardno deviacijo in koeficient poševnosti za vsako hitrost hoje (2. vrstica slike 4.5) in s pomočjo spodnje linearne transformacije [74] transformiramo prostor značilk ene hitrosti hoje v prostor značilk druge hitrosti, kar omogoča primerjavo subjektov, ki hodijo z različnimi hitrostmi:

$$\vec{s} = \frac{\vec{s} - \vec{\mu}_g}{\vec{\sigma}_g}, \quad (4.4)$$

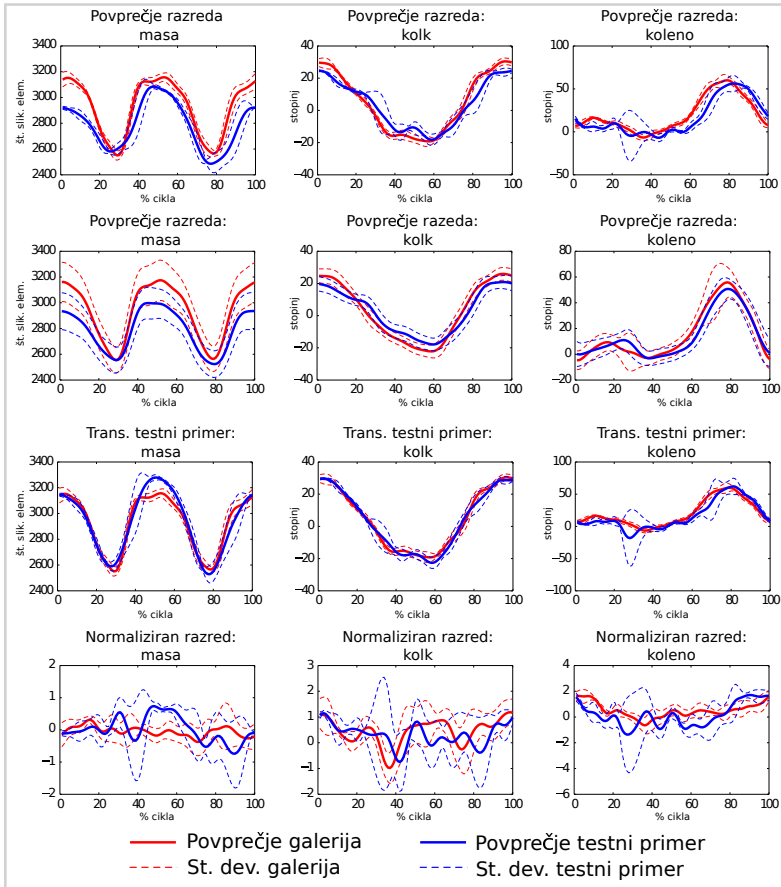
kjer je \vec{s} vektor signala določenega cikla neke značilke, $\vec{\mu}_g$ je povprečni vektor iste značilke za ciljno hitrost hoje in $\vec{\sigma}_g$ je vektor standardnih deviacij iste značilke za isto skupino. Pri tem smo predpostavili, da so vrednosti značilk simetrično porazdeljene okoli povprečja. Simetričnost porazdelitve lahko ugotovimo s pomočjo koeficienta poševnosti:

$$\vec{\gamma} = \frac{\frac{1}{N} \sum_{i=1}^N (\vec{s}_i - \vec{\mu}_g)^3}{\vec{\sigma}_g^3}, \quad (4.5)$$

kjer je N število primerov značilk \vec{s}_i enega subjekta. Bolj kot so vrednosti $\vec{\gamma}$ oddaljene od 0, bolj je porazdelitev nesimetrična [74]. V našem primeru se sicer izkaže, da vrednosti niso popolnoma simetrično porazdeljene okoli povprečja, a nesimetričnost porazdelitve kljub temu nima signifikantnega vpliva na rezultate transformacije, zato je v tem primeru zgornja poenostavitev sprejemljiva.

Takšna transformacija odstrani iz značilk glavne deviacije prostora skupine, ki predstavljajo določeno hitrost hoje, in ohrani le razlike med posameznikom ter povprečjem skupine, na osnovi katerih lahko razločujemo posameznike (slika 4.5, četrta vrstica).

Transformacijo izvedemo tako na galeriji kot na testnih primerkih, da dobimo normalizirano obliko, kar omogoča njuno neposredno primerjavo. Demonstracijo primera preslikave počasne hoje v hitro prikazuje slika 4.5, kjer smo s pomočjo opisanega postopka najprej preslikali počasno hojo v normalizirano obliko, nato pa z inverzno transformacijo preslikali značilke v prostor hitre hoje. V tretji vrstici slike 4.5 vidimo, da so značilke po transformaciji bolj primerljive.



Slika 4.5

Transformacija prostora značilke: prva vrstica prikazuje značilke istega subjekta za hitro (rdeča) in počasno (modra) hojo. Druga vrstica prikazuje srednje vrednosti celotne skupine za hitro in počasno hojo, tretja vrstica prikazuje preslikane značilke istega subjekta iz počasne hoje v hitro, četrta vrstica pa prikazuje normalizirano obliko značilke istega subjekta, potem ko smo jim odšteli skupinske elemente. Os y v prvih treh vrsticah vsebuje absolutne (izmerjene) vrednosti značilke, medtem ko četrta vrstica prikazuje razlike po normalizaciji, ki so relativno majhne v primerjavi z izmerjenimi vrednostmi. Os x označuje čas kot odstotek dela cikla hoje.



Rezultati in analiza

5.1 *Evalvacija normalizacije hitrosti hoje z geometrijskimi transformacijami*

Učinke normalizacije hitrosti hoje s pomočjo geometrijskih transformacij, opisane v podpoglavju 4.2, smo preverili s pomočjo zbirke Mobo [58], ki vsebuje 25 oseb, ki hodijo po tekoči stezi. Zajete so iz 6 različnih pogledov in vsebujejo 4 različne načine hoje (počasna hoja – 3,3 km/h, hitra hoja – 4,5 km/h, hoja v klanec in nošenje predmeta). Za dodatne parametre glej tabelo 2.1.

Izvedli smo več eksperimentov za ilustracijo razlik v učinkovitosti razpoznavanja pri enakih in spremenjenih pogojih hitrosti hoje za modele v galeriji ter za testne primere [73]. Zaradi pomanjkanja ločenih posnetkov za galerijo in testne primere smo v primeru eksperimentov z enakimi hitrostmi hoje razdelili posnetke na polovico, pri čemer smo eno polovico uporabili za galerijo, drugo pa za testiranje, medtem ko pri spremenjenih pogojih takšna delitev ni bila potrebna, saj gre pri različnih hitrostih hoje istih subjektov za različne posnetke.

V tabeli 5.1 navajamo rezultate razpoznavanja za metodi povprečne silhuete (angl. *Averaged Silhouette – AS*) in verjetnostnega prostorsko-časovnega modela (angl. *Probabilistic Spatio-temporal Model – PSTM*). Obe metodi (brez normalizacije) dajeta relativno dobre rezultate pri enakih hitrosti hoje, medtem ko so rezultati pri spremenjenih hitrostih hoje opazno slabši. Izboljšanje lahko opazimo pri uporabi normalizacije hitrosti hoje pri obeh metodah, kar potrjuje domnevo iz podpoglavja 4.2, da se spremembe hitrosti hoje odražajo v dolžini koraka, metode za razpoznavanje pa lahko z uporabo tovrstne normalizacije bistveno zmanjšajo negativen učinek spremembe hitrosti hoje na učinkovitost razpoznavanja. V primeru povprečnih silhuet lahko opazimo 16 % izboljšanje, medtem ko v primeru verjetnostnega prostorsko-časovnega modela lahko opazimo v povprečju 24 % izboljšanje.

Dodatno lahko opazimo, da so rezultati metode na osnovi povprečnih silhuet nekoliko boljši od rezultatov metode PSTM. Razliko pripisujemo predvsem vlogi, ki jo igra videz posameznika v metodah na osnovi silhuet. Spremembe hitrosti hoje namreč ne vplivajo toliko na obliko silhuet kot na dinamiko gibanja, ki igra pomembno vlogo v metodi PSTM, ki je zato bolj podvržena vplivom spremembe hitrosti.

V tabeli 5.2 so prikazani rezultati sorodnih metod, ki obravnavajo spremembo hitrosti hoje za enake oz. primerljive eksperimente. Nekateri avtorji uporabljajo druge evalvacijske zbirke z drugačnimi hitrostmi. Takšna primerjava je možna, če se osredo-

Tabela 5.1

Učinkovitost razpoznavanja (v %) za metodo povprečnih silhuet (AS) in metodo verjetnostnega prostorsko-časovnega modela (PSTM) pri nespremenjeni hitrosti hoje (F/F in S/S) ter pri spremenjeni hitrosti hoje (F/S in S/F) z normalizacijo in brez nje (S – počasna hoja, F – hitra hoja). Zapis G/T označuje galerijo/testni primer.

G/T	AS	AS+norm	PSTM	PSTM+norm
F/F	96	96	96	96
S/S	100	100	88	88
F/S	68	84(+16)	52	68(+16)
S/F	76	92(+16)	44	76(+32)

točimo na primerjavo učinkov, doseženih s posebno obravnavo hitrosti hoje, hkrati pa razlike med počasno in hitro hojo ne smejo bistveno odstopati od naših eksperimentov. Osnovni algoritem (angl. *Baseline – BL*) [12] ne obravnava spremembe hitrosti hoje, a služi kot osnova za oceno izboljšanja pri nekaterih algoritmihi. Drugi navedeni so časovna normalizacija s pomočjo skritih markovih modelov (HMM) [47], transformacija prostornine dinamike silhete (angl. *Gait Silhouette Volume Transformation – GSVT*) [44], normalizacija koraka (angl. *Stride Normalization – SN*) [24], variacija oblike na osnovi izometričnih vzorcev (angl. *Shape Variation Based Frieze Patterns – SVBFP*) [48] ter poprej opisani normalizirani različici povprečnih silhuet (ASn) in verjetnostnega prostorsko-časovnega modela (PSTMn). Vrstica IZB. prikazuje izboljšavo učinka razpoznavanja metod s posebnim obravnavanjem hitrosti hoje v primerjavi brez te posebne obravnave za eksperiment hitra/počasna hoja (F/S), saj tega navaja večina avtorjev. Pri metodah GSVT in SN so avtorji uporabili druge evalvacijske zbirke. V primeru GSVT (zbirka OU-ISIR A) so bile hitrosti hoje 3 km/h (počasna) in 4 km/h (hitra), v primeru SN (lastna zbirka) pa 2,6 km/h (počasna) in 3,4 km/h (hitra).

Rezultati v tabeli 5.2, predvsem spodnja vrstica, jasno nakazujejo, da je učinek normalizacije hitrosti hoje koristen tako za metode, ki temeljijo na videzu, kot tudi za metode, ki modelirajo dinamiko gibanja. Metode, ki ne obravnavajo spremembe hitrosti hoje, lahko s takšno normalizacijo izboljšajo učinkovitost razpoznavanja in dosežejo primerljiv nivo metod, ki na poseben način obravnavajo spremembe hitrosti hoje.

Tabela 5.2

Primerjava s sorodnimi deli, ki obravnavajo spremembo hitrosti hoje. Metode so označene s stolpci BL, HMM, GSVT, SN, SVBF, ASn in PSTMn. Oznaka – pomeni, da rezultatov za konkretni eksperiment avtorji niso navedli. Vrstica IZB. prikazuje izboljšanje učinkovitosti razpoznavanja zaradi posebne obravnave spremembe hitrosti hoje, glede na učinkovitost brez takšne obravnave. Pri metodah HMM in GSVT smo uporabili rezultate osnovnega (BL) algoritma kot osnovo za izračun izboljšanja, saj vse temeljijo na silhuetah, avtorji pa ne navajajo drugih rezultatov za primerjavo.

<i>G/T</i>	<i>BL</i>	<i>HMM</i>	<i>GSVT</i>	<i>SN</i>	<i>SVBF</i>	<i>ASn</i>	<i>PSTMn</i>
<i>F/F</i>	–	–	–	–	100	96	96
<i>S/S</i>	92	–	–	–	100	100	88
<i>F/S</i>	72	84	96	88	80	84	68
<i>S/F</i>	–	–	84	96	82	92	76
<i>IZB.</i>	–	12	24	20	8	16	16

5.2 Evalvacija predlagane metode na osnovi skeleta

Razvito metodo na osnovi skeleta smo testirali na zbirki OU-ISIR [62], natančneje na zbirki A, ki jo sestavljajo silhete visoke kakovosti za 34 oseb, zajetih iz stranskega pogleda, ki hodijo po tekoči stezi z variacijami hitrosti od 2 km/h do 10 km/h v intervalih po 1 km/h. Za vsako osebo sta na voljo dve sekvenci hoje, ena za izgradnjo galerije, druga pa za testiranje.

Medhitrostni (angl. *cross-speed*) test definiramo kot eksperiment, pri katerem modeli v galeriji pripadajo eni hitrosti hoje, testni subjekti pa drugi hitrosti hoje. Izvedli smo večje število eksperimentov za različne kombinacije hitrosti hoje [75], pri čemer smo ločili medhitrostne eksperimente, ki ustrezajo hoji (2 km/h do 7 km/h) od tistih, ki ustrezajo teku (8,9 in 10 km/h), saj se te v antropoloških spremembah toliko razlikujejo, da jih v literaturi [21, 76] obravnavamo kot dve ločeni akciji (po analogiji s področjem razpoznavanja akcij). Razlike so tolikšne, da je primerjava med subjekti, ki tečejo, s subjekti, ki hodijo, praktično nesmiselna.

Rezultate podajamo v obliki kumulativne karakteristike ujemanja (angl. *Cumulative Match Characteristic – CMC*). To je mera uspešnosti delovanja biometričnega sistema v nalogi identifikacije na zaprti množici. Testne primere primerjamo z modeli v galeriji, ki jih za vsak testni primer razvrstimo po mestih glede na neko izbrano mero podobnosti (npr. razdaljo). CMC-krivulja nam pove, s kakšno verjetnostjo se pravilni model galerije nahaja med R -najboljšimi ujemanji. Mestu R pravimo rang. V našem

primeru navajamo rezultate za rang $R = 1$, ki pove, koliko testnih primerov iz celotne testne množice smo pravilno klasificirali že v prvem poskusu. Takšna mera uspešnosti razpoznavanja je zelo odvisna od števila modelov v galeriji, zato v eksperimentih uporabimo razdelitev 25:9, kot jo predlaga [44]. 25 oseb uporabimo za izgradnjo galerije in testiranje (v zbirki sta na voljo dve sekvenci za vsako osebo, ena za galerijo, druga za testiranje), preostalih 9 oseb pa služi za namene učenja, kadar je to potrebno (npr. pri transformaciji prostorov značilk). Takšno razdelitev so povzeli tudi drugi avtorji in je postala standard za ocenjevanje uspešnosti metod, ki obravnavajo vpliv spremembe hitrosti hoje na razpoznavanje.

Eksperimente smo razdelili v več skupin, da omogočimo analizo vpliva vseh treh načinov klasificiranja in transformacije prostora značilk.

5.2.1 Rezultati eksperimentov brez transformacije prostora značilk

Najprej podajamo rezultate eksperimentov za vse tri načine klasifikacije brez transformacije prostora značilk.

Rezultati v tabeli 5.3 prikazujejo celostni pristop klasifikacije. Opazimo lahko, da so precej skromni celo takrat, ko primerjamo modele v galeriji in testne primere, ki hodijo z isto hitrostjo. V primeru večjih hitrosti dosegamo nekoliko boljše rezultate. Razlogov za to je več. Višje hitrosti hoje razkrivajo več posameznikovih značilnosti (npr. pri teku so razlike med ljudmi že na oko bolj vidne kot pri počasni hoji), znotrajrazredna varianca je pri višjih hitrostih manjša, torej so si cikli istega subjekta bolj podobni. Poleg tega je značilnosti gibanja lažje zajeti pri večjih hitrostih, saj so takrat silhuete v fazah dvojne opore bolj razločne, manj podvržene samoprekrivanju in drugim šumom, kar vodi do manjših napak pri segmentaciji skeleta in v končni fazi omogoča, da zajamemo več informacije o gibanju.

Nadalje izvedemo eksperimente za klasifikacijo s fuzijo značilk in klasifikacijo na osnovi okvirjev. Rezultati v tabelah 5.4 in 5.5 kažejo opazno izboljšanje, pri čemer so drugi še nekoliko boljši. Za lažjo primerjavo glej povzetek rezultatov vseh eksperimentov na sliki 5.1.

Rezultate za vse tri načine klasifikacije brez transformacije prostora značilk za višje hitrosti, ki ustrezajo teku, podajamo v tabeli 5.6. Rezultati potrjujejo naše opažanje za primere večjih hitrosti gibanja: višje hitrosti hoje razkrivajo več posameznikovih značilnosti, cikli istega subjekta so si bolj podobni, silhuete pa so v fazah dvojne opore bolj razločne ter manj podvržene šumom.

Tabela 5.3

Uspešnost razpoznavanja (v %): rezultati medhitrostnih eksperimentov za celostni pristop klasificiranja brez transformacije prostora značilik za različne hitrosti hoje. G in T označujeta hitrost hoje modelov v galeriji ter hitrost hoje testnih primerov.

G/T	2km/h	3km/h	4km/h	5km/h	6km/h	7km/h
2km/h	84	80	44	24	16	16
3km/h	36	80	72	28	28	12
4km/h	44	68	92	84	52	28
5km/h	40	36	68	100	88	64
6km/h	16	24	28	68	92	92
7km/h	28	20	28	56	76	96

Tabela 5.4

Uspešnost razpoznavanja (v %): rezultati medhitrostnih eksperimentov za klasifikacijo s fuzijo značilik brez transformacije prostora značilik za različne hitrosti hoje.

G/T	2km/h	3km/h	4km/h	5km/h	6km/h	7km/h
2km/h	88	88	96	72	56	52
3km/h	68	92	96	84	76	52
4km/h	72	80	96	92	88	72
5km/h	56	72	96	100	100	96
6km/h	56	64	76	96	100	100
7km/h	36	56	52	80	92	100

5.2.2 Rezultati eksperimentov s transformacijo prostora značilik

Rezultati naslednjih eksperimentov, kjer uporabljamo transformacijo prostora značilik za vse tri načine klasifikacije, so podani v tabelah 5.7, 5.8 in 5.9. Transformacija prostora značilik izboljša rezultate razpoznavanja za vse tri načine klasifikacije. Če modeli iz galerije in testni primeri hodijo z isto hitrostjo, razlike sicer niso tako očitne (sprememba hitrosti o km/h v tabeli povzetka rezultatov 5.11 in sliki 5.1). Kljub temu tudi v teh primerih transformacija prostora značilik pozitivno vpliva na rezultate. Razlog je predvsem v učinku, ki ga dosežemo s tem, ko odstranimo povprečja celotne skupi-

Tabela 5.5

Uspešnost razpoznavanja (v %): rezultati medhitrostnih eksperimentov za klasifikacijo na osnovi okvirjev brez transformacije prostora značilk za različne hitrosti hoje.

G/T	2km/h	3km/h	4km/h	5km/h	6km/h	7km/h
2km/h	92	96	96	84	64	52
3km/h	92	88	84	80	64	64
4km/h	80	88	96	92	92	76
5km/h	68	80	92	100	92	76
6km/h	72	80	92	96	100	92
7km/h	44	64	76	96	92	100

Tabela 5.6

Uspešnost razpoznavanja (v %): rezultati medhitrostnih eksperimentov brez transformacije prostora značilk za različne hitrosti teka.

G/T	celostni			fuzija			okvirji		
	8km/h	9km/h	10km/h	8km/h	9km/h	10km/h	8km/h	9km/h	10km/h
8km/h	96	92	92	96	96	96	100	92	88
9km/h	92	100	100	100	100	100	80	92	88
10km/h	76	92	96	92	96	96	72	96	96

ne za določeno hitrost iz značilk posameznih subjektov in s tem v opisnikih značilk modeliramo le znotrajrazredne razlike, ki ostanejo v signalih.

Na koncu podajamo še rezultate transformacije prostora značilk pri hitrostih teka za vse tri načine klasifikacije (tabela 5.10). V primeru fuzije značilk dosegamo 100 % v vseh eksperimentih.

5.2.3 Povzetek rezultatov medhitrostnih eksperimentov

Tabela 5.11 in slika 5.1 prikazujeta povprečno uspešnost razpoznavanja medhitrostnih eksperimentov hoje in teka. Podani so povprečni rezultati za hojo, tek in različne stopnje spremembe hitrosti hoje (od 0 km/h do 5 km/h). Hojo in tek v navedenih rezultatih strogo ločujemo, podobno kot avtorji sorodnih del (glej podpoglavje 5.2.4), saj to

Tabela 5.7

Uspešnost razpoznavanja (v %): rezultati medhitrotnih eksperimentov za celostni pristop klasifikacije s transformacijo prostora značilk za različne hitrosti hoje.

G/T	2km/h	3km/h	4km/h	5km/h	6km/h	7km/h
2km/h	92	84	84	72	44	32
3km/h	68	88	84	72	60	48
4km/h	72	92	96	88	88	64
5km/h	64	76	100	100	96	92
6km/h	40	60	72	88	100	96
7km/h	36	36	48	76	96	92

Tabela 5.8

Uspešnost razpoznavanja (v %): rezultati medhitrotnih eksperimentov za klasifikacijo s fuzijo značilk in s transformacijo prostora značilk za različne hitrosti hoje.

G/T	2km/h	3km/h	4km/h	5km/h	6km/h	7km/h
2km/h	92	96	92	92	60	68
3km/h	80	96	100	84	76	64
4km/h	80	92	100	96	96	72
5km/h	80	92	100	100	96	100
6km/h	76	92	100	100	100	100
7km/h	72	60	80	96	100	100

omogoča pravičnejšo primerjavo. Razlog za tovrstno ločevanje v literaturi pa je predvsem v težavnosti primerjave med obema načinoma gibanja, saj so razlike med njima prevelike za učinkovito izvajanje navzkrižnega razpoznavanja.

Povprečno uspešnost razpoznavanja posamezne kategorije izračunamo kot povprečje rezultatov vseh eksperimentov uvrščenih v isti kategoriji na osnovi določene lastnosti po formuli:

$$\bar{R} = \frac{1}{N} \sum_{i=1}^N r_{i,} \quad (5.1)$$

Tabela 5.9

Uspešnost razpoznavanja (v %): rezultati medhitrostnih eksperimentov za klasifikacijo na osnovi okvirjev in s transformacijo prostora značilik za različne hitrosti hoje.

G/T	2km/h	3km/h	4km/h	5km/h	6km/h	7km/h
2km/h	96	100	92	100	88	68
3km/h	96	100	100	100	88	96
4km/h	92	100	96	96	92	88
5km/h	88	96	96	100	92	92
6km/h	92	92	92	96	100	96
7km/h	84	92	84	92	96	96

Tabela 5.10

Uspešnost razpoznavanja (v %): rezultati medhitrostnih eksperimentov s transformacijo prostora značilik za različne hitrosti teka.

G/T	celostni			fuzija			okvirji		
	8km/h	9km/h	10km/h	8km/h	9km/h	10km/h	8km/h	9km/h	10km/h
8km/h	96	96	96	100	100	100	100	92	88
9km/h	88	100	100	100	100	100	88	96	100
10km/h	92	96	96	100	100	100	96	100	100

kjer je N število eksperimentov v kategoriji s to lastnostjo, r_i pa rezultat uspešnosti razpoznavanja posameznega eksperimenta v kategoriji. Npr. kategorija z razliko v hitrosti hoje med galerijo in testnimi primeri $\Delta = 4$ km/h vsebuje eksperimente parov hitrosti galerije in testnih primerov: 2 km/h in 6 km/h, 3 km/h in 7 km/h, 6 km/h in 2 km/h ter 7 km/h in 3 km/h, kategorija tek pa vsebuje vse kombinacije hitrosti hoje od 8 km/h do 10 km/h tako za galerijo kot za testne primere (skupaj 9 eksperimentov). 0 km/h ponazarja eksperimente kategorije, kjer modeli v galeriji in testni primeri hodijo z isto hitrostjo. V stolpcih so podani rezultati za vse tri načine klasifikacije brez transformacije prostora značilik (stolpci od 1 do 3) in s transformacijo (stolpci od 4 do 6).

Tabela 5.12 in slika 5.2 prikazujeta učinek različnih načinov klasifikacije na uspeš-

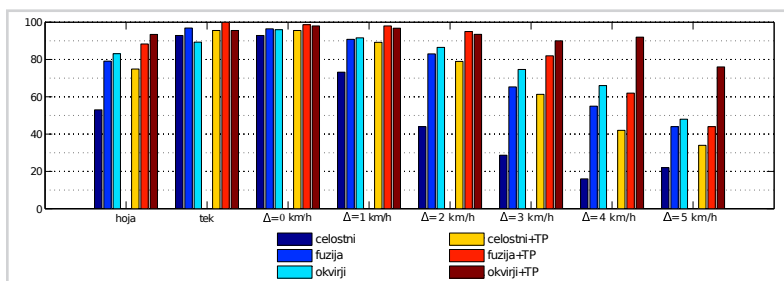
Tabela 5.11

Povprečna uspešnost razpoznavanja (v %) medhitrostnih eksperimentov hoje in teka glede na različne stopnje spremembe hitrosti.

	brez transformacije			s transformacijo		
	celostni	fuzija	okvirji	celostni	fuzija	okvirji
hoja	53,0	79,1	83,1	74,9	88,3	93,4
tek	92,9	96,9	89,3	95,6	100,0	95,6
$\Delta=0\text{ km/h}$	92,9	96,4	96,0	95,6	98,7	98,0
$\Delta=1\text{ km/h}$	73,2	90,8	91,6	89,2	98,0	96,8
$\Delta=2\text{ km/h}$	44,0	83,0	86,5	79,0	95,0	93,5
$\Delta=3\text{ km/h}$	28,7	65,3	74,7	61,3	82,0	90,0
$\Delta=4\text{ km/h}$	16,0	55,0	66,0	42,0	62,0	92,0
$\Delta=5\text{ km/h}$	22,0	44,0	48,0	34,0	44,0	76,0

Slika 5.1

Povprečna uspešnost razpoznavanja (v %) medhitrostnih eksperimentov hoje in teka glede na različne stopnje spremembe hitrosti.



nost razpoznavanja. Učinek posameznega načina klasifikacije je podan kot razlika v % glede na celostni način klasifikacije. Opazimo lahko znatno izboljšanje predvsem v primeru hoje in v primerih večje stopnje spremembe hitrosti hoje, pri čemer se klasifikacija na osnovi okvirjev obnese še nekoliko bolje kot klasifikacija s fuzijo značilik, predvsem pri večjih stopnjah spremembe hitrosti. V primeru transformacije prostora značilik opazamo podobno sliko z nekoliko manjšimi učinki izboljšave. Po drugi strani pa so učinki izboljšav precej manjši v primeru teka in hoje z nespremenjeno hitrostjo, kar je razumljivo, saj že celostni pristop deluje relativno dobro. V primeru teka gre

namreč le za eksperimente s stopnjami spremembe hitrosti hoje od 0 km/h do 2 km/h. Dodatno lahko opazimo, da klasifikacija na osnovi okvirjev v teh primerih deluje nekoliko slabše kot klasifikacija s fuzijo značilk, v ostalih primerih pa bolje. Razlog za boljše rezultate klasifikacije s fuzijo značilk tako pri manjših spremembah hitrosti kot teku je predvsem v tem, da večina klasifikatorjev, ki ustrezajo posameznim značilkam in tvorijo fuzijo, daje pri nespremenjenih pogojih dobre rezultate že na nivoju posameznega klasifikatorja. Ko te dobre rezultate združimo s fuzijo, se maloštevilne napake posameznih klasifikatorjev porazgubijo, zato je učinek fuzije zelo pozitiven. Po drugi strani pri klasifikaciji na osnovi okvirjev ne izkoriščamo fuzije, zaradi česar učinki te niso prisotni, izkoriščamo pa informacijo o razmerjih in korelacijah med posameznimi značilkami, ki se bolje ohranjajo tudi pri večjih spremembah hitrosti hoje. Večja kot je stopnja spremembe hitrosti, bolj očiten je učinek izboljšave klasifikacije na osnovi okvirjev.

Tabela 5.12

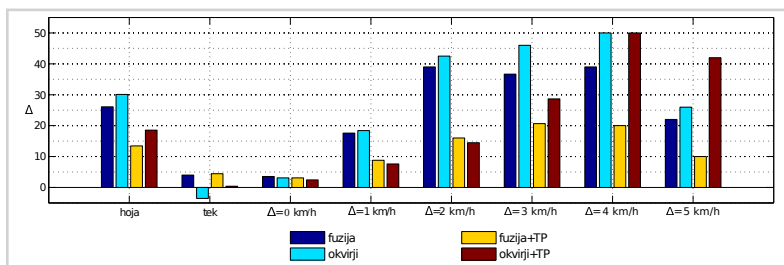
Učinek različnih načinov klasifikacije kot razlika v % uspešnosti razpoznavanja v primerjavi s celostnim načinom klasifikacije pri eksperimentih s transformacijo prostora značilk in brez nje glede na različne stopnje spremembe hitrosti.

	brez transformacije		s transformacijo	
	fuzija	okvirji	fuzija	okvirji
hoja	$\Delta=26,1$	$\Delta=30,1$	$\Delta=13,4$	$\Delta=18,6$
tek	$\Delta=4,0$	$\Delta=-3,6$	$\Delta=4,4$	$\Delta=0,0$
$\Delta=0\text{km/h}$	$\Delta=3,6$	$\Delta=3,1$	$\Delta=3,1$	$\Delta=2,4$
$\Delta=1\text{km/h}$	$\Delta=17,6$	$\Delta=18,4$	$\Delta=8,8$	$\Delta=7,6$
$\Delta=2\text{km/h}$	$\Delta=39,0$	$\Delta=42,5$	$\Delta=16,0$	$\Delta=14,5$
$\Delta=3\text{km/h}$	$\Delta=36,7$	$\Delta=46,0$	$\Delta=20,7$	$\Delta=28,7$
$\Delta=4\text{km/h}$	$\Delta=39,0$	$\Delta=50,0$	$\Delta=20,0$	$\Delta=50,0$
$\Delta=5\text{km/h}$	$\Delta=22,0$	$\Delta=26,0$	$\Delta=10,0$	$\Delta=42,0$

Tabela 5.13 in slika 5.3 prikazujeta učinek transformacije prostora značilk na uspešnost razpoznavanja. Učinek transformacije je podan kot razlika v % med uspešnostjo razpoznavanja s transformacijo in brez nje. V stolpcih so navedeni učinki izboljšav za vse tri načine klasifikacije. Opazimo lahko, da je učinek transformacije največji

Slika 5.2

Učinek različnih načinov klasifikacije kot razlika v % uspešnosti razpoznavanja v primerjavi s celostnim načinom klasifikacije pri eksperimentih s transformacijo prostora značilk in brez nje glede na različne stopnje spremembe hitrosti.

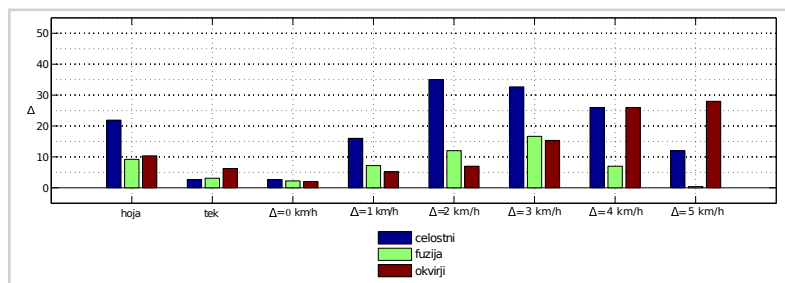


v primeru celostnega načina klasifikacije in nekoliko manjši v ostalih dveh primerih. Prav tako se po določeni stopnji spremembe hitrosti hoje učinek začne zmanjševati (celostni pri 3 km/h, fuzija pri 4 km/h), medtem ko se učinek pri pristopu na osnovi okvirjev povečuje do vključno 5 km/h stopnje spremembe hitrosti. To nakazuje na pomanjkljivost celostnega načina klasifikacije in klasifikacije na osnovi fuzije značilk pri obravnavi večjih sprememb hitrosti hoje, ki pa jo s klasifikacijo na osnovi okvirjev uspešno odpravimo.

Tabela 5.13

Učinek transformacije prostora značilk kot razlika v % uspešnosti razpoznavanja s transformacijo in brez nje pri eksperimentih za vse tri načine klasifikacije glede na različne stopnje spremembe hitrosti.

	celostni	fuzija	okvirji
<i>hoja</i>	$\Delta=21,9$	$\Delta=9,2$	$\Delta=10,3$
<i>tek</i>	$\Delta=2,7$	$\Delta=3,1$	$\Delta=6,2$
$\Delta=0\text{ km/h}$	$\Delta=2,7$	$\Delta=2,2$	$\Delta=2,0$
$\Delta=1\text{ km/h}$	$\Delta=16,0$	$\Delta=7,2$	$\Delta=5,2$
$\Delta=2\text{ km/h}$	$\Delta=35,0$	$\Delta=12,0$	$\Delta=7,0$
$\Delta=3\text{ km/h}$	$\Delta=32,7$	$\Delta=16,7$	$\Delta=15,3$
$\Delta=4\text{ km/h}$	$\Delta=26,0$	$\Delta=7,0$	$\Delta=26,0$
$\Delta=5\text{ km/h}$	$\Delta=12,0$	$\Delta=0,0$	$\Delta=28,0$



Slika 5.3

Učinek transformacije prostora značilik kot razlika v % uspešnosti razpoznavanja s transformacijo in brez nje pri eksperimentih za vse tri načine klasifikacije glede na različne stopnje spremembe hitrosti.

5.2.4 Primerjava s sorodnimi deli

Za namene preverjanja učinkovitosti metod razpoznavanja po načinu hoje pri različnih spremembah hitrosti so avtorji v [44] definirali dva eksperimenta na zbirki OU-ISIR-A, ki omogočata primerjavo s sorodnimi deli. Našo metodo primerjamo z naslednjimi sorodnimi deli, ki vsa posvečajo posebno pozornost obravnavanju spremembe hitrosti: časovna normalizacija s pomočjo skritih markovih modelov (HMM) [47], normalizacija koraka (angl. *Stride Normalization – SN*) [24], transformacija volumenske silhuete (angl. *Silhouette Volume Transformation – SVT*) [44], konfiguracija oblik višjega reda (angl. *High-order Shape Configuration – HSC*) [51] in model razlik v sestavi (angl. *Differential Composition Model – DCM*) [52]. Rezultati metode HMM in SN so pridobljeni na drugih bazah, medtem ko so rezultati SVT, HSC in DCM pridobljeni na isti bazi, kot jo uporabljamo v naših eksperimentih. V obeh primerih so uporabljene podobne hitrosti hoje, da je možna primerjava. Za scenarij manjše spremembe hitrosti so pri metodi HMM uporabili 3,3 km/h in 4,5 km/h, medtem ko pri zbirki OU-ISIR-A uporabimo 3 km/h in 4 km/h. Za scenarij večje spremembe hitrosti pa so pri metodi SN uporabili 2,5 km/h in 5,8 km/h, kar približno ustreza hitrostim 2 km/h in 6 km/h v primeru zbirke OU-ISIR-A. Oba eksperimenta izvedemo dvakrat, prvič uporabimo za galerijo nižjo hitrost, za testne primere pa večjo hitrost, drugič obratno, končni rezultat pa predstavlja povprečje obeh izvedb. Dodatno navajamo še tretji eksperiment, in sicer kot povprečje vseh medhitrostnih eksperimentov za različne kombinacije hitrosti hoje (od 2 km/h do 7 km/h) galerije in testnih primerov, za primerjavo z metodama HSC in DCM, katerih avtorji so objavili rezultate za isti tip eksperimenta.

Rezultati v tabeli 5.14 prikazujejo učinkovitost razpoznavanja različnih metod za oba poprej opisana testa in povprečno učinkovitost za vse eksperimente z različnimi

Tabela 5.14

Učinkovitost razpoznavanja (v %): primerjava s sorodnimi deli, ki obravnavajo spremembe hitrosti hoje. Oznaka – pomeni, da rezultatov za konkretni eksperiment avtorji niso navedli.

Scenarij	Fuzija značilk	Okvirji	SN [24]	HMM [47]	SVT [44]	HSC [51]	DCM [52]
Manjša sprememba hitrosti (3 km/h / 4 km/h)	96	100	–	84	90	96	98
Večja sprememba hitrosti (2 km/h / 6 km/h)	68	90	34	–	58	68	82
Celotna zbirka – hoja (2 km/h – 7 km/h)	88,33	93,44	–	–	–	85,67	92,44

hitrostmi hoje (od 2 km/h do 7 km/h) zbirke OU-ISIR-A (scenarij: Celotna zbirka – hoja). Rezultate navajamo le za tiste metode, ki so objavile določen tip eksperimenta. Opazimo lahko, da je naša metoda s klasifikacijo s fuzijo značilk na 2. mestu. Boljše rezultate dosega le metoda DCM, predvsem v scenarijih večjih sprememb hitrosti (4 km/h). V primeru klasifikacije na osnovi okvirjev pa je učinkovitost razpoznavanja najboljša, saj v tem primeru dosegamo primerljivo učinkovitost tudi pri scenarijih z večjimi spremembami hitrosti.

Razvite metode na osnovi modela in značilk, ki opisujejo predvsem dinamiko gibanja, demonstrirajo torej primerljivo učinkovitost z drugimi sorodnimi deli, ki temeljijo na silhuetah, in s tem posredno izkoriščajo videz za boljše razločevanje med različnimi subjekti, kljub temu, da je dinamika gibanja bolj podvržena vplivom spremembe hitrosti kot sam videz. Opazimo lahko tudi, da v zgornji tabeli ni nobene druge metode na osnovi modela, predvsem zaradi tega, ker obstoječe metode na osnovi modela redko dosegajo primerljive rezultate z metodami brez modela, celo v primerih brez prisotnosti faktorjev vpliva. Naša razvita metoda predstavlja edinstven primer metode na osnovi modela, ki posebej obravnava faktor spremembe hitrosti hoje in v primeru uporabe klasifikacije na osnovi okvirjev celo presega učinkovitost delovanja vseh drugih sorodnih del. Kot taka predstavlja nov prispevek na raziskovalnem področju analize gibanja.

5.3 Vpliv velikosti galerije na učinkovitost razpoznavanja

Ker je zbirka OU-ISIR relativno majhna (34 oseb), smo želeli razvito metodo preizkusiti še na nekoliko večji zbirki, ki hkrati ni omejena na hojo po tekoči stezi. Zbirka CASIA (C) [61] vsebuje 153 oseb za tri različne hitrosti hoje, opredeljene kot hitra, počasna in normalna hoja, ki v poprečju približno ustrezajo hitrostim 4, 5 in 6 km/h.

Izvedli smo več eksperimentov za vse različne kombinacije hitrosti hoje za galerijo in testne primere: normalna/normalna (nn), hitra/hitra (hh), počasna/počasna (pp), normalna/hitra (nh), normalna/počasna (np) in hitra/počasna (hp). Dodatno je za vsako hitrost hoje na voljo več sekvenc: normalna hoja štiri sekvence, hitra hoja dve sekvenci ter počasna hoja dve sekvenci. V primerih, ko modeli v galeriji pripadajo eni hitrosti hoje, testni primeri pa drugi, so rezultati podani kot povprečje vseh eksperimentov obeh kombinacij (npr. hitra/počasna štirje eksperimenti in počasna/hitra štirje eksperimenti, skupni rezultat za kombinacijo hitra/počasna hoja je torej povprečje osmih eksperimentov).

Rezultati v tabeli 5.15 in na sliki 5.4 prikazujejo povprečno učinkovitost razpoznavanja opisanih eksperimentov na zbirki CASIA (C) v primerjavi s povprečno učinkovitostjo razpoznavanja eksperimentov zbirke OU-ISIR s klasifikacijo na osnovi okvirjev in s transformacijo prostora značilk. Za učenje transformacije prostora značilk smo uporabili 33 oseb za učenje in 120 za klasifikacijo.

Tabela 5.15

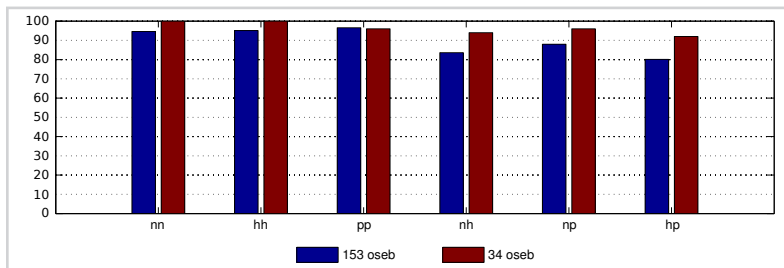
Primerjava učinkovitosti razpoznavanja (v %) večje zbirke CASIA (C) s 153 osebami z učinkovitostjo razpoznavanja manjše zbirke OU-ISIR s 34 osebami. Povprečni rezultat je izračunan na osnovi n eksperimentov iste skupine.

Zbirka	nm	hh	pp	nh	np	hp
CASIA (153 oseb)	94.61 (n=12)	95.09 (n=2)	96.56 (n=2)	83.61 (n=16)	88.01 (n=16)	80.11 (n=8)
OU-ISIR (34 oseb)	100.00 (n=1)	100.00 (n=1)	96.00 (n=1)	94.00 (n=2)	96.00 (n=2)	92.00 (n=2)

Opazimo lahko, da se učinkovitost razpoznavanja na zbirki CASIA (C) s 153 osebami nekoliko poslabša v primerjavi z manjšo zbirko OU-ISIR s 34 osebami, a predvsem pri eksperimentih z različno hitrostjo hoje. Pri eksperimentih z enako hitrostjo hoje je poslabšanje minimalno. Razlog za večje poslabšanje pri medhitrostnih eksperimentih je predvsem v tem, da je zbirka CASIA (C) namenjena analizi vpliva velikosti galerije

Slika 5.4

Primerjava učinkovitosti razpoznavanja (v %) večje zbirke CASIA (C) s 153 osebami z učinkovitostjo razpoznavanja manjše zbirke OU-ISIR s 34 osebami.



in ne toliko analizi spremembe hitrosti hoje. Osebe v zbirki so grobo razvrščene v tri hitrostne skupine (počasna, normalna in hitra), pri čemer se hitrosti hoje med posameznimi osebami iste skupine razlikujejo in le v povprečju celotne skupine ustrezajo vrednostim 4, 5 in 6 km/h. Takšna razpršenost hitrosti otežuje modeliranje prostora značilik in zato transformacija prostora značilik v tem primeru ne pride do izraza ter ne prispeva takšnega izboljšanja kot na zbirki OU-ISIR, kjer lahko posamezne hitrosti hoje zelo natančno modeliramo.

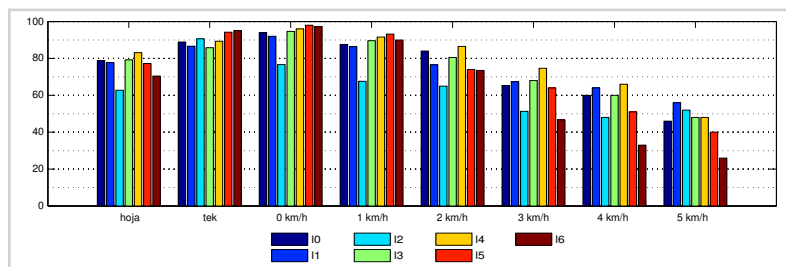
Učinkovitost razpoznavanja klasifikacije na osnovi okvirjev je torej tudi na skoraj petkrat večji zbirki, vsaj v primeru nespremenjene hitrosti hoje, primerljiva učinkovitosti manjše zbirke.

5.4 Vpliv valčkov na učinkovitost razpoznavanja

V podglavju 3.5.6 smo podali način uporabe analize z valčki kot pripomoček za izboljšanje rezultatov razpoznavanja. Tovrstno analizo smo v predstavljenih eksperimentih uporabili le pri klasifikaciji na osnovi okvirjev namesto trigonometrične aproksimacije, ki smo jo uporabili v preostalih načinih klasifikacije. S tem smo želeli raziskati vpliv valčkov na rezultate razpoznavanja in razjasniti predvsem smiselnost tovrstne analize.

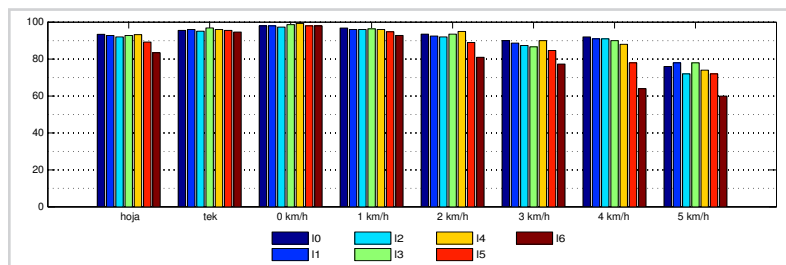
Na slikah 5.5 in 5.6 je prikazana učinkovitost razpoznavanja medhitrostnih eksperimentov pri uporabi različnih nivojev valčne analize, pri eksperimentih brez transformacije prostora značilik in z njo. Nivo 10 prikazuje učinkovitost uporabe celotnih signalov, torej brez aproksimacije, nivoji 11 do 16 pa prikazujejo učinkovitost uporabe aproksimiranih signalov z valčno analizo na nivojih od 1 do 6.

Opazimo lahko, da pride do razlik v učinkovitosti in s tem razlogov za smiselnost



Slika 5.5

Povprečna uspešnost razpoznavanja (v %) medhitrostnih eksperimentov glede na uporabljen nivo analize z valčki pri eksperimentih brez transformacije prostora značilk.



Slika 5.6

Povprečna uspešnost razpoznavanja (v %) medhitrostnih eksperimentov glede na uporabljen nivo analize z valčki pri eksperimentih s transformacijo prostora značilk.

uporabe tovrstne analize predvsem v primeru, ko ne uporabljamo transformacije prostora značilk. Na sliki 5.5 lahko opazimo, da učinkovitost do določenega nivoja I narašča, predvsem v primerih medhitrostnih eksperimentov z večjimi razlikami v hitrosti hoje, nato pa ponovno pada. Za primer hoje dosežemo največjo učinkovitost pri nivoju I_4 , zato smo ta nivo uporabili tudi pri eksperimentih brez transformacije prostora značilk (podpoglavje 5.2.1) za primere klasifikacije na osnovi okvirjev. Po drugi strani razlike v učinkovitosti na sliki 5.6 med različnimi nivoji valčne analize niso tako očitne. Kljub temu je možno zaznati upadanje učinkovitosti pri večjih spremembah hitrosti, od določenega nivoja I dalje, medtem ko je učinkovitost na prvih nekaj nivojih precej podobna. V primeru eksperimentov s transformacijo prostora značilk (podpoglavje 5.2.2) za primere klasifikacije na osnovi okvirjev zato valčne analize nismo uporabili in učinkovitost ustreza nivoju I_0 , torej uporabi celotnih, neaproximiranih signalov.

Analiza z valčki torej potrjuje predpostavko, da že sama transformacija prostora značilk omogoča modeliranje le tistih delov signalov, ki so najpomembnejši za razločevanje subjektov. Dodatno lahko potrdimo, da lahko s pomočjo takšne analize izboljšamo učinkovitost razpoznavanja pri spremenjenih hitrostih hoje predvsem v primerih, ko

ne želimo uporabiti oz. nimamo na voljo možnosti predhodnega učenja, ki je potrebno za transformacijo prostora značilk.

5.5 Vpliv faznega zamika ciklov na učinkovitost razpoznavanja

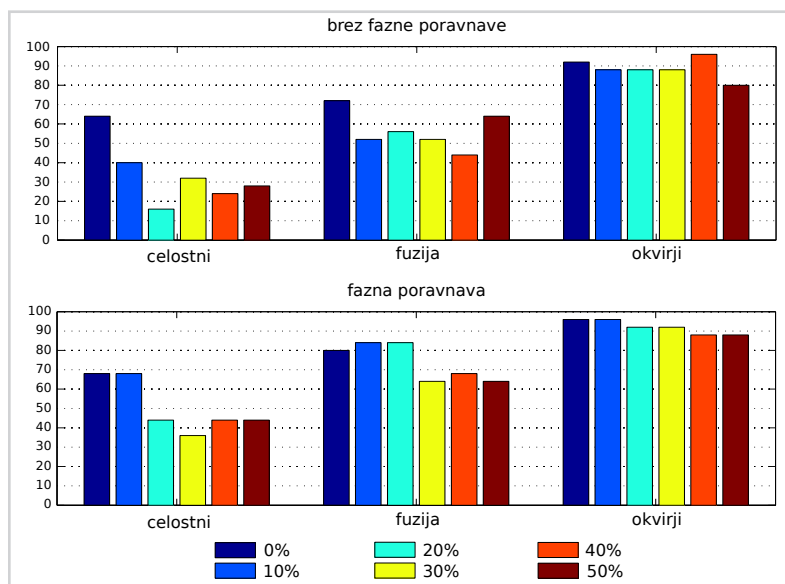
Ker primerjava značilk pri celostnem pristopu in fuziji značilk temelji na iskanju razdalje, je pred izvedbo klasifikacije nujno treba zagotoviti, da so primerjani cikli med seboj poravnani, saj že manjša odstopanja močno vplivajo na velikost razdalje. V teh primerih dosežemo poravnavo s pomočjo fazne poravnave, opisane v podpoglavju 3.5.1. V primeru klasifikacije na osnovi okvirjev ta omejitve ne obstaja, s čimer dosežemo neodvisnost klasifikacije od začetne točke cikla.

Za demonstracijo zadnje trditve smo izvedli eksperiment na bazi OU-ISIR, kot je to opisano v podpoglavju 5.2, pri čemer smo izvajali le en medhitrostni test, kjer učno množico predstavlja skupina subjektov, ki hodi s hitrostjo 5 km/h, testno množico pa skupina subjektov s hojo 3 km/h. Pri eksperimentu smo uporabili transformacijo prostora značilk. Predhodno smo signale vseh primerov galerije in testnih primerov zamaknili za določene deleže dolžine cikla. Npr. pri deležu zamika 20 % smo pri vsakem primeru iz množice subjektov naključno izbrali zamik v intervalu od -20 % do 20 % dolžine cikla, s čimer smo vsakemu primeru predstavili začetke ciklov. Nivo neporavnosti med dvema subjektoma je torej v tem primeru lahko od 0 % do 40 % dolžine cikla. Izvedli smo dve različici eksperimenta. V prvi različici nismo uporabljali fazne poravnave ciklov za nobenega od načinov klasifikacije in smo subjekte primerjali takšne, kot so, torej neporavnane. V drugi različici pa smo izvajali fazno poravnavo, kot je opisano v podpoglavju 3.5.1, s to razliko, da smo interval poravnavanja prilagajali deležem zamika.

Rezultati eksperimentov, prikazani na sliki 5.7, prikazujejo vpliv faznega zamika na učinkovitost razpoznavanja. Opazimo lahko, da se v primeru celostnega pristopa in fuzije značilk vpliv faznega zamika povečuje sorazmerno s povečevanjem deleža zamika, v primeru klasifikacije na osnovi okvirjev pa ta vpliv ni razviden. Dodatno lahko v primeru uporabe fazne poravnave opazimo, kako fazna poravnava, izvedena pred klasifikacijo, vpliv zamika nekoliko omili, a tudi v tem primeru učinkovitost razpoznavanja pade pri večjih deležih zamika. Vzrok je predvsem v tem, da pri fazni poravnavi večjih obsegov prihaja do zamenjave delov cikla leve in desne noge, kot je opisano v podpoglavju 3.4.3, zato ta ni primerna za večje fazne zamike in jo lahko učinkovito uporabljamo le pri manjših zamikih. Iz tega razloga že v fazi iskanja ciklov poiščemo

faze dvojne opore, ki se začnejo z isto nogo, kar zagotavlja grobo poravnano ciklov, ki kasneje omogoča učinkovito uporabo fazne poravnave.

Med rezultati lahko pri nekaterih zamikih opazimo tudi občasna odstopanja od trendov. Prvi razlog so zamiki blizu 50 % cikla, ki omogočijo ponovno poravnano nekaterih subjektov. Primer vidimo pri fuziji značilnik brez fazne poravnave, ko se rezultati pri 50 % nenadoma izboljšajo. Drugi razlog pa je predvsem v izvedbi enega samega testa, kar v kombinaciji z naključno izbiro zamikov ne nudi dovolj statistične moči za popolnoma točne rezultate. Primer takšnega odstopanja vidimo npr. na rezultatih klasifikacije na osnovi okvirjev brez fazne poravnave pri 40 % zamika.



Slika 5.7

Učinkovitost razpoznavanja pri različnih nivojih faznih zamikov (od 0 % do 50 %) začetkov ciklov za vse tri načine klasifikacije, tako brez fazne poravnave (zgoraj) kot s fazno poravnavo (spodaj).

Kljub statistični nenatančnosti so razlike v trendih rezultatov očitne. V obeh različnih izvedbah eksperimenta lahko opazimo podobnost v učinkovitosti delovanja klasifikacije na osnovi okvirjev, ki je relativno neobčutljiva na fazni zamik ciklov subjektov, kar odpravi potrebo po zapletenem in težavnem postopku fazne poravnave ter iskanju faz dvojne opore, ki se začnejo z isto nogo. Razvito metodo tako poenostavimo, s čimer izboljšamo robustnost delovanja in njeno učinkovitost.

5.6 Vpliv šuma, ločljivosti in netipičnih meritev

Zaradi občutljive narave segmentacije skeleta človeškega telesa je razvita metoda podvržena več tipom napak, ki jih lahko povzroči šumno zajetje videoposnetka, slaba kakovost silhuet, pridobljenih s postopkom segmentacije gibanja, ali objekti v sceni, ki prekrivajo subjekt zajemanja.

Najpogostejše med njimi so majhne napake v meritvah položajev kosti in sklepov zaradi šumnih silhuet. Npr. naključne ploskve, ki se začasno pojavijo v posameznih okvirjih videoposnetka, pojavljanje lukenj v notranjosti silhete, ki ločijo del telesa ali silhete od preostalega dela, začasna zapolnitev prostora med obema nogama v višini kolen ali gležnjev ipd. Tovrstne napake se pojavljajo le na nekaj zaporednih okvirjih, nato izginejo in se morda spet pojavijo na nekem drugem mestu. Rezultat takšnih napak je šum v signalih, ki jih pridobimo iz skeleta. Ta šum učinkovito obvladujemo z detekcijo izstopajočih konic v signalu (angl. *spike detection*), odstranitvijo teh konic in interpolacijo odstranjenih vrednosti, kot je opisano v podpoglavju 3.4.5. Dodatno so vsi signali filtrirani z aproksimacijo s trigonometričnimi polinomi (glej podpoglavje 3.5.2) oz. z aproksimacijo z valčki (glej podpoglavje 3.5.6) v primeru klasifikacije na osnovi okvirjev.

Večje motnje lahko povzročajo prekrivajoči objekti v sceni, manjkajoči večji deli silhuet, delov telesa ipd. Ko se tovrstne napake pojavijo le za kratek čas v le nekaj okvirjih videoposnetka, večino negativnih učinkov izločijo prijemi za obravnavanje šuma opisani v prejšnjem odstavku. Po drugi strani pa lahko daljša prisotnost tovrstnih motenj povzroči napačne meritve hkrati na več značilkah, skozi celoten cikel ali celo več ciklov gibanja. Negativne učinke manjkajočih oz. slabo izmerjenih značilk obvladujemo s pomočjo pristopa fuzije značilk (glej podpoglavje 3.5.4), kjer predstavlja vsaka značilka svoj klasifikator in vsak klasifikator vpliva le v manjšem deležu na končni rezultat klasifikacije, medtem ko večino končnega rezultata še vedno prispevajo preostale nepoškodovane značilke. Če napake povzročijo uničenje celotnih gibalnih ciklov, se ti cikli pojavljajo kot izstopajoče vrednosti v postopku izračuna razdalje med testnimi in učnimi primeri, katerih učinke izničimo s pomočjo mediane razdalj med vsemi cikli gibanja (glej podpoglavje 3.5.3).

V primeru klasifikacije na osnovi okvirjev, kjer ne uporabljamo aproksimacije s trigonometričnimi polinomi, fuzije značilk in ne izračunavamo razdalj med cikli gibanja, dosežemo podobne učinke s pomočjo aproksimacije z valčki ter s hierarhičnim

združevanjem rezultatov klasifikacije (glej podpoglavje 3.5.5), ki na najnižjem nivoju omogoča eliminacijo kratkotrajnih motenj (eliminacija okvirjev z izstopajočimi vrednostmi) pri določanju razreda cikla po večinskem pravilu iz rezultatov klasifikacije posameznih okvirjev ter na višjem nivoju omogoča še eliminacijo celotnih izstopajočih ciklov pri določanju razreda videoposnetka ponovno po večinskem pravilu iz rezultatov klasifikacije posameznih ciklov.

V nadaljevanju navajamo rezultate eksperimentov, ki prikazujejo vpliv različnih vrst šuma na rezultate razpoznavanja pri klasifikaciji na osnovi okvirjev. Vsi eksperimenti so bili izvedeni na bazi OU-ISIR, kot je to opisano v podpoglavju 5.2, pri čemer smo izvajali le en medhitrostni test, kjer učno množico predstavlja skupina subjektov, ki hodi s hitrostjo 5 km/h, testno množico pa skupina subjektov s hojo 3 km/h. To kombinacijo smo izbrali kot tipičnega predstavnika medhitrostnih eksperimentov, katerega razlika v hitrosti hoje ni niti prevelika, da bi preveč pokvarila izhodiščne rezultate in s tem otežila analizo pri prisotnosti večjih količin šuma, niti premajhna, da bi vpliv šuma ne bil vsaj do neke mere razviden tudi pri manjših količinah prisotnosti. Dodatno prisotnost dejavnika spremembe hitrosti hoje v analizo zajame tudi dele metode, kot je transformacija prostora značilk, katere delovanje pride do izraza predvsem v medhitrostnih eksperimentih, s čimer zagotovimo pokritost celotne metode v analizi. Pri vseh eksperimentih smo uporabili klasifikacijo na osnovi okvirjev ter transformacijo prostora značilk.

5.6.1 Vpliv ločljivosti silhuet na učinkovitost razpoznavanja

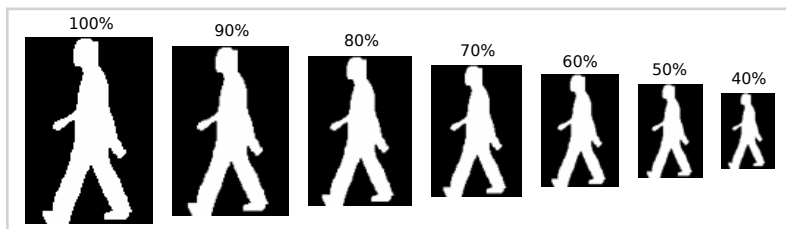
Glede na relativno kompleksen proces segmentacije skeleta se v prvi meri pojavi vprašanje, kakšna je minimalna ločljivost vhodnih slik, da je takšna segmentacija sploh možna, ter v kolikšni meri ločljivost vpliva na rezultate klasifikacije. Postopali smo tako, da smo vhodno sliko silhuete pomanjšali za določen odstotek, kot je to razvidno iz tabele 5.16. Ločljivost osnovne silhuete (100 %) je enaka 128×88 slikovnih elementov, najmanjša ločljivost, pri kateri algoritem segmentacije skeleta še deluje, pa ustreza 40 % osnovne ločljivosti, torej 52×36 slikovnih elementov. Manjše ločljivosti niso možne, saj so deli telesa premajhni za uspešno določitev orientacije posameznih delov. Primeri silhuet v posameznih ločljivostih so prikazani na sliki 5.8.

Na sliki 5.9 je prikazana učinkovitost klasifikacije na različnih ločljivostih za posamezne nivoje hierarhičnega združevanja rezultatov klasifikacije. Rezultati klasifikacije okvirjev predstavljajo najnižji nivo, nivo višje so rezultati klasifikacije ciklov gibanja, na

Tabela 5.16

Ločljivost slik silhuet.

delež	širina	višina
100 %	128	88
90 %	116	80
80 %	103	71
70 %	90	62
60 %	77	53
50 %	64	44
40 %	52	36



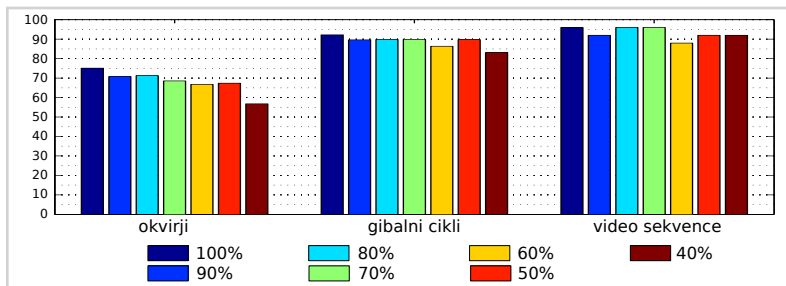
Slika 5.8

Ločljivosti silhuet.

najvišjem nivoju pa so rezultati klasifikacije celotnih videosekvenc hoje. Vpliv ločljivosti na učinkovitost razpoznavanja je razvidna predvsem na nivoju okvirjev, medtem ko lahko na rezultatih višjih nivojev opazimo, kako hierarhično združevanje učinkovito izniči negativen vpliv zmanjšanja ločljivosti. Ugotovimo lahko, da je kljub negativnemu vplivu ločljivosti vhodnih silhuet na rezultate razpoznavanja možno izvajati razvito metodo tudi na slikah z relativno nizko ločljivostjo.

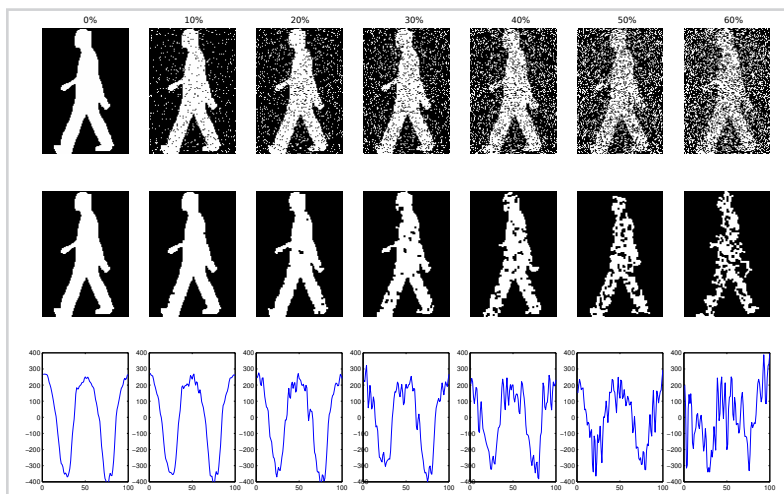
Slika 5.9

Učinkovitost razpoznavanja pri različnih ločljivostih silhuet za različne nivoje hierarhičnega združevanja rezultatov. Najnižji nivo: klasifikacija okvirjev, srednji nivo: klasifikacija ciklov, najvišji nivo: klasifikacija videosekvenc.



5.6.2 Vpliv šuma v silhuetah na učinkovitost razpoznavanja

Za ta eksperiment slikam učnih in testnih subjektov dodamo šum sol in poper (angl. *salt & pepper*), kjer vhodnim slikam dodamo določen delež črnih ali belih slikovnih elementov, glede na celotno število točk v sliki. Lokacije dodanih točk pa se izbirajo naključno. Izvedli smo eksperimente za deleže dodanega šuma od 0 % do 60 %. Višji deleži šuma silhueto popačijo do te mere, da segmentacija skeleta ni več mogoča. Uporaba tovrstnega šuma (namesto Gaussovega belega šuma) je smiselna predvsem zato, ker gre za binarne slike silhuet, poleg tega se večja količina prisotnosti šuma odraža v silhuetah podobno kot motnje, ki jih povzročijo nekoliko slabša segmentacija gibanja. Primeri silhuet z dodanim šumom sol in poper so prikazani v zgornji vrstici slike 5.10. Tovrsten šum v prvi meri omilimo z morfološkimi operacijami na silhuetah, kot je prikazano v sredinski vrstici slike. Opazimo lahko, da morfološke operacije učinkovito odstranijo bele točke, medtem ko luknje v silhuetah ostajajo, kar je značilno za tovrsten šum. Hkrati je namreč možno odstraniti le en tip šuma, ali sol ali poper, ne pa tudi obeh. Rezultat morfoloških operacij je podoben dejanskemu stanju slabšega rezultata segmentacije gibanja. Spodnja vrstica prikazuje vpliv šuma na signale, ki jih pridobimo iz skeleta človeškega telesa.



Slika 5.10

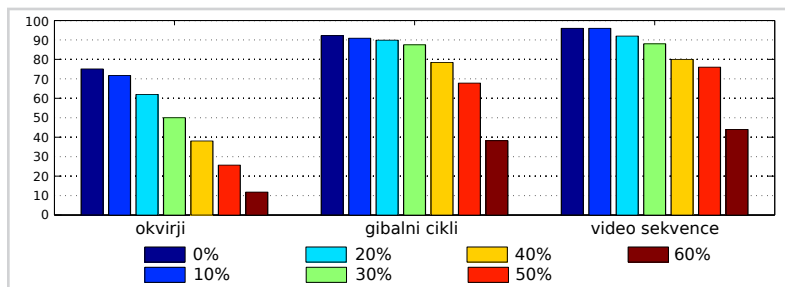
Primeri silhuet z dodanim šumom sol in poper (zgornja vrstica), obdelava silhuet z morfološkimi operacijami (srednja vrstica) in primeri signalov, pridobljenih iz skeleta človeškega telesa za deleže dodanega šuma od 0 % do 60 %.

Na sliki 5.11 je prikazana učinkovitost klasifikacije za različne deleže dodanega šuma

za posamezne nivoje hierarhičnega združevanja rezultatov klasifikacije. Vpliv šuma na učinkovitost razpoznavanja je v tem primeru razvidna na vseh treh nivojih, kljub temu pa je možno opaziti omiljenje učinkov šuma na višjih nivojih z bistvenim preskokom na nivoju klasifikacije ciklov gibanja. Ugotovimo lahko, da je izvajanje razpoznavanja smiselno pri deležih šuma, nižjih od 50 %.

Slika 5.11

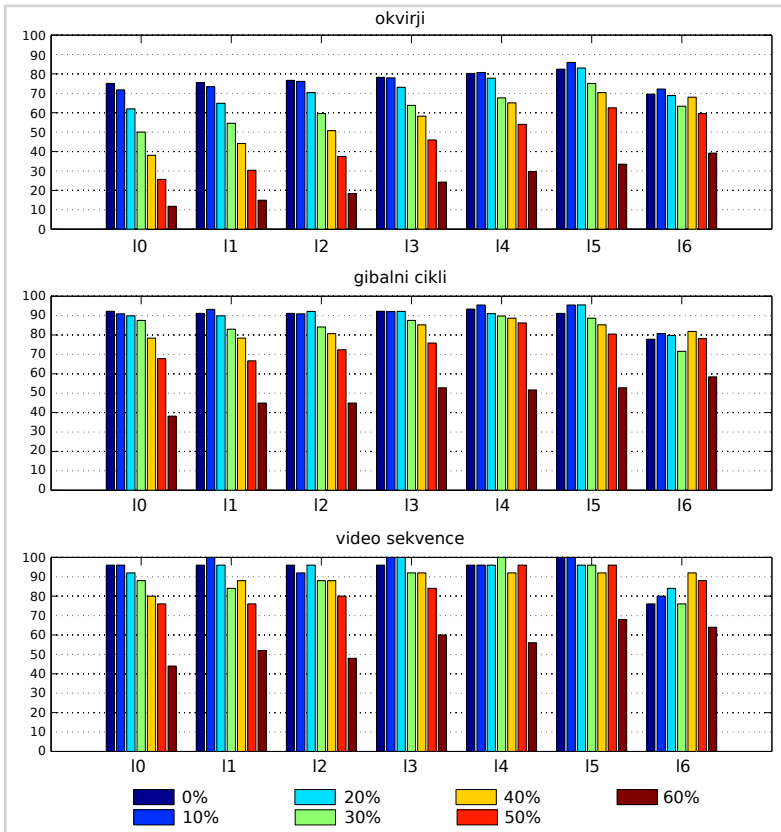
Učinkovitost razpoznavanja pri prisotnosti šuma sol in poper v silhuetah za različne nivoje hierarhičnega združevanja rezultatov.



Dodatno je na sliki 5.12 prikazano, kako lahko vpliv šuma omilimo z uporabo valčkov. Izboljševanje učinkovitosti razpoznavanja lahko opazimo na rezultatih valčkov nekje do nivoja I5. Še več, če opazujemo klasifikacijo videosekvenc (spodnja vrstica), lahko ugotovimo, da smo negativni učinek šuma do 50 % skoraj popolnoma izničili že na nivoju valčkov I4. S pomočjo valčkov je torej možno učinkovito odstraniti negativen vpliv šuma sol in poper na rezultate razpoznavanja.

5.6.3 Vpliv šuma v skeletu na učinkovitost razpoznavanja

Glede na to, da je sama segmentacija skeleta človeškega telesa aktivno področje raziskovanja [6, 77, 78], ki je dobilo v zadnjem času še dodaten zagon s pojavom tehnologij, kot je Microsoft Kinect [79], je na mestu vprašanje, kako kakovostno segmentacijo skeleta človeškega telesa potrebujemo za relativno učinkovito delovanje predlagane metode. Ali nekoliko natančneje, kako kakovostni morajo biti pridobljeni signali s pomočjo skeleta človeškega telesa, da je razpoznavanje še lahko učinkovito. Dodatno temelji segmentacija skeleta človeškega telesa, opisana v poglavju 3.3, na uporabi povprečnih anatomskih značilnosti človeškega telesa. Te pa so lahko pri dejanskih subjektih nekoliko drugačne, saj vsak posameznik do neke mere odstopa od povprečja. Napaka iz tega naslova se prav tako odraža kot šum v koraku segmentacije skeleta, zato z odgovorom na zgornje vprašanje vsaj do neke mere opišemo tudi vpliv napake, zaradi odstopanja



Slika 5.12

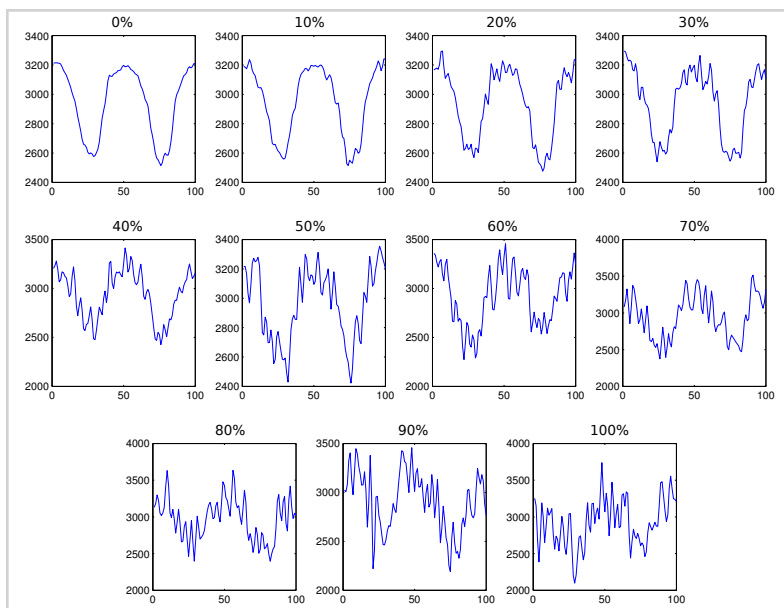
Učinkovitost razpoznavanja pri prisotnosti šuma sol in poper v silhuetah za različne nivoje aproksimacije z valčki (od I0 do I6) na različnih nivojih hierarhičnega združevanja. I0 ustreza rezultatom brez aproksimacije.

posameznikovih anatomskih značilnosti od povprečja, na učinkovitost razpoznavanja. Odgovor na oboje poskušamo torej poiskati z dodajanjem določenega deleža Gaussovega šuma pridobljenim signalom značilk:

$$y_n = y + w\sigma\phi, \tag{5.2}$$

kjer je y signal, w delež dodanega šuma, σ standardna deviacija signala in ϕ generator naključnih števil z normalno porazdelitvijo s srednjo vrednostjo 0 in standardno deviacijo 1. Signalu torej dodamo Gaussov šum s srednjo vrednostjo 0 in standar-

dno deviacijo $w\sigma$. Slika 5.13 prikazuje primere signalov z dodanimi različnimi deleži Gaussovega šuma.

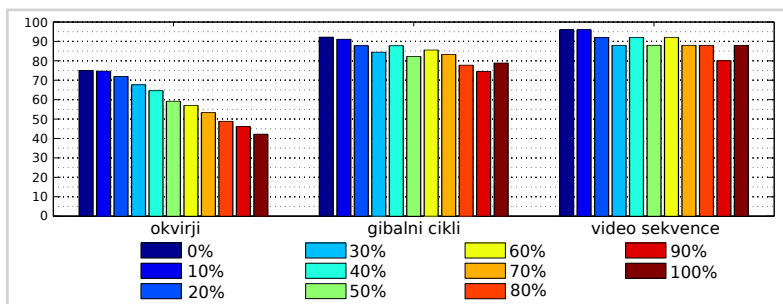


Slika 5.13

Primer signalov z dodanim Gaussovimi šumom.

Na sliki 5.14 je prikazana učinkovitost klasifikacije za različne deleže dodanega Gaussovega šuma za posamezne nivoje hierarhičnega združevanja rezultatov klasifikacije. Vpliv šuma na učinkovitost razpoznavanja je v tem primeru razvidna predvsem na najnižjem nivoju okvirjev. Podobno kot v primeru šuma sol in poper je možno tudi v tem primeru opaziti omiljenje učinkov šuma na višjih nivojih z bistvenim preskokom na nivoju klasifikacije ciklov gibanja. Za razliko od šuma sol in poper so negativni učinki šuma v tem primeru nekoliko milejši in je možno razpoznavanje celo pri dodanem Gaussovem šumu s standardno deviacijo enako kar standardni deviaciji signala (100 %).

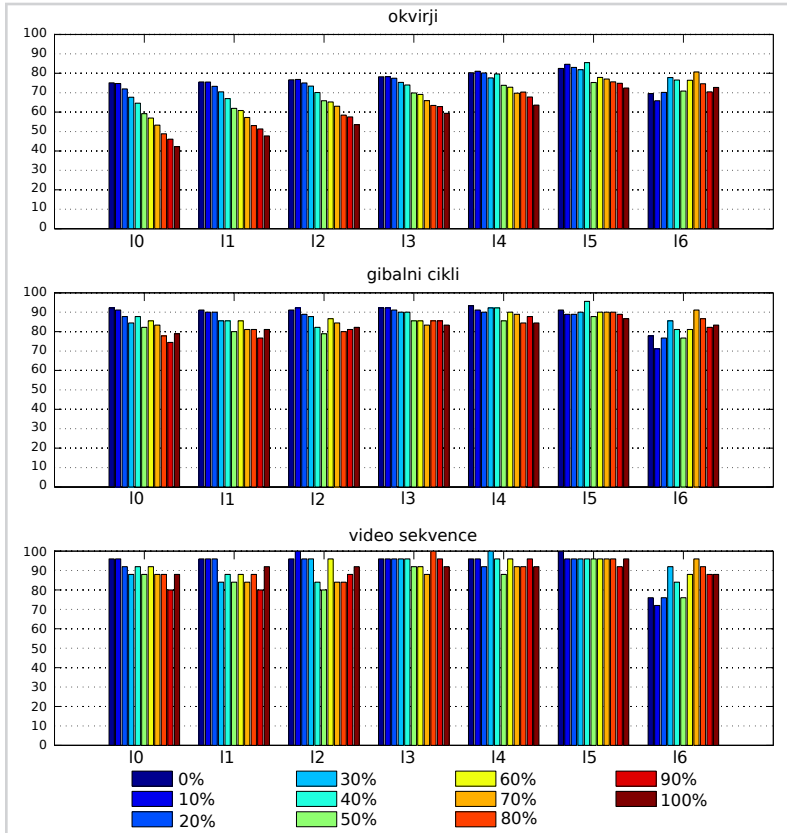
Dodatno je na sliki 5.15 prikazano, kako lahko vpliv šuma omilimo z uporabo valčkov. Podobno kot v prejšnjem primeru lahko izboljševanje učinkovitosti razpoznavanja opazimo na rezultatih valčkov nekje do nivoja 15, hkrati pa rezultati klasifikacije videosekvenc prikazujejo, da je možno negativen učinek Gaussovega šuma skoraj po-



Slika 5.14

Učinkovitost razpoznavanja pri prisotnosti Gaussovega šuma v signalih za različne nivoje hierarhičnega združevanja rezultatov.

polnoma izničili na nivoju valčkov 15. Torej je s pomočjo valčkov možno učinkovito odstraniti negativen vpliv tako šuma sol in poper kot tudi Gaussovega šuma.



Slika 5.15

Učinkovitost razpoznavanja pri prisotnosti Gaussovega šuma v signalih za različne nivoje aproksimacije z valčki (I0 do I6) na različnih nivojih hierarhičnega združevanja. I0 ustreza rezultatom brez aproksimacije.

Zaključek

V tem delu smo predstavili metodo za razpoznavanje ljudi po načinu hoje na osnovi modela človeškega telesa z glavnim poudarkom na značilkah, ki opisujejo dinamiko gibanja. Predlagamo nekatere rešitve za specifične probleme segmentacije skeleta, kot so: poravnava značilk na osnovi slikovnih momentov, ki je kritična za uspešno primerjavo signalov časovnih vrst, detekcija faze dvojne opore, detekcija križanja nog in fuzija klasifikatorjev. Predlagamo klasifikacijo s fuzijo značilk in klasifikacijo na osnovi okvirjev, ki skupaj z aproksimacijo signalov z valčki dosegata in presegata uspešnost delovanja sorodnih najboljših (angl. *state of the art*) metod področja.

Nadalje smo naslovili problem spremembe hitrosti hoje in predlagali normalizacijo hitrosti hoje na osnovi geometrijskih transformacij ter jo posplošili in nadgradili na transformacijo prostora značilk, ki je se je izkazala za zelo učinkovito za primer dinamičnih značilk, pridobljenih iz modela skeleta. Obe metodi omilita negativne učinke vpliva spremembe hitrosti hoje in s tem omogočata primerjavo subjektov, ki hodijo z različnimi hitrostmi.

Dodatno so značilke, predstavljene v tem delu, razumljive človeškemu strokovnjaku, npr. s področja medicine, kineziologije in športa, ter imajo kot take perspektivo za prispevanje novih vpogledov v analizo gibanja na omenjenih področjih.

Učinkovitost razpoznavanja razvitih postopkov smo evalvirali na zbirki OU–ISIR in pokazali, da lahko tudi pristopi na osnovi modela, ki se osredotočajo zgolj na opis dinamike gibanja, dosegajo primerljive ali celo boljše rezultate kot najboljše (angl. *state of the art*) metode področja, ki niso omejene le na dinamiko gibanja, ampak za razločevanje posameznikov uporabljajo tudi videz. Dodatno smo s podanimi eksperimenti prikazali tudi robustnost delovanja metode v prisotnosti različnih vrst motenj in šuma, ki so pogosto prisotne v aplikacijah računalniškega vida. S pomočjo teh smo določili tudi spodnjo mejo ločljivosti videosekvenc, zgornjo mejo prisotnosti šuma v silhuetah ter šuma v segmentiranem skeletu.

Kolikor vemo, je predlagana metoda edina v literaturi, ki temelji na modelu in hkrati učinkovito obravnava spremembe hitrosti hoje. Čeprav zbirka OU–ISIR vsebuje relativno majhno število subjektov (34), vsebuje posnetke za širok nabor različnih hitrosti hoje in je zato najprimernejša trenutno obstoječa zbirka za raziskave vpliva hitrosti hoje na gibanje. Splošna zmožnost razpoznavanja pristopov na osnovi modela je sicer že bila prikazana v literaturi (npr. [9]), kjer so avtorji dosegli 84 % učinkovitost razpoznavanja na 100 osebah, vendar ne pri spremenjenih pogojih, kot smo demonstrirali v tem delu.

6.1 Prispевki k znanosti

Glavni prispevki našega dela so:

1. Razvoj metode na osnovi modela skeleta človeškega telesa, iz katerega pridobimo značilke hoje, ki se osredotočajo predvsem na opis dinamike gibanja, s čimer dosežemo relativno neodvisnost od posameznikovega videza:
 - (a) Detekcija ciklov hoje na osnovi slikovnih momentov (angl. *image moments*), ki omogoča skoraj popolno poravnavo faze ciklov med različnimi videoposnetki, kar je bistvenega pomena pri klasifikaciji na osnovi razdalje z značilkami, opisanimi s signali časovnih vrst (angl. *time-series signals*).
 - (b) Fuzija značilk, ki omogoča uporabo več klasifikatorjev na osnovi posameznih značilk, katerih rezultati so združeni v končni klasifikator s pomočjo glasovanja.
 - (c) Klasifikacija na osnovi okvirjev, ki s pomočjo klasificiranja vzorcev signalov posameznih okvirjev videoposnetka omogoča neodvisnost od začetne točke cikla hoje.
2. Zmanjšanje vpliva spremembe hitrosti hoje na učinkovitost razpoznavanja:
 - (a) Normalizacija hitrosti hoje s pomočjo geometrijskih transformacij z aplikacijo v metodah brez modela.
 - (b) Transformacija prostora značilk, ki temelji na statističnem modeliranju prostorov hitrosti hoje z aplikacijo na dinamičnih značilkah predlagane metode modela skeleta človeškega telesa.

6.2 Nadaljnje delo

Klasifikacija na osnovi okvirjev v osnovi ne potrebuje deljenja signalov na cikle. Posplošitev delovanja metode na celotne videoposnetke bi močno poenostavila postopek obdelave signalov in s tem prispevala k robustnejšemu delovanju metode, pri tem pa ne bi bistveno vplivala na učinkovitost razpoznavanja.

Križno razpoznavanje med tekom in hojo predstavlja velik problem, saj se posameznikov tek bistveno razlikuje od hoje, tako po dinamiki gibanja kot po videzu. Identifikacija tekačev s pomočjo vzorcev galerije, zajete z normalno hitrostjo hoje, pa lahko

predstavlja velik prispevek k aplikacijam s področja nadzora in varnosti ter forenzičnim preiskavam.

Posplošitev transformacije prostora značilnik na druge faktorje vpliva, kot so obutev, obleka in sprememba pogleda, je vsaj do neke mere možna že v sedanji obliki. Za natančnejšo analizo bi bilo treba izvesti eksperimente in analize na zbirkah, ki obravnavajo omenjene faktorje.

Konec koncev lahko zaključimo, da današnje najboljše (angl. *state of the art*) metode področja, vključno z metodo, razvito v sklopu tega dela, predstavljajo velik potencial tako z vidika biometrične identifikacije kot tudi drugih aplikacijskih domen, ki uporabljajo analizo gibanja, a je praktična raba v realnem okolju še vedno mogoča le v redkih, zelo specifičnih primerih uporabe, s strogimi omejitvami glede okolja zajemanja videoposnetkov in gibanja posameznikov. Za splošnejšo rabo idealiziranega primera uporabe razpoznavanja ljudi na prostem, v množici, bo treba nasloviti še številne danes nerešene probleme, kot so spremembe gibanja zaradi obutve, obleke, vpliva staranja, bolezni ipd., ter nazadnje širše probleme področja računalniškega vida, kot so sprememba pogleda, sledenje objektov in segmentacija gibanja. Danes torej ta biometrika, ne predstavlja resne konkurence splošno razširjenim biometrikam kot so prstni odtisi, šarenice, obrazi ipd., lahko pa predstavlja dobro dopolnilno orodje za kombinacijo z drugimi biometrikami oz. kot sredstvo identifikacije v primerih, ko druge biometrike niso na voljo.

LITERATURA

- [1] G. Johansson. Visual motion perception. *Scientific American*, 232(6):76–88, 1976. doi: [10.1038/scientificamerican0675-76](https://doi.org/10.1038/scientificamerican0675-76).
- [2] J. E. Boyd and J. J. Little. Biometric gait recognition. In *Advanced Studies in Biometrics*, volume 3161 of *Lecture Notes in Computer Science*, pages 19–42, Berlin, Heidelberg, 2005. Springer-Verlag. doi: [10.1007/11493648_2](https://doi.org/10.1007/11493648_2).
- [3] I. Bouchrika and M. S. Nixon. People detection and recognition using gait for automated visual surveillance. In *Crime and Security*, pages 576–581, London, UK, 2006. IET.
- [4] B. Pogorelc, Z. Bosnić, and M. Gams. Automatic recognition of gait-related health problems in the elderly using machine learning. *Multimedia Tools and Applications*, 58(2):333–354, 2011. doi: [10.1007/s11042-011-0786-1](https://doi.org/10.1007/s11042-011-0786-1).
- [5] R. Bartlett. Artificial intelligence in sports biomechanics: New dawn or false hope? *Journal of Sports Science and Medicine*, 5(4):474–479, 2006.
- [6] T. B. Moeslund, A. Hilton, and V. Kruger. A survey of advances in vision-based human motion capture and analysis. *Computer Vision and Image Understanding*, 104(2-3):90–126, 2006. doi: [10.1016/j.cviu.2006.08.002](https://doi.org/10.1016/j.cviu.2006.08.002).
- [7] I. Bouchrika, J. N. Carter, M. S. Nixon, R. Morzinger, and G. Thallinger. Using gait features for improving walking people detection. In *Pattern Recognition*, pages 3097–3100. IEEE, 2010. doi: [10.1109/ICPR.2010.758](https://doi.org/10.1109/ICPR.2010.758).
- [8] G. Ariyanto and M. S. Nixon. Marionette mass-spring model for 3d gait biometrics. In *Biometrics*, pages 354–359, New Delhi, 2012. IEEE. doi: [10.1109/ICB.2012.6199832](https://doi.org/10.1109/ICB.2012.6199832).
- [9] J. Yoo and M. S. Nixon. Automated markerless analysis of human gait motion for recognition and classification. *ETRI Journal Information, Telecommunications and Electronics*, 33(2):259–266, 2011.
- [10] H. Lu, K. Plataniotis, and A. Venetsanopoulos. A full-body layered deformable model for automatic model-based gait recognition. *Journal of Advances in Signal Processing*, 2008:1–14, 2008. doi: [10.1155/2008/261317](https://doi.org/10.1155/2008/261317).
- [11] Z. Liu and S. Sarkar. Simplest representation yet for gait recognition: Averaged silhouette. In *International Conference on Pattern Recognition*, volume 4, pages 211–214. IEEE, 2004. doi: [10.1109/ICPR.2004.1333741](https://doi.org/10.1109/ICPR.2004.1333741).
- [12] S. Sarkar, P. J. Phillips, Z. Liu, I. R. Vega, P. Grother, and K. W. Bowyer. The HumanID gait challenge problem: data sets, performance, and analysis. *Transactions on Pattern Analysis and Machine Intelligence*, 27(2):162–177, 2005. doi: [10.1109/TPAMI.2005.39](https://doi.org/10.1109/TPAMI.2005.39).
- [13] Z. Wang, X. Sun, L. Sun, and Y. Huang. Manifold adaptive kernel semisupervised discriminant analysis for gait recognition. *Advances in Mechanical Engineering*, 2013:1–12, 2013.
- [14] M. Peternel and A. Leonardis. Visual learning and recognition of a probabilistic spatio-temporal model of cyclic human locomotion. In *International Conference on Pattern Recognition*, volume 4, pages 146–149. IEEE, 2004. doi: [10.1109/ICPR.2004.1333725](https://doi.org/10.1109/ICPR.2004.1333725).
- [15] W. Stone and M. Skubic. Capturing habitual, in-home gait parameter trends using an inexpensive depth camera. In *Engineering in Medicine and Biology Society*, pages 5106–5109. IEEE, 2012. doi: [10.1109/EMBC.2012.634714](https://doi.org/10.1109/EMBC.2012.634714).
- [16] T. Banerjee, J. M. Keller, and M. Skubic. Resident identification using Kinect depth image data and fuzzy clustering techniques. In *Engineering in Medicine and Biology Society*, pages 5102–5105. IEEE, 2012. doi: [10.1109/EMBC.2012.6347141](https://doi.org/10.1109/EMBC.2012.6347141).
- [17] M. Whittle. *Gait analysis: an introduction*. Butterworth-Heinemann, 4 edition, 2007.

- [18] M. R. Lim, R. C. Huang, A. Wu, F. P. Girardi, and F. P. Cammisa. Evaluation of the elderly patient with an abnormal gait. *The Journal of the American Academy of Orthopaedic Surgeons*, 15(2):107–117, 2007.
- [19] V. Boulgouris, D. Hatzinakos, and K. N. Plataniotis. Gait recognition: A challenging signal processing technology for biometric. *Signal Processing Magazine*, 22(6):78–90, 2005.
- [20] L. Liu and W. Jia. Survey of gait recognition. In *Emerging intelligent computing technology and applications*, volume 5755 of *Lecture Notes in Computer Science*, pages 652–659, Berlin, Heidelberg, 2009. Springer-Verlag.
- [21] M. S. Nixon, T. N. Tan, and R. Chellappa. *Human Identification based on Gait*. International Series on Biometrics. Springer-Verlag, New York, USA, 2005.
- [22] M. S. Nixon and J. N. Carter. Automatic recognition by gait. *Proceedings of the IEEE*, 94(11):2013–2024, 2006. doi: [10.1109/JPROC.2006.886018](https://doi.org/10.1109/JPROC.2006.886018).
- [23] J. Wang, M. She, S. Nahavandi, and A. Kouzani. A review of vision-based gait recognition methods for human identification. In *International Conference on Digital Image Computing: Techniques and Applications*, pages 320–327. IEEE, 2010. doi: [10.1109/DIC-TA.2010.62](https://doi.org/10.1109/DIC-TA.2010.62).
- [24] R. Tanawongsuwan and A. Bobick. Modelling the effects of walking speed on appearance-based gait recognition. In *Computer Vision and Pattern Recognition*, volume 2, pages 783–790. IEEE, 2004. doi: [10.1109/CVPR.2004.1315244](https://doi.org/10.1109/CVPR.2004.1315244).
- [25] C. Stauffer and W. E. Grimson. Adaptive background mixture models for real-time tracking. In *Computer Vision and Pattern Recognition*, volume 2, pages 2246–2252, 1999. doi: [10.1109/CVPR.1999.784637](https://doi.org/10.1109/CVPR.1999.784637).
- [26] N. V. Boulgouris and X. Huang. Gait recognition using hmms and dual discriminative observations for sub-dynamics analysis. *Transactions on Image Processing*, 22(9):3636–3647, 2013. doi: [10.1109/TIP.2013.2266578](https://doi.org/10.1109/TIP.2013.2266578).
- [27] X. Huang and N. Boulgouris. Model-based human gait recognition using fusion of features. In *Acoustics, Speech and Signal Processing*, pages 1469–1472. IEEE, 2009. doi: [10.1109/ICASSP.2009.4959872](https://doi.org/10.1109/ICASSP.2009.4959872).
- [28] S. U. Jung and M. S. Nixon. Heel strike detection based on human walking movement for surveillance analysis. *Pattern Recognition Letters*, 34(8):895–902, 2013. doi: [10.1016/j.patrec.2012.08.007](https://doi.org/10.1016/j.patrec.2012.08.007).
- [29] T. Yu and J. H. Zou. Automatic human gait imitation and recognition in 3d from monocular video with an uncalibrated camera. *Mathematical Problems in Engineering*, 2012:1–35, 2012.
- [30] X. He, D. Cai, S. Yan, and H. J. Zhang. Neighborhood preserving embedding. In *International Conference on Computer Vision*, volume 2, pages 1208–1213. IEEE, 2005. doi: [10.1109/ICCV.2005.167](https://doi.org/10.1109/ICCV.2005.167).
- [31] A. Quattoni, S. Wang, L. P. Morency, M. Collins, and T. Darrell. Hidden conditional random fields. *Transactions on pattern analysis and machine intelligence*, 29(10):1848–1853, 2007. doi: [10.1109/TPAMI.2007.1124](https://doi.org/10.1109/TPAMI.2007.1124).
- [32] N. V. Boulgouris and Z. X. Chi. Human gait recognition based on matching of body components. *Pattern Recognition*, 40(6):1763–1770, 2007. doi: [10.1016/j.patcog.2006.11.012](https://doi.org/10.1016/j.patcog.2006.11.012).
- [33] X. Li, S. J. Maybank, S. Yan, D. Tao, and D. Xu. Gait components and their application to gender recognition. *Transactions on Systems, Man, and Cybernetics: Applications and Reviews*, 38(2):145–155, 2008.
- [34] I. Venkat and P. Wilde. Robust gait recognition by learning and exploiting sub-gait characteristics. *International Journal of Computer Vision*, 91(1):7–23, 2010. doi: [10.1007/s11263-010-0362-6](https://doi.org/10.1007/s11263-010-0362-6).
- [35] M. A. Giese and T. Poggio. Neural mechanisms for the recognition of biological movements. *Neuroscience*, 4(3):179–192, 2003. doi: [10.1038/nrn1057](https://doi.org/10.1038/nrn1057).
- [36] M. Escobar and P. Kornprobst. Action recognition via bio-inspired features: The richness of center-surround interaction. *Computer Vision and Image Understanding*, 116(5):593–605, 2012. doi: [10.1016/j.cviu.2012.01.002](https://doi.org/10.1016/j.cviu.2012.01.002).
- [37] H. Jhuang, T. Serre, and L. Wolf. A biologically inspired system for action recognition. In *International Conference on Computer Vision*, pages 1–8. IEEE, 2007. doi: [10.1109/ICCV.2007.4408988](https://doi.org/10.1109/ICCV.2007.4408988).
- [38] Y. Mu. Biologically inspired feature manifold for gait recognition. *Neurocomputing*, 73(4-6):895–902, 2010. doi: [10.1016/j.neucom.2009.09.017](https://doi.org/10.1016/j.neucom.2009.09.017).
- [39] K. Bashir, T. Xiang, and S. Gong. Gait recognition without subject cooperation. *Pattern Recognition Letters*, 31(13):2052–2060, 2010. doi: [10.1016/j.patrec.2010.05.027](https://doi.org/10.1016/j.patrec.2010.05.027).
- [40] S. Singh and K. Biswas. Biometric gait recognition with carrying and clothing variants. In *Pattern Recognition and Machine Intelligence*, pages 446–451, Berlin, Heidelberg, 2009. Springer-Verlag. doi: [10.1007/978-3-642-11164-8_72](https://doi.org/10.1007/978-3-642-11164-8_72).
- [41] W. Kusakunniran and Q. Wu. Support vector regression for multi-view gait recognition based on local motion feature selection. In *Computer vision and pattern recognition*, pages 974–981, 2010.

- [42] W. Kusakunniran, Q. Wu, and J. Zhang. Cross-view and multi-view gait recognitions based on view transformation model using multi-layer perceptron. *Pattern Recognition Letters*, 33(7):882–889, 2011. doi: [10.1016/j.patrec.2011.04.014](https://doi.org/10.1016/j.patrec.2011.04.014).
- [43] Y. Makihara, R. Sagawa, and Y. Mukaigawa. Gait recognition using a view transformation model in the frequency domain. In *European Conference on Computer Vision*, pages 151–163, Berlin, Heidelberg, 2006. Springer-Verlag. doi: [10.1007/11744078_12](https://doi.org/10.1007/11744078_12).
- [44] A. Tsuji, Y. Makihara, and Y. Yagi. Silhouette transformation based on walking speed for gait identification. In *International Conference on Computer Vision and Pattern Recognition*, pages 717–722, 2010.
- [45] H. Lee, S. Hong, and E. Kim. An efficient gait recognition with backpack removal. *Journal on Advances in Signal Processing*, 2009:1–7, 2009. doi: [10.1155/2009/384384](https://doi.org/10.1155/2009/384384).
- [46] R. Martin-Felez and T. Xiang. Gait recognition by ranking. In *European Conference on Computer Vision*, pages 328–341, 2012.
- [47] Z. Liu and S. Sarkar. Improved gait recognition by gait dynamics normalization. *Transactions on pattern analysis and machine intelligence*, 28(6):863–76, 2006.
- [48] S. Lee and R. Collins. Shape variation-based frieze pattern for robust gait recognition. In *Computer Vision and Pattern Recognition*, pages 1–8. IEEE, 2007. doi: [10.1109/CVPR.2007.383138](https://doi.org/10.1109/CVPR.2007.383138).
- [49] A. Veeraraghavan, A. Srivastava, A. K. Roy-Chowdhury, and R. Chellappa. Rate-invariant recognition of humans and their activities. *Transactions on image processing*, 18(6):1326–39, 2009.
- [50] T. Kobayashi and N. Otsu. Three-way auto-correlation approach to motion recognition. *Pattern Recognition Letters*, 30(3), 2009. 212–221.
- [51] W. Kusakunniran, Q. Wu, J. Zhang, and H. Li. Speed-invariant gait recognition based on procrustes shape analysis using higher-order shape configuration. In *Image Processing*, pages 545–548. IEEE, 2011. doi: [10.1109/ICIP.2011.6116403](https://doi.org/10.1109/ICIP.2011.6116403).
- [52] W. Kusakunniran, Q. Wu, J. Zhang, and H. Li. Differential composition model. *Systems, man, and Cybernetics*, 42(6):1654–1668, 2012.
- [53] Y. Iwashita, R. Baba, K. Ogawara, and R. Kurazume. Person identification from spatio-temporal 3D gait. In *International Conference on Emerging Security Technologies*, pages 30–35. IEEE, 2010. doi: [10.1109/EST.2010.19](https://doi.org/10.1109/EST.2010.19).
- [54] E. Ribnick, R. Sivalingam, N. Papanikolopoulos, and K. Daniilidis. Reconstructing and analyzing periodic human motion from stationary monocular views. *Computer Vision and Image Understanding*, 116(7):816–826, 2012. doi: [10.1016/j.cviu.2012.03.004](https://doi.org/10.1016/j.cviu.2012.03.004).
- [55] K. Bashir, T. Xiang, and S. Gong. Cross view gait recognition using correlation strength. In *British Machine Vision Conference*, pages 109.1–109.11. BMVA Press, 2010. doi: [10.5244/C.24.109](https://doi.org/10.5244/C.24.109).
- [56] M. Altab Hossain, Y. Makihara, J. Wang, and Y. Yagi. Clothing-invariant gait identification using part-based clothing categorization and adaptive weight control. *Pattern Recognition*, 43(6):2281–2291, 2010. doi: [10.1016/j.patcog.2009.12.020](https://doi.org/10.1016/j.patcog.2009.12.020).
- [57] H. Iwama, M. Okumura, Y. Makihara, and Y. Yagi. The ou-isir gait database comprising the large population dataset and performance evaluation of gait recognition. *Transactions on Information Forensics and Security*, 7(5):1511–1521, 2012. doi: [10.1109/TIFS.2012.2204253](https://doi.org/10.1109/TIFS.2012.2204253).
- [58] R. Gross and J. Shi. The CMU motion of body (MO-BO) database. Technical Report CMU-RI-TR-01-18, Robotics Institute, Pittsburgh, PA, 2001.
- [59] J. D. Shutler, M. G. Grant, M. S. Nixon, and J. N. Carter. On a large sequence-based human gait database. In Ahamad Lotfi and Jonathan M. Garibaldi, editors, *Applications and Science in Soft Computing*, volume 24 of *Advances in Soft Computing*, pages 339–346. Springer-Verlag, Berlin, Heidelberg, 2004.
- [60] S. Yu, D. Tan, and T. Tan. A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition. In *International Conference on Pattern Recognition*, pages 441–444. IEEE, 2006. doi: [10.1109/ICPR.2006.67](https://doi.org/10.1109/ICPR.2006.67).
- [61] D. Tan, K. Huang, S. Yu, and T. Tan. Efficient night gait recognition based on template matching. In *International Conference on Pattern Recognition*, pages 1000–1003. IEEE, 2006. doi: [10.1109/ICPR.2006.478](https://doi.org/10.1109/ICPR.2006.478).
- [62] Y. Makihara, H. Mannami, A. Tsuji, M. A. Hossain, K. Sugiura, A. Mori, and Y. Yagi. The OU-ISIR gait database comprising the treadmill dataset. *Transactions on Computer Vision and Applications*, 4:53–62, 2012.
- [63] M. Hofmann, S. Sural, and G. Rigoll. Gait recognition in the presence of occlusion: A new dataset and baseline algorithms. In *Computer Graphics, Visualization and Computer Vision*, pages 99–104, 2011.
- [64] L. Zappella, X. Llado, and J. Salvi. *Pattern Recognition*, chapter New Trends in Motion Segmentation. InTech, 2009.
- [65] D. A. Winter. *The biomechanics and motor control of human movement*. John Wiley & Sons, 1990.

- [66] M. P. Murray. Gait as a total pattern of movement. *American Journal of Physical Medicine*, 46(1):290–333, 1967.
- [67] D. Weinland, R. Ronfard, and E. Boyer. Free viewpoint action recognition using motion history volumes. *Computer Vision and Image Understanding*, 104(2): 249–257, 2006. doi: [10.1016/j.cviu.2006.07.013](https://doi.org/10.1016/j.cviu.2006.07.013).
- [68] A. Bissacco, A. Chiuso, and S. Soatto. Classification and recognition of dynamical models: the role of phase, independent components, kernels and optimal transport. *Transactions on Pattern Analysis and Machine Intelligence*, 29(11):1958–1972, 2007. doi: [10.1109/TPAMI.2007.1101](https://doi.org/10.1109/TPAMI.2007.1101).
- [69] J. E. Cutting, D. R. Proffitt, and L. T. Kozlowski. A biomechanical invariant for gait perception. *Journal of experimental psychology. Human perception and performance*, 4(3):357–72, 1978.
- [70] M. Aung, S. Thies, L. Kenney, D. Howard, R. Selles, A. Findlow, and J. Goulermas. Automated detection of instantaneous gait events using time frequency analysis and manifold embedding. *Transactions on neural systems and rehabilitation engineering*, 21(6), 2013. doi: [10.1109/TNSRE.2013.2239313](https://doi.org/10.1109/TNSRE.2013.2239313).
- [71] S. Rahati, R. Moravejian, and F. M. Kazemi. Gait recognition using wavelet transform. In *International Conference on Information Technology: New Generations*, pages 932–936. IEEE, 2008. doi: [10.1109/ITNG.2008.124](https://doi.org/10.1109/ITNG.2008.124).
- [72] Z. Xue, D. Ming, W. Song, B. Wan, and S. Jin. Infrared gait recognition based on wavelet transform and support vector machine. *Pattern Recognition*, 43(8): 2904–2910, 2010. doi: [10.1016/j.patcog.2010.03.011](https://doi.org/10.1016/j.patcog.2010.03.011).
- [73] J. Kovač and P. Peer. Transformation based walking speed normalization for gait recognition. *KSIJ Transactions on Internet and Information Systems*, 7(11): 2690–2701, 2013.
- [74] N. Pavešić. *Razpoznavanje vzorcev: uvod v analizo in razumevanje vidnih in slušnih signalov*. Fakulteta za elektrotehniko, Ljubljana, Slovenija, 2000.
- [75] J. Kovač and P. Peer. Human skeleton model based dynamic features for walking speed invariant gait recognition. *Mathematical Problems in Engineering*, 2014:1–15, 2014. ISSN 1024-123X. doi: [10.1155/2014/484320](https://doi.org/10.1155/2014/484320).
- [76] C. Yam, M. S. Nixon, and J. N. Carter. Automated person recognition by walking and running via model-based approaches. *Pattern Recognition*, 37(5): 1057–1072, 2004. doi: [10.1016/j.patcog.2003.09.012](https://doi.org/10.1016/j.patcog.2003.09.012).
- [77] M. Ye, Q. Zhang, L. Wang, J. Zhu, R. Yang, and J. Gall. A survey on human motion analysis from depth data. In *Time-of-Flight and Depth Imaging. Sensors, Algorithms, and Applications*, volume 8200 of *Lecture Notes in Computer Science*, pages 149–187, Berlin, Heidelberg, 2013. Springer-Verlag. doi: [10.1007/978-3-642-44964-2_8](https://doi.org/10.1007/978-3-642-44964-2_8).
- [78] M. Brubaker, D. J. Fleet, and A. Hertzmann. Physics-based person tracking using the anthropomorphic walker. *International Journal of Computer Vision*, 87(1-2):140–155, 2009. doi: [10.1007/s11263-009-0274-5](https://doi.org/10.1007/s11263-009-0274-5).
- [79] Z. Zhang. Microsoft Kinect sensor and its effect. *Multimedia*, 19(2):4–10, 2012. doi: [10.1109/MMUL.2012.24](https://doi.org/10.1109/MMUL.2012.24).