

UNIVERZITET U BEOGRADU
ELEKTROTEHNIČKI FAKULTET

DIPLOMSKI RAD

Analiza glavnih komponenti elektromiografskih
signala za klasifikaciju pokreta šake

mentor

prof. dr Mirjana B. Popović

komentor

dr Nadica S. Miljković

kandidat

Milica S. Isaković
2010/0032

Beograd, septembar 2014.



PREDGOVOR

Signali koji su korišćeni i analizirani u ovom diplomskom radu su snimljeni u okviru NinaPro projekta (<http://www.idiap.ch/project/ninapro/>). Signali su javno dostupni u bazi podataka na internet stranici projekta (<http://ninapro.hevs.ch/>), odakle su i preuzeti.

Diplomski rad ima 40 strana, 20 slika i jednu tabelu. Lista referenci sadrži 24 bibliografske jedinice. Na kraju rada je u Prilogu A dat sažet opis kodova i funkcija projektovanih u programskom paketu MATLAB, kao i uputstvo za korišćenje projektovanog korisničkog interfejsa.

REZIME RADA

Jedan od osnovnih ciljeva neuralnog inženjerstva je projektovanje proteza namenjenih osobama nakon amputacije ekstremiteta, kako bi se u što je moguće većoj meri zamenio amputirani deo tela i omogućilo obavljanje svakodnevnih aktivnosti. Prvi korak u razvoju proteze predstavlja akvizicija i analiza signala pomoću kojih će se upravljati protezom. Ovaj rad se bavi upravo definisanjem kontrolnih signala za potrebe upravljanja veštačkom šakom. Konkretno, akcenat je na pretprocesiranju signala i analizi glavnih komponenti (eng. *Principal Component Analysis – PCA*) sa ciljem detekcije pojedinih pokreta. Osim toga, razmatrana je i klasifikacija na izdvojenom setu pokreta.

Električni signali mereni na mišićima podlaktice (elektromiografski – EMG signali) tokom pokreta su preuzeti iz baze podataka NINAPRO projekta [1]. U ovom diplomskom radu je dat opis projekta i baze podataka, akvizicije podataka, izvedenih pokreta i protokola merenja. Signali iz baze su originalni, neobrađeni, te ih je najpre potrebno pretprocesirati. Ovaj korak je neophodan radi pripreme signala za dalju analizu i sastoji se od sinhronizacije, filtriranja, segmentacije i usrednjavanja. Obradeni signali su nakon pretprocesiranja podesni za primenu PCA. U diplomskom radu je dat kratak istorijat analize glavnih komponenti, definisane su glavne komponente i njihove osobine, opisana je primena metode. Opisan je kvadratni klasifikator koji je primenjen za klasifikaciju pokreta. Rezultati pretprocesiranja, analize glavnih komponenti (sa varijacijama broja pokreta i ispitanika) i klasifikacije pokreta (za izdvojeni set pokreta) realizovani u programskom paketu MATLAB (The Mathworks, Natick, USA) su grafički prikazani. Dobijeni rezultati analize glavnih komponenti i klasifikacije pokreta su upoređeni sa rezultatima iz literature [2] i dati su opšti zaključci. Predstavljena su ograničenja primenjenih metoda, opisani su problemi i predlozi rešenja. Konačno, data su finalna razmatranja – predlozi za unapređenje primenjenih metoda i dalja istraživanja u ovoj oblasti. Sažet opis kodova i funkcija projektovanih u MATLAB-u, kao i uputstvo za korišćenje projektovanog korisničkog interfejsa su prikazani u Prilogu A.



ZAHVALNICA

Veliku zahvalnost dugujem svom mentoru prof. dr Mirjani Popović za predlog ove teme, korisne savete i sugestije, kao i za sve što sam do sada naučila o akviziciji i analizi elektrofizioloških signala. Ta znanja su mi jako puno značila tokom izrade diplomskog rada, i verujem da će mi biti korisna u daljem školovanju i budućem radu.

Zahvaljujem se svom komentoru dr Nadici Miljković na svim korisnim savetima, konsultacijama, želji i volji da odgovori na sva pitanja i pomogne da saznam što više, kako tokom izrade ovog diplomskog rada, tako i na svim predmetima iz oblasti biomedicinskog inženjerstva tokom osnovnih studija.

Svim kolegama i koleginicama sa Odseka za signale i sisteme hvala na originalnim idejama, savetima i lepom druženju tokom prethodnih godina.

Na kraju, želela bih da se zahvalim majci Gordani, ocu Srđanu i sestri Jeleni na bezuslovnoj podršci, ljubavi i razumevanju. Posebno bih istakla sestru, koja me je više od bilo koga drugog naučila o snazi, volji, istrajnosti i prevazilaženju prepreka i predstavlja moju najveću inspiraciju.

Milica S. Isaković

U Beogradu, septembar 2014.

SADRŽAJ

PREDGOVOR	2
REZIME RADA.....	3
ZAHVALNICA.....	4
SADRŽAJ	5
1 UVOD	7
2 METODOLOGIJA RADA	8
2.1 Baza snimljenih signala NINAPRO	8
2.2 Pretprocesiranje signala.....	13
2.3 Analiza glavnih komponenti.....	14
2.3.1 Uvod.....	14
2.3.2 Definicija i osobine	15
2.3.3 Primena PCA	18
2.4 Klasifikacija pokreta.....	21
3 REZULTATI.....	22
3.1 Rezultati pretprocesiranja	22
3.2 Rezultati PCA	23
3.3 Rezultati klasifikacije pokreta	27
4 DISKUSIJA I ZAKLJUČCI.....	29
4.1 Diskusija dobijenih rezultata i opšti zaključci	29
4.2 Ograničenja primenjenih metoda.....	31
4.3 Predlog za budući rad	32

Milica S. Isaković

Analiza glavnih komponenti elektromiografskih signala za klasifikaciju pokreta šake



5 LITERATURA.....	34
PRILOG A – Sažet opis projektovanih kodova i funkcija i uputstvo za upotrebu korisničkog interfejsa projektovanog u programskom paketu MATLAB	37

1 UVOD

Svakodnevni život ljudi se nakon amputacije šake u potpunosti menja. Uprkos napretku u oblasti neuralnog inženjerstva, neuralne proteze još uvek imaju do tri stepena slobode, skupe su, robusne i estetski nedovoljno usavršene, a proces učenja upravljanja protezom pomoću površinskih elektromiografskih signala (eng. *Surface Electromyography* – sEMG) je kompleksan, tako da je moguće realizovati samo ograničen broj jednostavnih pokreta i hvatova. Zbog svega toga, veliki broj amputiraca ne koristi redovno proteze, pa postoji potreba i veliko interesovanje za rad na njihovom poboljšanju, posebno na polju upravljanja protezama. Uprkos tehničkim i tehnološkim unapređenjima, rezultati studija pokazuju da tokom prethodnih decenija nije bilo napretka na polju proteza za šaku – nove generacije proteza su dale lošije rezultate od onih testiranih 1987. godine [3]. Ovo je posebno izraženo za proteze u obliku šake (proteze u obliku kuke su dale bolje rezultate), koje su teže i neudobnije, ali estetski poželjnije. Kao prvi korak potrebno je usavršiti postupke pretprocesiranja i klasifikacije signala da bi se sa što većim uspehom mogao definisati i projektovati željeni pokret. Sa tim ciljem je nastala NINAPRO baza podataka [2], iz koje su preuzeti signali koji su u daljem radu analizirani.

Metod glavnih komponenti je *metod multivarijacione analize koji se koristi za smanjenje dimenzije skupa podataka (sačinjava ga veliki broj uzajamno korelisanih promenljivih) uz istovremeno zadržavanje maksimalnog mogućeg varijabiliteta koji je prisutan u tim podacima* [4]. Iz same definicije se može uočiti da je ova metoda pogodna za višedimenzionalni set podataka koji je na raspolaganju – vrednosti sEMG signala za veliki broj elektroda, ispitanika i pokreta. Veliki broj radova ukazuje na to da je PCA odličan izbor za pripremu EMG signala za klasifikaciju [5], [6], [7].

Ovaj rad prikazuje analizu sEMG signala koja prethodi klasifikaciji – pretprocesiranje i primena PCA, kao i rezultate klasifikacije pokreta za izdvojeni set podataka. Cilj rada je pokazati da primena PCA povećava separabilnost između različitih pokreta i olakšava vizuelizaciju i klasifikaciju. Da bi se omogućilo poređenje dobijenih rezultata, razmatrane su iste grupe pokreta kao u [2], ali i druge grupe pokreta za koje smatramo da mogu biti reprezentativne i dati zadovoljavajuće rezultate.

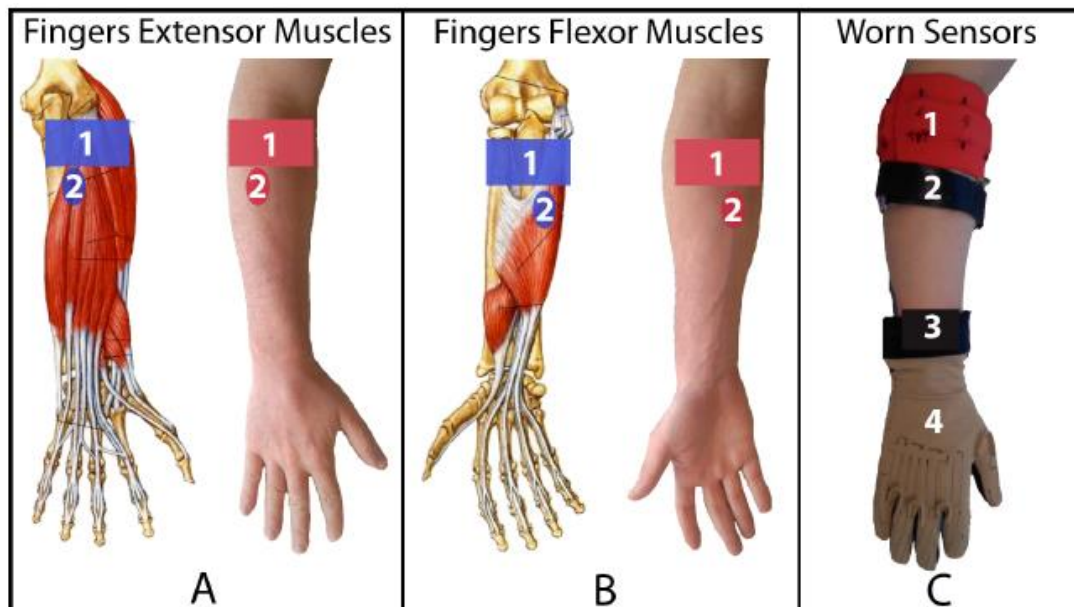
2 METODOLOGIJA RADA

2.1 Baza snimljenih signala NINAPRO¹

NINAPRO baza podataka sadrži sEMG i kinematičke signale sa 27 zdravih (neamputiranih) ispitanika koji izvode 52 različita pokreta, podatke o subjektima, protokole merenja i akvizicije, dobijene rezultate nakon pretprocesiranja i primena metoda mašinskog učenja za klasifikaciju pokreta [2]. Baza je nastala kao rezultat projekta NinaPro (eng. *Non-Invasive Adaptive Hand Prothesis*), koji traje od početka 2011. godine sa ciljem da se razviju familije algoritama koji bi povećali efikasnost proteza kontrolisanih sEMG signalima i skratili vreme obučavanja [8].

Akvizicija signala koji potiču od mišićne aktivnosti (sEMG) je izvršena pomoću deset *MyoBock 13E200* (Otto Bock HealthCare GmbH, Duderstadt, Germany) površinskih elektroda, koje daju pojačan, filtriran filtrom propusnikom opsega i ispravljen signal u opsegu 0-25 Hz. Osam uniformno raspoređenih elektroda je postavljeno ispod lakta, na jednakim udaljenostima od zgloba, dok su dve elektrode postavljene na mišiće ekstenzore i fleksore prstiju. Raspored elektroda je prikazan na Sl. 2.1.A i Sl. 2.1.B. Osim mišićne aktivnosti, baza sadrži i kinematičke signale snimljene sa *Cyberglove II dataglove* (Cyberglove Systems LLC, San Jose, CA, USA), rukavicom sa 22 senzora koji snimaju pozicije prstiju. Pored toga, subjekat tokom snimanja na ručnom zglobu nosi inklinometar za snimanje orijentacije zgloba. Celokupna postavka senzora je prikazana na Sl. 2.1.C.

¹ U ovom odeljku je dat sažet prikaz NinaPro projekta i baze podataka, akvizicionog postupka i protokola merenja u meri u kojoj je to neophodno za ovaj rad. Za detalje pogledati [2] i [4].



Sl. 2.1 Položaj elektroda: A. sEMG elektrode na mišićima ekstenzorima prstiju (A.1 Osam uniformno raspoređenih elektroda, A.2 Dve dodatne elektrode); B. sEMG elektrode na mišićima fleksorima prstiju (B.1 Osam uniformno raspoređenih elektroda, B.2 Dve dodatne elektrode); C. Svi senzori (C.1 Osam uniformno raspoređenih elektroda, C.2 Dve dodatne elektrode, C.3 Inklinometar, C.4 Cyberglove II dataglove). Preuzeto iz [2].

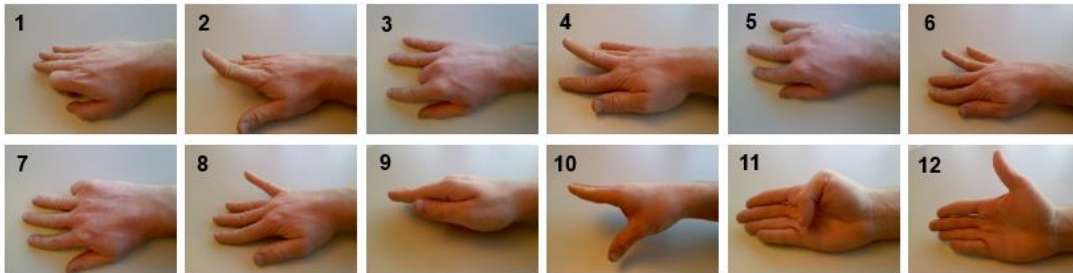
Ispitanici su izvodili ukupno 52 pokreta, podeljena u četiri klase:

- 12 osnovnih pokreta prstiju – fleksija (savijanje)/ekstenzija (opružanje) i adukcija (primicanje)/abdukcija (odmicanje)
- 8 izometrijskih i izotoničnih položaja šake
- 9 osnovnih pokreta ručnog zgloba – fleksija (savijanje)/ekstenzija (opružanje), adukcija (primicanje)/abdukcija (odmicanje) i supinacija (izvrtanje)/pronacija (uvrtanje)
- 23 hvata i funkcionalna pokreta – hvatanje predmeta sa kojima se ljudi susreću u svakodnevnim aktivnostima

Slike 2.2 – 2.5 grafički prikazuju pokrete iz sve četiri klase, dok su opisi pojedinačnih pokreta dati u Tabeli 1.

Tabela 1. Opisi svih 52 pokreta iz protokola NINAPRO baze

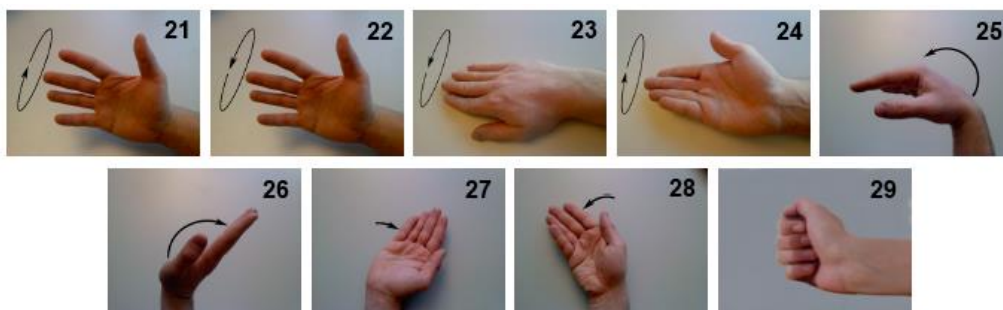
	Redni broj	Opis pokreta
Pokreti prstiju	1-2	Fleksija (savijanje) i ekstenzija (opružanje) kažiprsta
	3-4	Fleksija (savijanje) i ekstenzija (opružanje) srednjeg prsta
	5-6	Fleksija (savijanje) i ekstenzija (opružanje) domalog prsta
	7-8	Fleksija (savijanje) i ekstenzija (opružanje) malog prsta
	9-10	Adukcija (primicanje) i abdukcija (odmicanje) palca
	11-12	Fleksija (savijanje) i ekstenzija (opružanje) palca
Položaji šake	13	Palac gore
	14	Fleksija (savijanje) palca, domalog i malog prsta
	15	Fleksija (savijanje) domalog i malog prsta
	16	Palac savijen ka osnovi malog prsta
	17	Svi prsti rašireni i opruženi
	18	Svi prsti savijeni zajedno (pesnica)
	19	Ispružen kažiprst
	20	Svi prsti zajedno skupljeni i opruženi
Pokreti zgloba	21-22	Supinacija i pronacija zgloba – osa rotacije prolazi kroz srednji prst
	23-24	Supinacija i pronacija zgloba – osa rotacije prolazi kroz mali prst
	25-26	Fleksija i ekstenzija zgloba
	27-28	Radikalna i ularna devijacija
	29	Ekstenzija zgloba sa skupljenom šakom
Hvatovi	30-31	Predmeti velikog i malog prečnika
	32	Fiksirani „kukasti“ hvat
	33	Ekstenzija kažiprsta
	34	Obuhvatanje predmeta srednjeg prečnika
	35	Palac i kažiprst formiraju prsten
	36	Četiri prsta (sem palca) u formi prizme
	37	Štap
	38	Pisanje (hvat olovke sa tri prsta)
	39-41	Loptica – celom šakom, sa tri prsta i precizan hvat
	42	Loptica – precizan hvat sa tri prsta
	43-44	Prstohvat – u formi prizme i standardan
	45	Hvat sa 4 prsta
	46	Bočni (lateralni) hvat
	47-48	Paralelna ekstenzija i fleksija
	49	Disk celom šakom
	50	Otvaranje flaše sa tri prsta
	51	Okretanje šrafcižera (hvat 37)
	52	Sečenje (hvat 33)



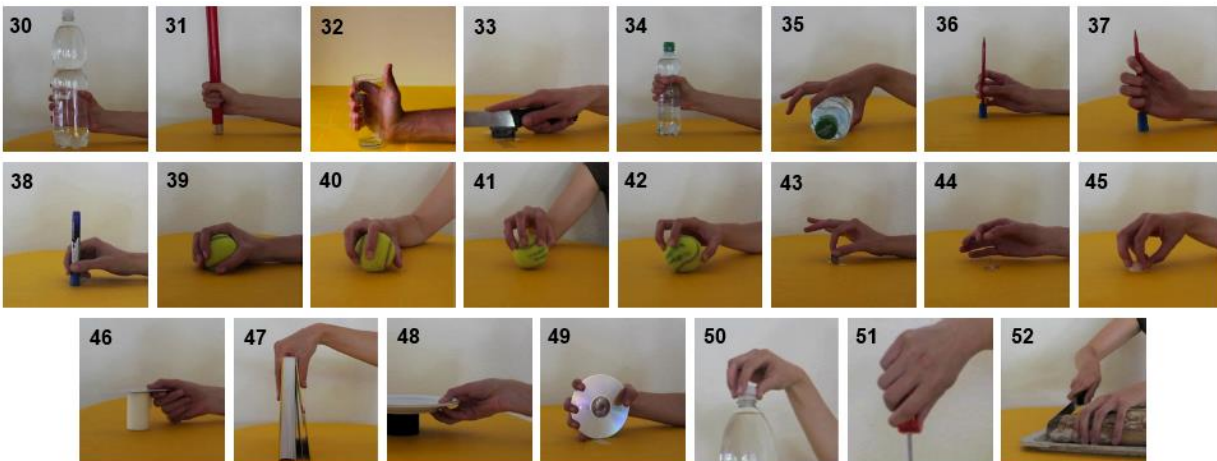
Sl. 2.2 Prva klasa pokreta – pokreti prstiju. Preuzeto i modifikovano iz [2].



Sl. 2.3 Druga klasa pokreta – položaji šake. Preuzeto i modifikovano iz [2].



Sl. 2.4 Treća klasa pokreta – pokreti zgloba. Preuzeto i modifikovano iz [2].



Sl. 2.5 Četvra klasa pokreta – hvatovi i funkcionalni pokreti. Preuzeto i modifikovano iz [2].

U toku merenja, ispitanici sa postavljenim sEMG elektrodama, rukavicom za snimanje kinematičkih signala i inklinometrom na desnoj ruci sede ispred ekrana na kome se puštaju kratki filmovi sa pokretima koje treba što tačnije da izvedu. Svaki ispitanik najpre prolazi kroz fazu treninga, tokom koje se ne snimaju signali. Nakon toga sledi snimanje, tokom koga se izvodi deset ponavljanja svakog pokreta. Svaki pokret traje pet sekundi, a između ponavljanja pokreta je pauza u trajanju od tri sekunde. Četiri klase pokreta su grupisane u tri vežbanja (druga i treća klasa zajedno čine drugo vežbanje) koja traju 16, 23 i 31 minut, respektivno. Da bi se izbegao zamor mišića, koji negativno utiče na kvalitet sEMG signala, postoje petominutne pauze između faze treninga i snimanja, kao i između pojedinačnih vežbanja.

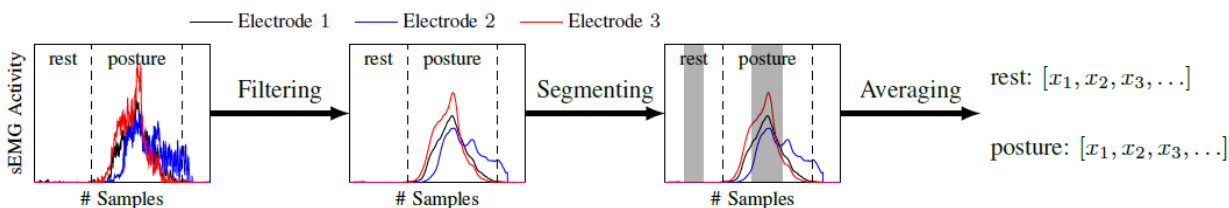
sEMG elektrode i inklinometar su povezani sa DAQ karticom (učestanost odabiranja je 100 Hz), koja je povezana sa laptop računaru. Ovi podaci su smešteni u „emg.txt“ fajlovima, za svako vežbanje i svakog ispitanika. Podaci o stimulusu (redni broj pokreta koji je izveden) imaju učestanost 25 Hz i nalaze se u „movie.txt“ fajlovima. Podaci sa rukavice, koja je sa računaru povezana preko *Bluetooth*-a na učestanosti od 25 Hz, su u „glove.txt“ fajlovima. Podaci nisu pretprocesirani niti sinhronizovani – svaki odbirak ima svoje podatke o datumu i vremenu.

Prva verzija NINAPRO baze, koja je razmatrana u ovom radu, sadrži signale sa 27 ispitanika koji nisu imali amputaciju (pol: 20 M/7 Ž, dominantna ruka: 25 desna/2 leva, starost: $28,0 \pm 3,4$ god.). Podaci o ispitanicima se mogu naći na [1].

2.2 Pretprocesiranje signala

S obzirom na različite učestanosti odabiranja sEMG signala (100 Hz) i podataka o stimulusu (25 Hz), početni korak pretprocesiranja je sinhronizacija podataka. Podaci o stimulusu su prvo dopunjeni nulama, produženi ili skraćeni, u zavisnosti od konkretnog slučaja, da bi obuhvatali isti vremenski period kao sEMG signali. Nakon toga sledi linearna interpolacija na 100 Hz. Dalji tok pretprocesiranja sEMG signala je prikazan na Sl. 2.6 i sastoji se od tri koraka:

- **Filtriranje** – signali mišićne aktivnosti se filtriraju niskopropusnim *Butterworth*-ovim filtrom drugog reda, sa graničnom učestanošću jednakom 1 Hz. Samo niskofrekventna komponenta signala, koja ostaje nakon filtriranja, je od interesa za dalju analizu.
- **Segmentacija** – Da bi se uklonile „dvosmislenosti“ u signalu koje su rezultat prelaza između pauze i pokreta, svako izvođenje pokreta i pauze je podeljeno na tri jednaka segmenta. Zbog distinktivnosti se zadržavaju samo centralni segmenti svih ponavljanja.
- **Usrednjavanje** – Odbirci izdvojenih centralnih segmenata se usrednjavaju. Nakon ovoga, svaki pokret za svakog subjekta sadrži 20 odbiraka (po dva za svako od 10 ponavljanja), od kojih polovina pripada pokretima, a druga polovina pauzama između ponavljanja pokreta.



Sl. 2.6 Tok pretprocesiranja sEMG signala – filtriranje, segmentacija, usrednjavanje. Preuzeto iz [2].

Kao poslednji korak sledi normalizacija signala (za svakog ispitanika posebno), tako da svaki sEMG signal ima nultu srednju vrednost i jediničnu standardnu devijaciju. Ovo je zapravo prvi korak u primeni PCA.

2.3 Analiza glavnih komponenti

2.3.1 Uvod

PCA je prvi definisao Karl Pearson 1901. godine po analogiji sa glavnim osama u mehanici [9], ali je opis izračunavanja i naziv pod kojim je metoda danas poznata dao Harold Hotelling 30-ih godina prošlog veka [10]. Prva izračunavanja su prikazana još 1958. godine, mada su svi rani radovi bili matematički orijentisani (detaljniji istorijat u [11]). Široka primena ove metode je počela sa razvojem računara, koji su omogućili da ona ostvari svoj potencijal za rešavanje problema sa velikim brojem promenljivih. Danas se upotrebljava u velikom broju naučnih oblasti (obrada signala, mehanika, kontrola kvaliteta, meteorologija, hemometrija, genetika i bioinformatika, linearna algebra...) u kojima je poznata pod različitim nazivima.

Analiza glavnih komponenti je jedna od najjednostavnijih multivarijacionih tehnika. Primenjuje se kada je veliki broj varijabli u datom skupu redundantan, odnosno kada se više varijabli odnosi na istu dimenziju, te ne pružaju dodatne informacije koje nisu već sadržane u nekoj drugoj varijabli. Zadatak PCA je određivanje linearnih kombinacija originalnih promenljivih, koje se nazivaju glavne komponente, koje imaju maksimalnu varijansu i međusobno su nekorelisane, pritom gubeći u najmanjoj mogućoj meri informacije sadržane u originalnom skupu podataka. Ovim postupkom se originalni, korelisani skup podataka transformiše u skup nekorelisanih promenljivih sa opadajućim vrednostima varijansi. Želja je da varijanse većeg broja glavnih komponenti budu zanemarljivo male, te da se kovarijaciona struktura skupa originalnih podataka može uspešno aproksimirati sa svega nekoliko glavnih komponenti.

Ovim se postižu dva cilja: redukcija dimenzija originalnog skupa podataka i pojednostavljenje njegove interpretacije. Dakle, osim što predstavlja izvestan stepen uštede, pošto se broj promenljivih koje se koriste u daljoj analizi značajno smanjuje, metoda glavnih komponenti generiše hipoteze o proučavanom fenomenu i pomaže u razumevanju njegove strukture.

Analiza glavnih komponentata je izuzetno pogodna metoda za primenu u biomedicini, naročito kada je u pitanju analiza pokreta. Na primer, Coffey i dr. [12] su primenili funkcionalnu analizu glavnih komponenti (eng. *Functional Principal Component Analysis*), proširenje standardne multivarijantne PCA u domenu funkcionalne analize podataka (eng. *Functional Data Analysis*), na setu biomehaničkih podataka u cilju ispitivanja povreda Ahilove tetive i funkcionalnih efekata ortoza. Maurer i dr. [13] su koristili PCA u analizi kinematike hoda, testirajući različite hipoteze zasnovane na prvoj i višim glavnim komponentama.

2.3.2 Definicija i osobine²

Neka je \mathbf{X} p – dimenzionalni slučajni vektor sa kovarijacionom matricom $\mathbf{\Sigma}$, a $Y_1 = \alpha_{11}X_1 + \alpha_{12}X_2 + \dots + \alpha_{1p}X_p = \mathbf{\alpha}'_1\mathbf{X}$ linearna kombinacija elemenata vektora \mathbf{X} , gde su $\alpha_{11}, \alpha_{12}, \dots, \alpha_{1p}$ koeficijenti linearne kombinacije. Po pravilima linearne algebre, poznato je da je $Var(Y_1) = Var(\mathbf{\alpha}'_1\mathbf{X}) = \mathbf{\alpha}'_1\mathbf{\Sigma}\mathbf{\alpha}_1$. Potrebno je odrediti vektor koeficijenata $\mathbf{\alpha}_1$ tako da se maksimizira varijansa promenljive Y_1 , uz ograničenje da je traženi vektor jedinične dužine, tj. da je $\mathbf{\alpha}'_1\mathbf{\alpha}_1 = 1$. Do rešenja se dolazi maksimizacijom *Lagrange*-ove funkcije

$$\mathbf{\alpha}'_1\mathbf{\Sigma}\mathbf{\alpha}_1 - \lambda(\mathbf{\alpha}'_1\mathbf{\alpha}_1 - 1),$$

gde je λ *Lagrange*-ov množitelj. Diferenciranjem *Lagrange*-ove funkcije po koeficijentima $\mathbf{\alpha}_1$ i izjednačavanjem dobijenog izraza sa nulom dobija se jednakost

$$\mathbf{\Sigma}\mathbf{\alpha}_1 - \lambda\mathbf{\alpha}_1 = 0,$$

odnosno

$$(\mathbf{\Sigma} - \lambda\mathbf{I})\mathbf{\alpha}_1 = 0,$$

gde je \mathbf{I} jedinična matrica dimenzija $p \times p$. Uslov za dobijanje netrivialnog rešenja za $\mathbf{\alpha}_1$ je

$$|\mathbf{\Sigma} - \lambda\mathbf{I}| = 0,$$

što znači da λ mora biti jedan od karakterističnih korena kovarijacione matrice $\mathbf{\Sigma}$. Pošto je cilj maksimizacija varijanse, kao logičan izbor za λ se nameće najveći karakteristični koren. Označimo ga kao λ_1 . Iz uslova $(\mathbf{\Sigma} - \lambda\mathbf{I})\mathbf{\alpha}_1 = 0$ sledi da je $\mathbf{\alpha}_1$ karakteristični vektor pridružen karakterističnom korenu λ_1 .

U slučaju da je potrebno odrediti više od jedne glavne komponente, za drugu glavnu komponentu se sprovodi analogan postupak, uz dodatni uslov da kovarijansa prve i druge glavne komponente bude jednaka nuli. Neka je $Y_2 = \alpha_{21}X_1 + \alpha_{22}X_2 + \dots + \alpha_{2p}X_p = \mathbf{\alpha}'_2\mathbf{X}$ linearna kombinacija elemenata vektora \mathbf{X} čije koeficijente $\alpha_{21}, \alpha_{22}, \dots, \alpha_{2p}$ treba odrediti uz dva uslova: normirajući uslov $\mathbf{\alpha}'_2\mathbf{\alpha}_2 = 1$ i dodatni uslov nekorelisanosti glavnih komponenti, do kojeg se dolazi polazeći od izraza za kovarijansu:

$$Cov(Y_1, Y_2) = Cov(\mathbf{\alpha}'_1\mathbf{X}, \mathbf{\alpha}'_2\mathbf{X}) = \mathbf{\alpha}'_2\mathbf{\Sigma}\mathbf{\alpha}_1 = \mathbf{\alpha}'_1\mathbf{\Sigma}\mathbf{\alpha}_2 = \mathbf{\alpha}'_2\mathbf{\alpha}_1\lambda_1 = \mathbf{\alpha}'_1\mathbf{\alpha}_2\lambda_1.$$

² Preuzeto i modifikovano iz [4]

Pošto je $\Sigma \alpha_1 = \lambda \alpha_1$, kovarijansa prve i druge glavne komponente može biti jednaka nuli samo u slučaju da je $\alpha_2' \alpha_1 = 0$. Diferenciranjem *Lagrange*-ove funkcije sa dva množitelja

$$\alpha_2' \Sigma \alpha_2 - \lambda (\alpha_2' \alpha_2 - 1) - \phi \alpha_2' \alpha_1$$

po α_2 , izjednačavanjem dobijenog izraza sa nulom i množenjem dobijenog izraza sa leve strane sa α_1' dobija se:

$$\alpha_1' \Sigma \alpha_2 - \lambda \alpha_1' \alpha_2 - \phi \alpha_1' \alpha_1 = 0.$$

Prva dva člana ove jednakosti su jednaka nuli, pa iz uslova $\alpha_1' \alpha_1 = 1$ sledi da je $\phi = 0$. Na osnovu toga je $\Sigma \alpha_2 - \lambda \alpha_2 = 0$, odnosno $(\Sigma - \lambda \mathbf{I}) \alpha_2 = 0$, što podrazumeva da je λ karakteristični koren kovarijacione matrice Σ , a α_2 odgovarajući karakteristični vektor. Kao i kod prve glavne komponente, za λ biramo što je moguće veću vrednost – drugi po veličini karakteristični koren λ_2 , α_2 je njemu pridruženi karakteristični vektor, a linearna kombinacija $Y_2 = \alpha_2' \mathbf{X}$ je druga glavna komponenta.

Do ostalih glavnih komponenti, kojih ima onoliko koliko ima različitih karakterističnih korena kovarijacione matrice, dolazi se na isti način. Kada su svi karakteristični koreni kovarijacione matrice Σ različiti i u opadajućem poretku $\lambda_1 > \lambda_2 > \dots > \lambda_p \geq 0$, tada postoji p glavnih komponenti Y_1, Y_2, \dots, Y_p , definisanih kao $Y_j = \alpha_j' \mathbf{X}, j = 1, 2, \dots, p$, gde su vektori koeficijentata α_j karakteristični vektori kovarijacione matrice Σ koji odgovaraju karakterističnim korenima λ_j .

Sledeće osobine glavnih komponenti direktno slede iz definicije:

- Matematičko očekivanje glavne komponente je jednako nuli $E(Y_j) = 0, j = 1, \dots, p$;
- Varijansa glavne komponente je jednaka odgovarajućem karakterističnom korenu $Var(Y_j) = \lambda_j, j = 1, \dots, p$, iz čega proizilazi $Var(Y_1) \geq Var(Y_2) \geq \dots \geq Var(Y_p) \geq 0$;
- Kovarijansa svakog para glavnih komponenti je jednaka nuli $Cov(Y_i, Y_j) = 0, i \neq j$.

Glavne komponente je moguće izvesti i na drugi način, koristeći matričnu algebru. Neka je ponovo \mathbf{X} p – dimenzionalni slučajni vektor promenljivih sa kovarijacionom matricom Σ , a \mathbf{Y} vektor glavnih komponenti takav da je $\mathbf{Y}' = [Y_1 \ Y_2 \ \dots \ Y_p]$. Transformacija originalnog skupa promenljivih, $\mathbf{Y} = \mathbf{A} \mathbf{X}$, se naziva ortogonalna transformacija ili rotacija, a sama matrica \mathbf{A} , matrica dimenzija $p \times p$ čiji su redovi karakteristični vektori kovarijacione matrice pridruženi odgovarajućim karakterističnim korenima ($\mathbf{A}' = [\alpha_1' \ \alpha_2' \ \dots \ \alpha_p']$), je ortogonalna matrica. Ovo je posledica osobine matice \mathbf{A} da je $\mathbf{A}' = \mathbf{A}^{-1}$, što proizilazi iz osobina karakterističnih vektora.



Transformacijom se vrši rotacija koordinatnih osa za određeni ugao, pri čemu ose ostaju međusobno upravne, a ugao između bilo koja dva vektora ostaje nepromenjen nakon transformacije, odakle potiče njen naziv.

Primenom ortogonalne transformacije može se izvršiti ortogonalna dekompozicija kvadratne simetrične matrice čiji su koreni različiti, a upravo takva je kovarijaciona matrica Σ . Tada je $\Sigma = \mathbf{A}'\Lambda\mathbf{A}$, gde je Λ dijagonalna matrica čiji su elementi karakteristični koreni matrice Σ :

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_p \end{bmatrix},$$

a matrica \mathbf{A} je ranije definisana ortogonalna matrica. Kovarijaciona matrica vektora glavnih komponenti je $Var(\mathbf{Y}) = \mathbf{A}\Sigma\mathbf{A}' = \mathbf{A}(\mathbf{A}'\Lambda\mathbf{A})\mathbf{A}' = \Lambda$, pošto je \mathbf{A} ortogonalna matrica i za nju važi $\mathbf{A}'\mathbf{A} = \mathbf{I}$. Ovim su na drugi način izvedene glavne komponente.

Na osnovu dobijenog rezultata se može odrediti generalizovana varijansa vektora \mathbf{Y} , i to na dva načina. Prema prvoj definiciji generalizovana varijansa jednaka je determinanti kovarijacione matrice, a prema drugoj tragu kovarijacione matrice. Na osnovu obe definicije se dolazi do istog zaključka: generalizovane varijanse originalnog i transformisanog skupa podataka su međusobno jednake (za detaljnije izvođenje pogledati [4]), što znači da glavne komponente obuhvataju sve varijacije originalnog skupa podataka.

Na osnovu dobijenih rezultata se intuitivno nameće da se relativni doprinos j -te glavne komponente može odrediti kao odnos odgovarajućeg karakterističnog korena λ_j i generalizovane varijanse (po drugoj definiciji):

$$\frac{\lambda_j}{\sum_{k=1}^p \lambda_k}, j = 1, 2, \dots, p.$$

2.3.3 Primena PCA

Dobijene glavne komponente najčešće predstavljaju tek prvi korak i koriste se kao polazni podatak u daljoj analizi. U našem slučaju, dalje sledi klasifikacija pokreta, a sve u cilju poboljšanja kvaliteta upravljanja protezom za šaku primenom površinskih elektromiografskih signala. Uzimajući u obzir dimenzije originalnog skupa podataka (27 ispitanika, 52 pokreta, 10 ponavljanja i 10 elektroda), neophodno je smanjiti taj broj, uz minimalne gubitke sadržanih informacija, da bi klasifikacija uopšte bila moguća i validna.

Analiza glavnih komponenti se, u opštem slučaju, vrši u sledećim koracima:

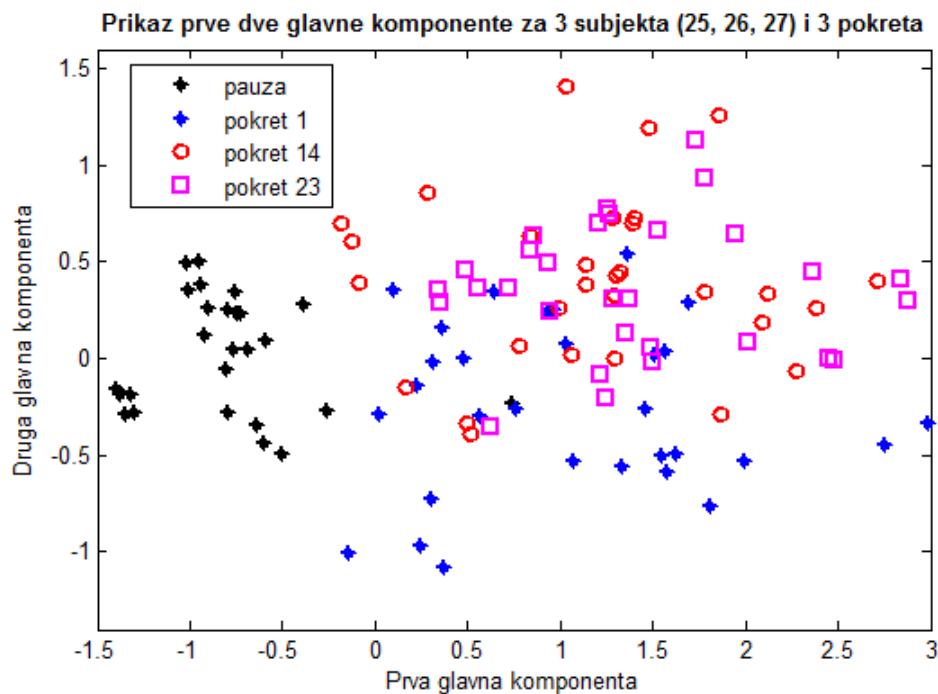
1. Standardizuju se originalni podaci tako da originalne promenljive imaju nultu aritmetičku sredinu i jediničnu varijansu. Ovaj korak je već opisan u odeljku o pretprocesiranju signala.
2. Izračunava se kovarijaciona matrica Σ .
3. Izračunavaju se sopstvene vrednosti $\lambda_1, \lambda_2, \dots, \lambda_p$ i odgovarajući sopstveni vektori $\alpha_1, \alpha_2, \dots, \alpha_p$, koji definišu glavne komponente.
4. Eliminišu se komponente koje obuhvataju zanemarljiv deo varijacija originalnog skupa podataka i bira se broj glavnih komponenti.

Izbor broja glavnih komponenti je problem za koji postoji više rešenja, a od konkretne situacije zavisi koje je najbolje primeniti. Osim formalnog postupka testiranja značajnosti karakterističnih korena, zasnovanog na korišćenju količnika verodostojnosti (eng. *Likelihood Ratio test*), postoje i manje formalni postupci: apriorno fiksiranje kumulativne proporcije ukupne varijanse „objašnjene“ izdvojenim glavnim komponentama (obično je dovoljno 80 – 90 % ukupne varijanse), izdvajanje onih glavnih komponenti čije su varijanse veće od prosečne varijanse ili geometrijske sredine svih karakterističnih korenova, kao i „*scree test*“ – grafički prikaz vrednosti karakterističnih korenova po njihovom rednom broju, određivanje preloma na krivoj i broja glavnih komponenti kao rednog broja glavne komponente čija vrednost karakterističnog korena leži poslednja na pravoj liniji nakon preloma [14].

U konkretnoj prikazanoj analizi, u obzir su uzete samo prve dve glavne komponente. Osim što je olakšana dalja klasifikacija, omogućen je i dvodimenzionalan grafički prikaz podataka, koji daje bolji vizuelni uvid u odnos između pokreta i sEMG signala.

Kao i za pretprocesiranje, za analizu glavnih komponenti je korišćen program MATLAB (The Mathworks, Natick, USA). Algoritam je realizovan tako da je na početku potrebno uneti ukupan željeni broj subjekata (n_s) i pokreta (n_p), a zatim iz ponuđenog menija odabrati konkretne subjekte i pokrete. Ukoliko su odabrani svi ispitanici ili svi pokreti meni se neće pojaviti, već će automatski biti učitani svi potrebni signali. Izgled projektovanog komandnog prozora i menija za

izbor pojedinačnih željenih ispitanika i pokreta je dat u Prilogu A (Sl. A.1, Sl. A.2 i Sl. A.3, respektivno). Nakon toga sledi učitavanje potrebnih podataka (.mat fajlovi dobijeni nakon pretprocesiranja) i formiranje dve trodimenzionalne matrice, jedne za pokrete i jedne za pauzu. Obe 3D matrice su istih dimenzija - $10n_s \times 10 \times n_p$, što odgovara 10 ponavljanja za svakog od n_s subjekata, 10 elektroda i n_p različitih pokreta. Teorijski opisana analiza glavnih komponenti (koraci 2, 3 i 4) se sprovodi nad svakom od n_p dvodimenzionalnih matrica, kako za pokrete tako i za pauze (odmor između ponavljanja pokreta). Finalni rezultat je dvodimenzionalni prikaz prve dve glavne komponente, crnom bojom za pauzu i različitim bojama i oznakama za različite pokrete, kao što je prikazano na Sl. 2.7 na primeru tri subjekta (subjekti br. 25, 26 i 27) i tri pokreta (pokreti br. 1, 14 i 23).



Sl. 2.7 Prikaz prve dve glavne komponente sEMG signala za 3 ispitanika (ispitanici br. 25, 26 i 27) i 3 pokreta (pokreti br. 1, 14 i 23). Podaci su preuzeti iz NinaPro baze podataka, a prikaz realizovan u projektovanoj GUI aplikaciji.

Analiza glavnih komponenti elektromiografskih signala za klasifikaciju pokreta šake



Iako se za grafički prikaz i klasifikaciju koriste samo prve dve komponente, da bi se stekao celokupan uvid u varijabilnost transformisanog seta podataka, potrebno je odrediti procentne udele svih glavnih komponenti. Kao što je ranije definisano, relativni doprinos j -te glavne komponente predstavlja odnos odgovarajućeg karakterističnog korena λ_j i sume svih karakterističnih korenova (generalizovane varijanse, po drugoj definiciji). Procentni udeo se dobija množenjem relativnog udela sa 100. Program je projektovan tako da izračunava i grafički prikazuje procentne udele glavnih komponenti kada su odabrane određene vrednosti ukupnog broja pokreta i ispitanika, odnosno u karakterističnim slučajevima koji ilustruju razlike koje postoje kod različitih ispitanika i pokreta. Ove informacije su izuzetno korisne za dalju analizu (izbor pokreta koji se mogu zajedno analizirati sa istim brojem razmatranih glavnih komponenti, izbor broja glavnih komponenti koje se uzimaju u obzir prilikom klasifikacije, izbor broja elektroda koje se koriste za snimanje sEMG signala).

2.4 Klasifikacija pokreta

Za klasifikaciju pokreta je korišćen kvadratni klasifikator. U opštem slučaju, kada je odbirke vektora opservacija X potrebno pridružiti dvema klasama ω_1 i ω_2 , diskriminaciona funkcija i pravilo odlučivanja kvadratnog klasifikatora su definisani jednakostima:

$$h(X) = X^T Q X + V^T X + v_0 > 0 \rightarrow X \in \omega_1$$

$$h(X) = X^T Q X + V^T X + v_0 < 0 \rightarrow X \in \omega_2,$$

gde su Q, V i v_0 matrica, vektor i skalar, respektivno. Iako uspešan za probleme sa dve klase, kvadratni klasifikator se ne može primeniti za probleme sa više klasa, već se koristi njegovo proširenje – “deo-po-deo” (eng. *piecewise*) kvadratni klasifikator. Najbolje rezultate, u smislu minimizacije greške klasifikacije, daje *Bayes-ov* test više hipoteza [15]. Pravilo odlučivanja, u slučaju L klasa, glasi:

$$q_k(X) = \max_i q_i(X) \rightarrow X \in \omega_k, k = 1, \dots, L,$$

gde je q_i aposteriorna verovatnoća klase ω_i ako je poznato X . Odnosno, po *Bayes-ovoj* teoremi:

$$P_k p_k(X) = \max_i P_i p_i(X) \rightarrow X \in \omega_k, k = 1, \dots, L,$$

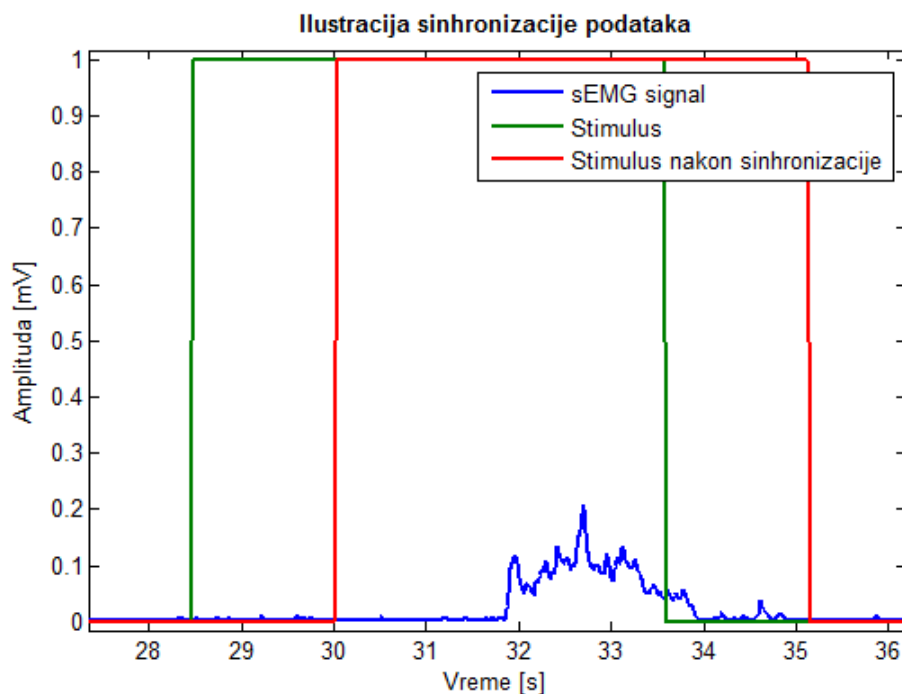
gde je P_i apriorna verovatnoća klase ω_i , a p_i uslovna funkcija gustine verovatnoće.

Za klasifikaciju je korišćena funkcija *classify* iz MATLAB Statistics Toolbox-a. Klasifikacije je sprovedena za tri pokreta i pauzu, sa ciljem da se uporede rezultati dobijeni primenom kvadratnog klasifikatora i LS-SVM (eng. *Least-Squares Support Vector Machine*) korišćenog u [2]. Takođe, ovaj set pokreta omogućava pregledan grafički prikaz. Izračunata je ukupna greška klasifikacije, kao i greške klasifikacije za pauzu i pojedinačne pokrete.

3 REZULTATI

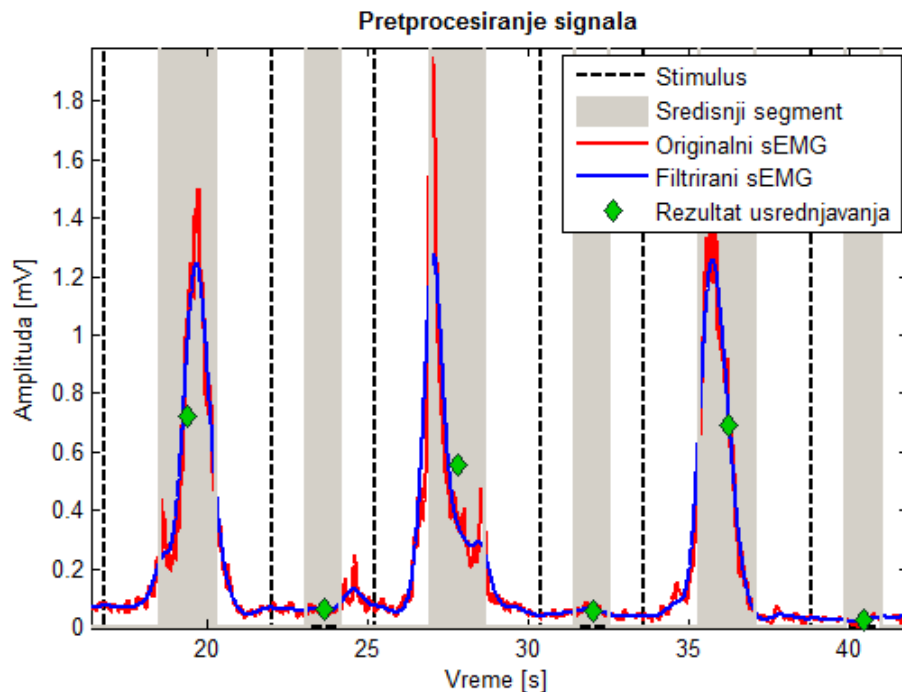
3.1 Rezultati pretprocesiranja

Sl. 3.1 ilustruje neophodnost vremenske sinhronizacije podataka. Izdvojen je sEMG signal koji odgovara jednom izvođenju prvog pokreta (vrednost stimulusa je jedan tokom trajanja pokreta, a nula dok traje pauza) kod prvog ispitanika. Kao što je rečeno, podaci nisu sinhronizovani, već svaki odbirak sadrži informaciju o tačnom vremenu kada je snimljen. Dalje procesiranje je moguće tek nakon što su signali vremenski usklađeni, a ovo je naročito bitno za uspešnu segmentaciju sEMG signala.



Sl. 3.1 Efekat vremenske sinhronizacije podataka o mišićnoj aktivnosti i stimulusu. Prikaz podataka iz NinaPro baze realizovan u projektovanoj GUI aplikaciji.

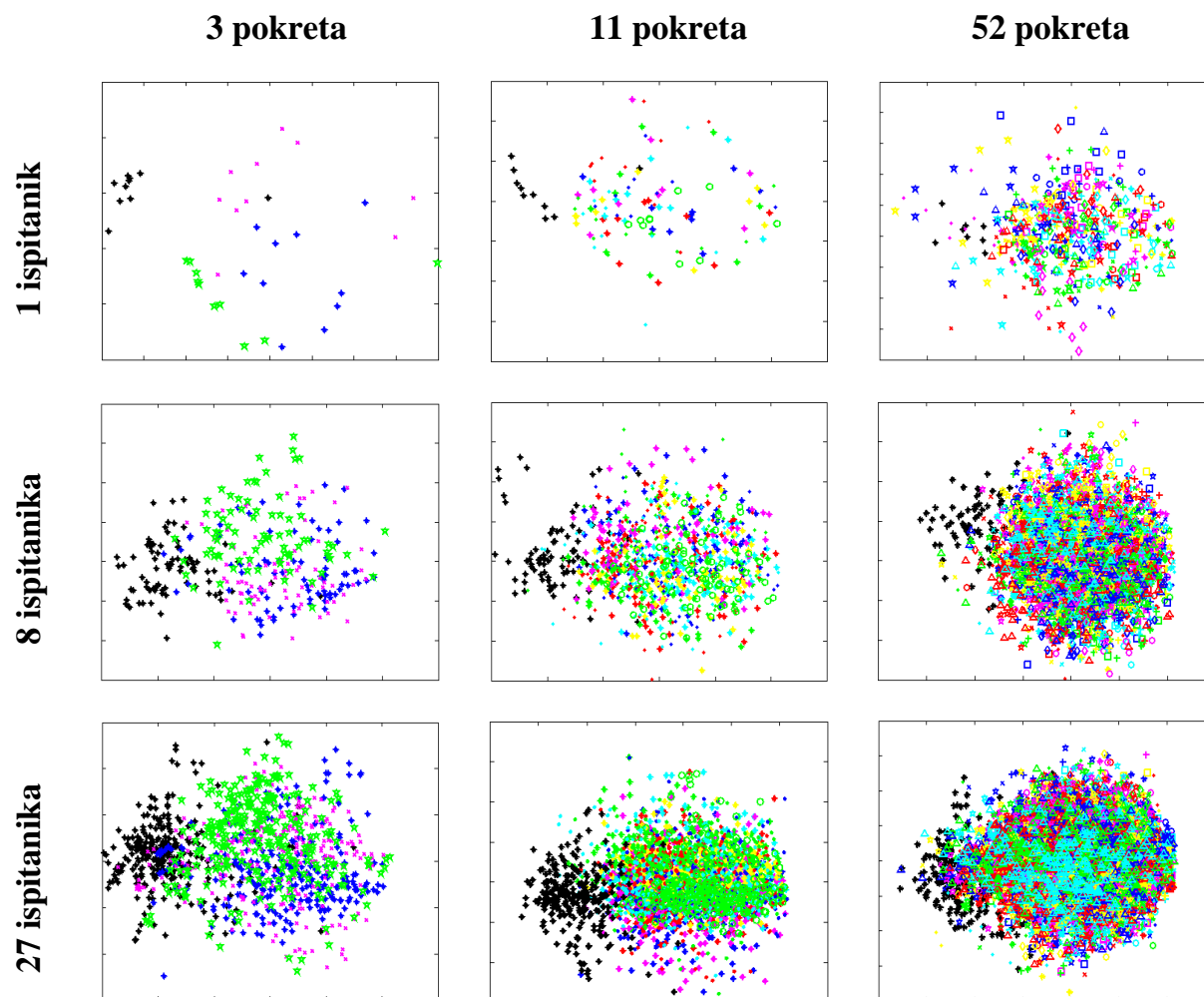
Na Sl. 3.2 je prikazan rezultat pretprocesiranja sEMG signala na primeru prvog subjekta i četvrtog pokreta. Izdvojena su tri ponavljanja pokreta sa samo jedne od deset elektroda radi bolje preglednosti.



Sl. 3.2 Rezultat pretprocesiranja signala za 1. subjekta, 4. pokret i 8. elektrodu. Prikaz podataka iz NinaPro baze realizovan u projektovanoj GUI aplikaciji.

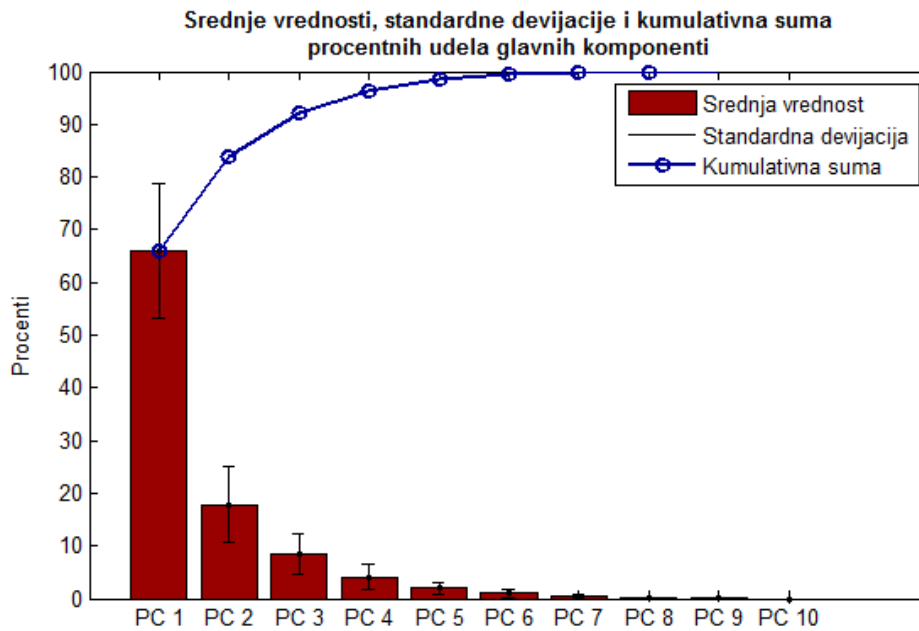
3.2 Rezultati PCA

Rezultat PCA je grafički prikaz prve i druge glavne komponente. Na Sl. 3.3 je razmatrano devet slučajeva, za tri seta pokreta i tri seta ispitanika. Setovi pokreta su: tri hvata (pokreti 39, 40 i 43) koji su ispitivani u [16], 11 pokreta prstiju (pokreti 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 i 15) koji su ispitivani u [17] i svih 52 pokreta. Setovi ispitanika su: jedan ispitanik (ispitanik br. 4), 8 ispitanika (ispitanici br. 10, 11, 12, 13, 14, 15, 16, 17) i svih 27 ispitanika. Radi bolje preglednosti sa slika su izostavljeni nazivi i numeracija osa, legenda i naslov.

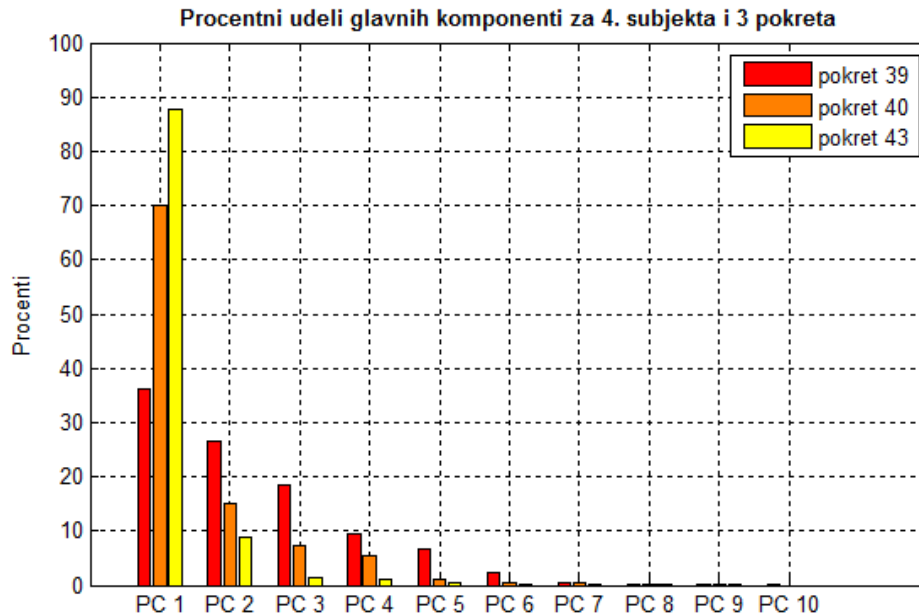


Sl. 3.3 Prve dve glavne komponente sEMG signala za različite kombinacije broja ispitanika i pokreta. Signali su preuzeti iz NinaPro baze podataka, a obrada i prikaz realizovani u projektovanoj GUI aplikaciji.

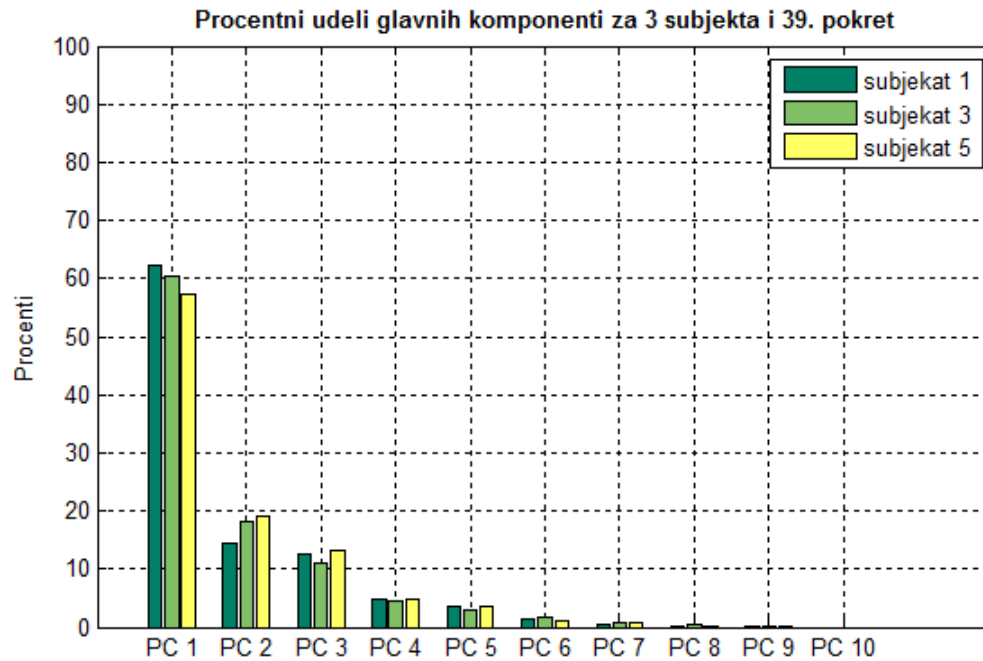
Na Sl. 3.4 su prikazane srednje vrednosti i standardne devijacije, kao i kumulativna suma procentnih doprinosa glavnih komponenti za prvog ispitanika i svih 52 pokreta. Na Sl. 3.5 je dat uporedni prikaz procentnih udela glavnih komponenti za jednog ispitanika (ispitanik br. 4) i tri pokreta (pokreti br. 39, 40 i 43), dok je na Sl. 3.6 dat prikaz za tri ispitanika (ispitanici br. 1, 3 i 5) i jedan pokret (pokret br. 39).



Sl. 3.4 Srednje vrednosti, standardne devijacije i kumulativna suma procentnih udela glavnih komponenti za prvog ispitanika i 52 pokreta. Signali su preuzeti iz NinaPro baze podataka, a obrada i prikaz realizovani u projektovanoj GUI aplikaciji.



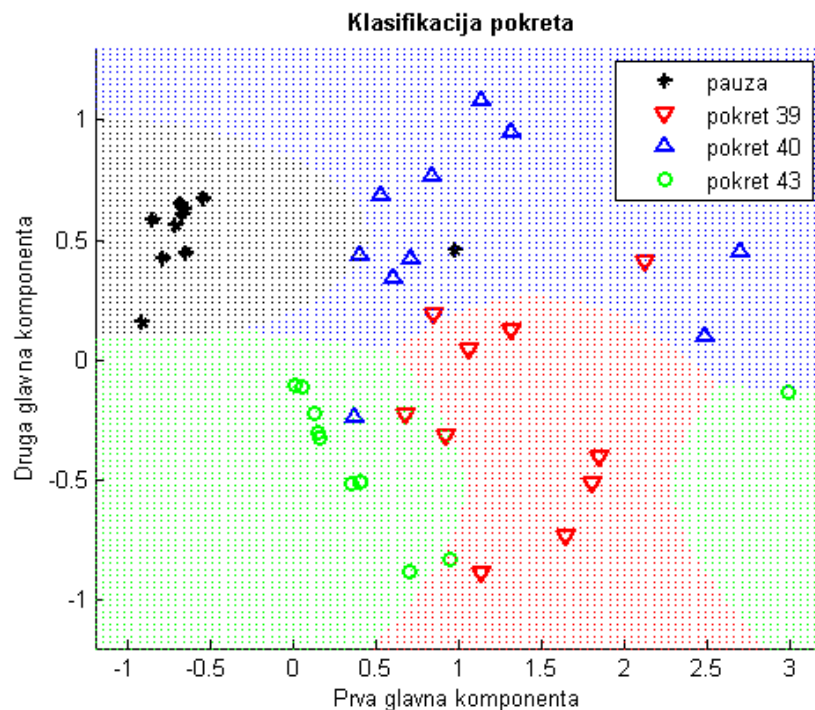
Sl. 3.5 Procentni udeli glavnih komponenti za četvrtog ispitanika i pokrete br. 39, 40 i 43. Signali su preuzeti iz NinaPro baze podataka, a obrada i prikaz realizovani u projektovanoj GUI aplikaciji.



Sl. 3.6 Procentni udeli glavnih komponenti za ispitanike br. 1, 3 i 5 i 39. pokret. Signali su preuzeti iz NinaPro baze podataka, a obrada i prikaz realizovani u projektovanoj GUI aplikaciji.

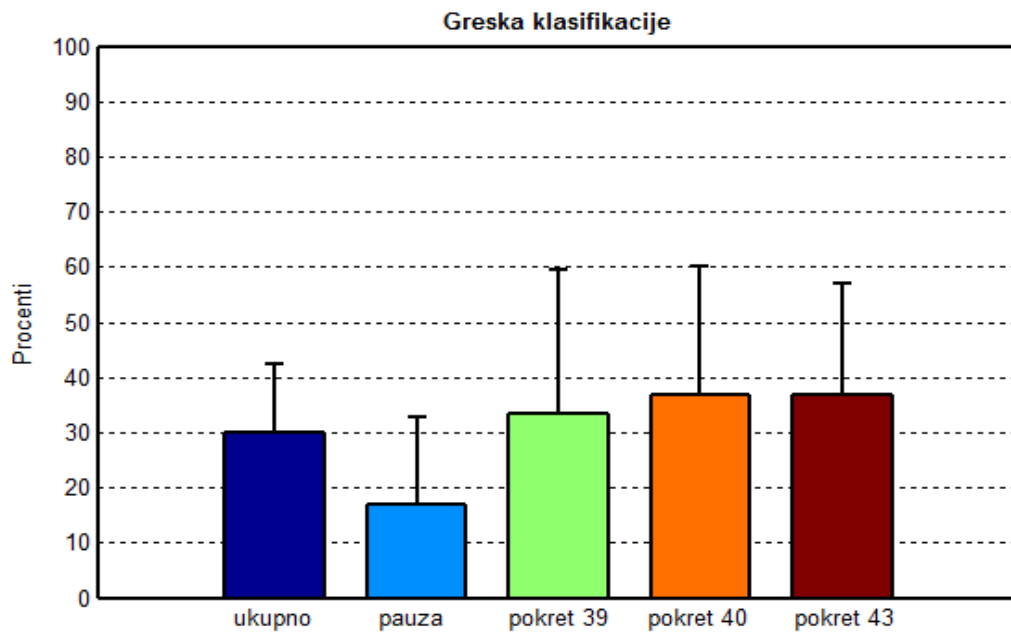
3.3 Rezultati klasifikacije pokreta

Rezultat klasifikacije pokreta je grafički prikazan na Sl. 3.7 u slučaju jednog ispitanika (ispitanik br. 4) i tri pokreta – br. 39, 40 i 43. U pitanju su isti ispitanik i tri hvata za koje je prikazan rezultat PCA i procentni udeli glavnih komponenti. Greška klasifikacije iznosi 20 %.



Sl. 3.7 Rezultat projektovanog klasifikatora za tri hvata (pokreti br. 39, 40 i 43) i pauzu u slučaju ispitanika br. 4. Greška klasifikacije iznosi 20 %. Signali su preuzeti iz NinaPro baze podataka, a obrada i prikaz realizovani u projektovanoj GUI aplikaciji.

Za svakog od 27 ispitanika je izvršena klasifikacija za ista tri pokreta (pokreti br. 39, 40 i 43). Izračunata je ukupna greška klasifikacije, kao i greške klasifikacije za pojedinačne pokrete i pauzu. Ukupna greške klasifikacije i greške klasifikacije za pojedine pokrete i pauzu, (u formi srednja vrednost \pm standardna devijacija za svih 27 ispitanika), su prikazane na Sl. 3.8.

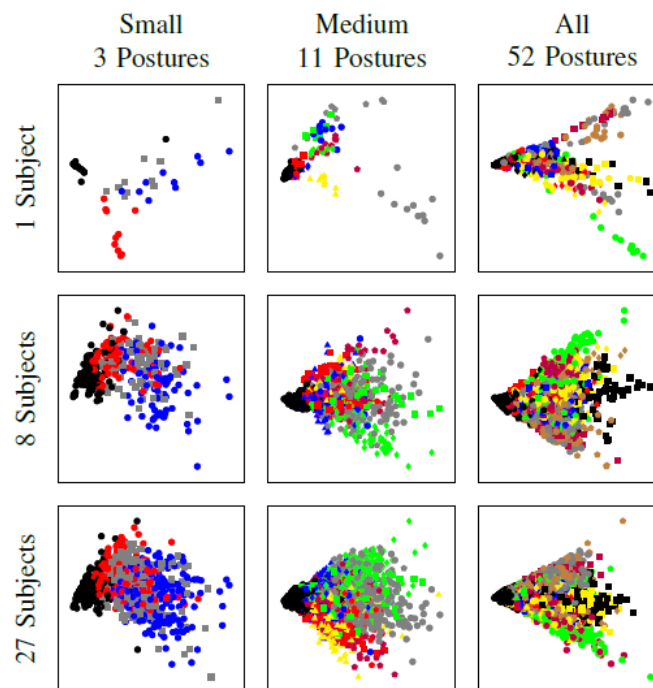


Sl. 3.8 Ukupna greška klasifikacije i greške klasifikacije za pauzu i 3 razmatrana hvata (pokreti br. 39, 40 i 43). Rezultati prikazuju srednju vrednost i standardnu devijaciju za svih 27 ispitanika. Signali su preuzeti iz NinaPro baze podataka, a obrada i prikaz realizovani u projektovanoj GUI aplikaciji.

4 DISKUSIJA I ZAKLJUČCI

4.1 Diskusija dobijenih rezultata i opšti zaključci

Vizuelnim poređenjem Sl. 3.3, koja prikazuje prve dve glavne komponente sEMG signala za različite setove ispitanika i pokreta, i Sl. 4.1, koja predstavlja isti rezultat dobijen u [2], uočavaju se očigledne razlike. S obzirom da u [2] nisu navedeni redni brojevi ispitanika za koje su prikazani rezultati, ove razlike su očekivane. Razlike takođe mogu poticati od mogućeg različitog načina normalizacije signala. Opis normalizacije naveden u [2], “*podaci za pojedinačne subjekte su normalizovani tako da svaki sEMG signal ima nultu srednju vrednost i jediničnu standardnu devijaciju*”, ne precizira da li su signali normalizovani za pojedinačne pokrete za svakog subjekta, kako je urađeno u ovom diplomskom radu, ili za sve pokrete koji pripadaju jednom ispitaniku zajedno. Ove razlike ni u kom slučaju ne znače da dobijeni rezultati nisu tačni. Glavni cilj PCA, smanjenje dimenzije početnog seta podataka i olakšavanje njegovog tumačenja, je uspešno ostvaren.



Sl. 4.1 Rezultati PCA dobijeni u [2].

Analiza glavnih komponenti elektromiografskih signala za klasifikaciju pokreta šake



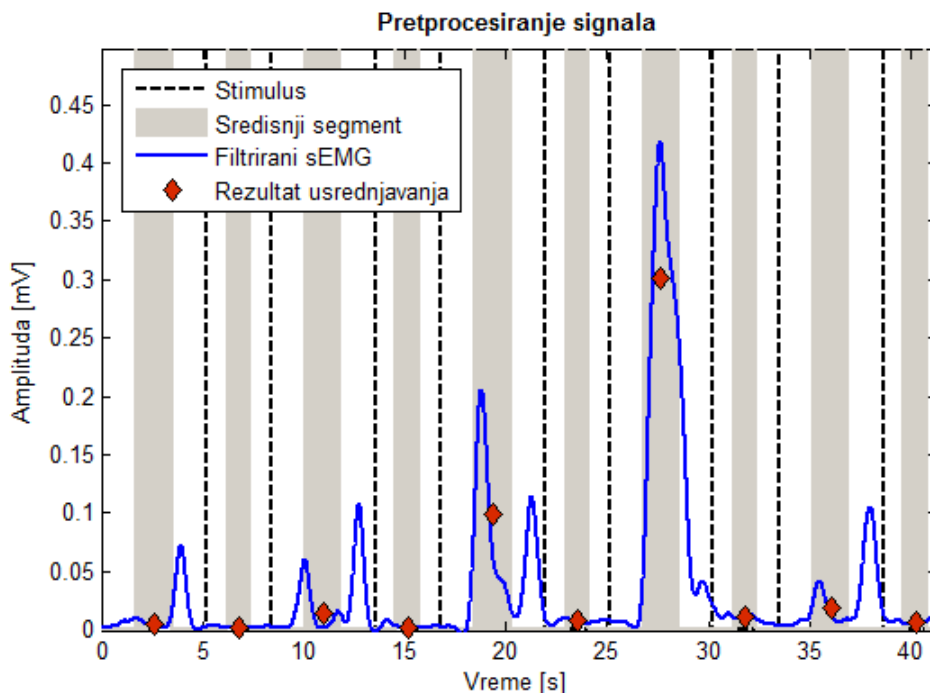
Greška klasifikacije, dobijena primenom kvadratnog klasifikatora na setu od tri pokreta i pauze, iznosi $30,1 \pm 12,3$ % (Sl. 3.8) U poređenju sa rezultatima dobijenim u [2] na istom setu podataka, koji iznose $7,5 \pm 7,2$ %, može se videti lošiji učinak kvadratnog klasifikatora u odnosu na LS-SVM. Ovaj rezultat je takođe očekivan, s obzirom da je kvadratni klasifikator jedan od najjednostavnijih, dok LS-SVM spada u red složenijih metoda mašinskog učenja. Relativno velika standardna devijacija greške klasifikacije ukazuje na značajne razlike u uspešnosti klasifikacije za različite subjekte. Ovo je problem u slučaju oba primenjena klasifikatora, te je potrebno projektovati napredniji klasifikacioni algoritam koji će moći da prevaziđe problem velike varijabilnosti u izvođenju istih pokreta među različitim subjektima.

Procentni udeli pojedinačnih glavnih komponenti u varijabilnosti transformisanog seta podataka predstavljaju bitnu informaciju koju treba imati u vidu prilikom tumačenja rezultata PCA i klasifikacije. Sa Sl. 3.4 se može videti da u slučaju ispitanika br. 1 prve dve glavne komponente obuhvataju 83% sadržanih informacija. Ovo se može smatrati prosečnom vrednošću, jer je kod svih ispitanika ta vrednost u opsegu 80 – 90%, kada se posmatra srednja vrednost za sve pokrete. Ipak, procentni udeo prve komponente u slučaju prvog ispitanika je 66 ± 12 %, a ova vrednost standardne devijacije govori o razlikama koje postoje u procentnim udelima glavnih komponenti u slučaju različitih pokreta. Sl. 3.5 potvrđuje tu tvrdnju: u slučaju ispitanika br. 4 procentni udeli prve glavne komponente su 36%, 70% i 88%, druge glavne komponente – 27%, 15% i 9%, i treće komponente – 18%, 9% i 2%, za pokrete br. 39, 40 i 43, respektivno. S obzirom da je procentni udeo prve glavne komponente za pokret br. 43 (88%), veći od zbira udela prve dve komponente za pokret br. 40 (85%) i zbira prve tri komponente za pokret br. 39 (80%), nameće se zaključak da procentni udeli imaju značajnog uticaja na rezultat klasifikacije koja sledi. Procentni udeo treće glavne komponente u slučaju pojedinih pokreta (kao što je pokret br. 39) ne može biti zanemaren, pošto sadrži više informacija od druge glavne komponente kod većine ostalih pokreta (kao što su pokreti br. 40 i 43). Na osnovu ove analize se može pretpostaviti da bi rezultati klasifikacije nesumnjivo bili bolji kada bi se razmatrale prve tri, umesto prve dve glavne komponente. Upravo je poređenje rezultata klasifikacije za dve i tri glavne komponente tema rada koji će se nastaviti na ovaj diplomski rad. Sa druge strane, rezultat sa Sl. 3.6 pokazuje da je razlika u procentnim udelima glavnih komponenti za jedan pokret i različite ispitanike jako mala. Ovaj rezultat ukazuje da je opravdano raditi klasifikaciju za iste pokrete i različite subjekte.

Podatke o procentnim udelima glavnih komponenti treba uzeti u obzir i prilikom razmatranja broja kanala za snimanje sEMG signala. Iako korišćenih 10 elektroda pokrivaju veliki deo, ako ne i svu mišićnu aktivnost podlaktice, sa činjenicom da je za sve subjekte i sve pokrete zbir procentnih udela poslednjih pet glavnih komponenti manji od 5%, nameće se pitanje o smanjenju broja analiziranih kanala. Efekat smanjenja broja kanala bi bio zanimljiv za analizu, pri čemu bi raspored manjeg broja elektroda trebalo pažljivo odabrati, uz obaveznu prethodnu analizu pokreta.

4.2 Ograničenja primenjenih metoda

Preprocesiranje i priprema signala za PCA imaju velikog uticaja na krajnji rezultat. Jedan od problema prilikom preprocesiranja je izdvajanje segmenata sEMG signala koji odgovaraju izvođenjima pokreta. Predloženi metod segmentacije signala na tri dela i izdvajanja središnjeg segmenta, čiji se odbirci nakon toga usrednjavaju, je jednostavan i uglavnom uspešan (Sl. 3.2). Međutim, pojedina ponavljanja signala nemaju karakteristični oblik, imaju različite amplitude ili su pomerena u odnosu na sredinu segmenta. Na to mogu da utiču vrsta samog pokreta, pomeranje elektrode prilikom snimanja sEMG signala, zamor ili trenutni nedostatak koncentracije kod ispitanika. Na Sl. 4.2 je prikazano pet ponavljanja signala sa takvim osobinama i efekat preprocesiranja. Kod tri od ovih pet ponavljanja nakon preprocesiranja nema razlike između odbiraka koji pripadaju pokretu i pauzi. Zbog toga bi za izdvajanje pokreta trebalo upotrebiti neki od algoritama za detekciju mišićnog oneta predloženih u [18], a zatim izvršiti usrednjavanje odbiraka izdvojenog segmenta.



Sl. 4.2 Efekat preprocesiranja signala. Izdvojeno je pet ponavljanja za 1. subjekta, 1. pokret i 10. elektrodu. Signali su preuzeti iz NinaPro baze podataka, a obrada i prikaz realizovani u projektovanoj GUI aplikaciji.

Nedostaci NINAPRO baze podataka su odsustvo adekvatne analize pokreta i ponovljivosti merenja. Ponovljivost merenja bi doprinela verodostojnosti rezultata klasifikacije i eliminisala negativne efekte zamora kod ispitanika, pomeranja elektroda i ostalih nepovoljnih situacija koje se mogu dogoditi tokom merenja. Takođe, sile koje se razvijaju tokom pokreta nisu merene u prvoj verziji baze koja je razmatrana u diplomskom radu. Podaci o sili, koja je korelisana sa električnom aktivnošću mišića, se mogu koristiti u cilju detekcije pokreta u sEMG signalu [16]. Osim toga, iako se sEMG signal najčešće koristi, sila može biti kontrolni signal za proteze [19]. Iz tog razloga su u drugoj verziji NINAPRO baze uključeni i podaci o sili koji su snimljeni tokom vežbanja sa fleksijom i ekstenzijom prstiju [20], [21].

4.3 Predlog za budući rad

Jedna od najbitnijih stavki prilikom unapređenja NINAPRO baze podataka je svakako uključivanje amputiranih ispitanika u protokol. Iako postoje rezultati za jednog [22] i tri [23] ispitanika sa amputiranom šakom, taj broj mora biti znatno veći za uspešnu analizu. Nakon toga bi pretprocesiranje, ekstrakciju obeležja i klasifikaciju trebalo prilagoditi dobijenim signalima, koji zapravo jesu kontrolni signali za upravljanje protezom.

U odnosu na prvu verziju NINAPRO baze podataka, druga verzija baze sadrži veliki broj unapređenja. Broj ispitanika je povećan sa 27 na 40, mada i dalje nema ispitanika sa amputiranom šakom. Osim većeg broja ispitanika, postoje značajne promene u pretprocesiranju i ekstrakciji obeležja, kao i u klasifikaciji pokreta [21]. Međutim, kao što je pokazano u ovom radu, postoji niz aspekata u kojima se mogu unaprediti ovde prikazane metode, bez potrebe da se u potpunosti napuste.

Sa stanovišta pretprocesiranja, poželjno je uključiti neki od algoritama za detekciju mišićnog onseta predloženih u [18] i iskoristiti korelisanost sile koja se razvija tokom pokreta i sEMG signala [16]. Nakon primene PCA treba razmotriti korišćenje prve tri glavne komponente kao obeležja koja se koriste za klasifikaciju. Imajući u vidu procentne udele pojedinačnih glavnih komponenti u varijabilnosti transformisanog seta podataka, ova promena bi doprinela poboljšanju rezultata klasifikacije. S obzirom na blizinu postavljenih elektroda prilikom snimanja sEMG signala i mogućeg „preslušavanja“, trebalo bi razmotriti analizu nezavisnih komponenti (eng. *Independent Component Analysis – ICA*), kao zamenu za PCA u pripremi signala za klasifikaciju. Sa stanovišta klasifikacije, u slučaju primene tri glavne komponente bi trebalo razmotriti primenu kaskadnih klasifikatora. Kaskadni klasifikator je sa velikim procentom uspešnosti korišćen za klasifikaciju pokreta u [24]. Kod kaskadnog klasifikatora je procedura klasifikacije podeljena na

Analiza glavnih komponenti elektromiografskih signala za klasifikaciju pokreta šake



više nivoa, dok su na svakom nivou obeležja odabrana tako da ima što manje preklapanja. Uzevši u obzir broj razmatranih pokreta iz NINAPRO baze i nivo preklapanja između njih u dvodimenzionalnom prostoru glavnih komponenti, kao i veliki broj subjekata i činjenicu da u protokol treba dodati i subjekte sa amputiranom šakom, razmišljanja u pravcu kaskadnog klasifikatora su sasvim opravdana.

5 LITERATURA

- [1] Internet stranica NinaPro projekta: <http://ninapro.hevs.ch/>. Baza podataka je preuzeta 27. januara 2014. godine.
- [2] Atzori M, Gijsberts A, Heynen S, Hager AG, Deriaz O, Vand der Smagt P, Castellini C, Caputo B, Müller H. Building the NINAPRO Database: A Resource for the Biorobotics Community, *Proceedings of the IEEE International Conference on Biomedical Robotics and Biomechatronics*, pp. 1258-1265, Rome, Italy, 2012.
- [3] Smit G, Bongers RM, Van der Sluis CK, Plettenburg DH. Efficiency of voluntary opening hand and hook prosthetic devices: 24 years of development?, *Journal of Rehabilitation Research and Development* 49(4):523-534, 2012.
- [4] Kovačić ZJ. Multivarijaciona analiza, Univerzitet u Beogradu, Ekonomski fakultet, Beograd, pp. 1-298, 1994.
- [5] Bosco G. Principal Component Analysis of Electromyographic Signals: An Overview, *The Open Rehabilitation Journal* 3:127-131, 2010.
- [6] Güler NF, Koçer S. Classification of EMG Signals Using PCA and FFT, *Journal of Medical Systems* 29(3):241-250, 2005.
- [7] Staudenmann D, Kingma I, Daffertshofer A, Stegeman DF, van Dieën JH. Improving EMG-based muscle force estimation by using a high-density EMG grid and principal component analysis, *IEEE Transactions on Biomedical Engineering* 53(4):712-719, 2006.
- [8] Atzori M, Gijsberts A, Heynen S, Hager AG, Castellini C, Caputo B, Müller H. Experiences in the Creation of an Electromyography Database to Help Hand Amputated Persons, *Studies in Health Technology and Informatics* 180:828-832, 2012.
- [9] Pearson K. On Lines and Planes of Closest Fit to Systems of Points in Space, *Philosophical Magazine* 2(6):559-572, 1901.
- [10] Hotelling H. Analysis of a Complex of Statistical Variables into Principal Components, *Journal of Educational Psychology* 24:417-441 and 498-520, 1933.



- [11] Esbensen KH, Geladi P. Principal component analysis: concept, geometrical interpretation, mathematical background, algorithms, history, practice, In: Brown SD, Tauler R, Walczak B. (Editors) *Comprehensive chemometrics: chemical and biochemical data analysis*, vol. 2, pp. 211-227 (chapter 2.13), Amsterdam: Elsevier Ltd., 2009.
- [12] Coffey N, Harrison AJ, Donoghue OA, Hayes K. Common functional principal components analysis: A new approach to analysing human movement data, *Human Movement Science* 30:40-45, 2011.
- [13] Maurer C, Federolf Peter, von Tscherner V, Stirling L, Nigg BM, Discrimination of gender-, speed-, and shoe-dependent movement patterns in runners using full-body kinematics, *Gait & Posture* 36(1):40-45, 2012.
- [14] Jolliffe IT. *Principal Component Analysis*, Springer, Ed. 2, pp. 1-488, 2002.
- [15] Fukunaga K. *Introduction to Statistical Pattern Recognition*, Academic Press, Ed. 2, pp. 1-591, 1990.
- [16] Castellini C, Fiorilla AE, Sandini G. Multi-subject/daily-life activity EMG-based control of mechanical hands, *Journal of Neuroengineering and Rehabilitation* 6:41, 2009.
- [17] Tenore FV, Ramos A, Fahmy A, Acharya S, Etienne-Cummings R, Thakor NV, Decoding of individuated finger movements using surface electromyography, *IEEE Transactions on Biomedical Engineering* 56(5):1427-1434, 2009.
- [18] Isaković M., Muscle Onset Detection: Methods for Online and Offline Applications, Poster presentation at the *First Conference on Brain-Computer Interface from Student-to-Student Interface (BCI from SSI)*, Academic Mind, University of Belgrade - School of Electrical Engineering, March 14, 2014.
- [19] Corbett EA, Perreault EJ, Kuiken TA. Comparison of electromyography and force as interfaces for prosthetic control, *Journal of Rehabilitation Research and Development* 48(6):629-642, 2011.
- [20] Gijssberts A, Atzori M, Castellini C, Müller H, Caputo B. Movement error rate for evaluation of machine learning methods for sEMG-based hand movement classification, *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 22(4):735-744, 2014.



- [21] Atzori M, Gijsberts A, Kuzborskij I, Heynen S, Hager AG, Deriaz O, Castellini C, Müller H, Caputo B. Characterization of a Benchmark Database for Myoelectric Movement Classification, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2014. DOI: 10.1109/TNSRE.2014.23284.
- [22] Atzori M, Baechler M, Müller H. Recognition of hand movements in a trans-radial amputated subject by sEMG. Proc of *IEEE International Conference on Rehabilitation Robotics (ICOR)*, June 24-26, 2013, Seattle, Washington, USA, pp 1-5.
- [23] Atzori M, Gijsberts A, Caputo B, Müller H. Natural Control Capabilities of Robotic Hands by Hand Amputated Subjects. *Proc of Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, August 24-30, 2014, Chicago, Illinois, USA
- [24] Tang X, Liu Z, Lv C, Sun D. Hand Motion Classification Using a Multi-Channel Surface Electromyography Sensor, *Sensors* 12:1130-1147, 2012.

PRILOG A – Sažet opis projektovanih kodova i funkcija i uputstvo za upotrebu korisničkog interfejsa projektovanog u programskom paketu MATLAB

S obzirom na obim projektovanih kodova, u ovom poglavlju će biti opisane samo glavne celine projektovanih kodova za pretprocesiranje signala, analizu glavnih komponenti i klasifikaciju signala, ako i funkcije koje su projektovane u programskom paketu MATLAB (The Mathworks, Natick, USA). Svi projektovani kodovi su dostavljeni mentoru. Ako postoji potreba za projektovanim kodovima, zainteresovani se mogu obratiti autoru na milicaisakovic@hotmail.com.

Pretprocesiranje sEMG signala

Za pretprocesiranje signala je za svakog subjekta projektovan jednostavni kod *sX_obrađa.m*, gde je X redni broj subjekta (npr. *s2_obrađa.m* za drugog ispitanika). U ovom kodu se učitavaju po tri odgovarajuća „*emg.txt*“ i „*movie.txt*“ fajla, koji odgovaraju vežbanjima na koje je merenje bilo podeljeno, poziva se projektovana funkcija *pretprocesiranje.m*, koja vrši pretprocesiranje, i rezultat pretprocesiranja se čuva u *.mat* fajlu.

Za samo pretprocesiranje signala je projektovana funkcija *pretprocesiranje.m*. Ulazne promenljive ove funkcije su po tri matrice sa podacima o sEMG signalu i stimulusu, za svako od tri vežbanja. Izlazna promenljiva je krajnji rezultat pretprocesiranja signala – 3D matrica dimenzija $20 \times 10 \times 52$, što odgovara po 10 ponavljanja pokreta i pauze, 10 elektroda i 52 pokreta. U okviru ove funkcije se koriste i druge projektovane funkcije:

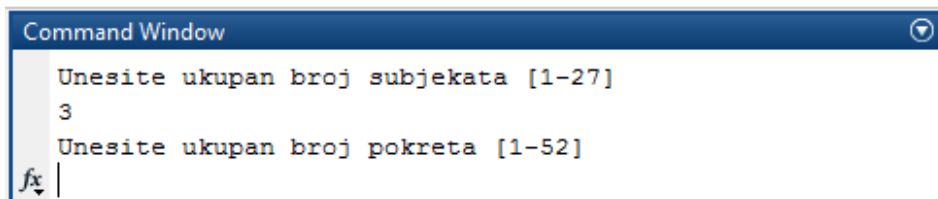
- *LPfilt_1hz.m* – filtracija sEMG signala niskopropusnim *Butterworth*-ovim filtrom drugog reda, sa graničnom učestanošću jednakom 1 Hz.
- *movie_dopuna.m* – vremenska sinhronizacija podataka o sEMG signalu i stimulusu na osnovu informacija o datumu i vremenu koji svaki odbirak poseduje. Podaci o stimulusu su prvo dopunjeni nulama, produženi ili skraćeni, u zavisnosti od konkretnog slučaja, da bi obuhvatali isti vremenski period kao sEMG signali.
- *podela_e1.m* / *podela_e2.m* / *podela_e3.m* – Ove tri funkcije su suštinski iste, ali se razlikuju po broju pokreta, pošto odgovaraju prvom, drugom i trećem vežbanju, respektivno. Pronalaze se početak i kraj svake serije od 10 ponavljanja jednog pokreta. Pronalaze se početak i kraj svakog ponavljanja pokreta i pauze. Svako ponavljanje pokreta i pauze se deli na tri segmenta. Odbirci koji pripadaju središnjim segmentima

se usrednjavaju. Vršiti se normalizacija, tako da svaki sEMG signal ima nultu srednju vrednost i jediničnu standardnu devijaciju. Izlazne promenljive svake od tri funkcija su 3D matrice dimenzija $20 \times 10 \times 12$, $20 \times 10 \times 17$ i $20 \times 10 \times 23$, respektivno.

Konkatenacijom dobijenih 3D matrica se dobija krajnji rezultat funkcije *pretprocesiranje.m* i ulazni signal za PCA.

PCA i klasifikacija pokreta

Analiza glavnih komponenti, računanje procentnih udela pojedinačnih komponenti i klasifikacija određenog seta pokreta su rezultat programa *pca.m*. Prilikom pokretanja programa, u MATLAB komandni prozor je potrebno uneti željeni broj subjekata i pokreta za koje će biti izvršena PCA – Sl. A.1. Nakon toga je iz ponuđenih menija potrebno odabrati određene ispitanike i pokrete – Sl. A.2 i Sl. A.3. U naslovu menija je istaknuto koji je po redu trenutni ispitanik/ pokret, koliko ih je ukupno potrebno odabrati i koji ispitanici/pokreti su već odabrani. U slučaju da su odabrani svi ispitanici (svih 27) ili svi pokreti (svih 52), odgovarajući meni se neće pojaviti, već će automatski biti učitani svi podaci.



```

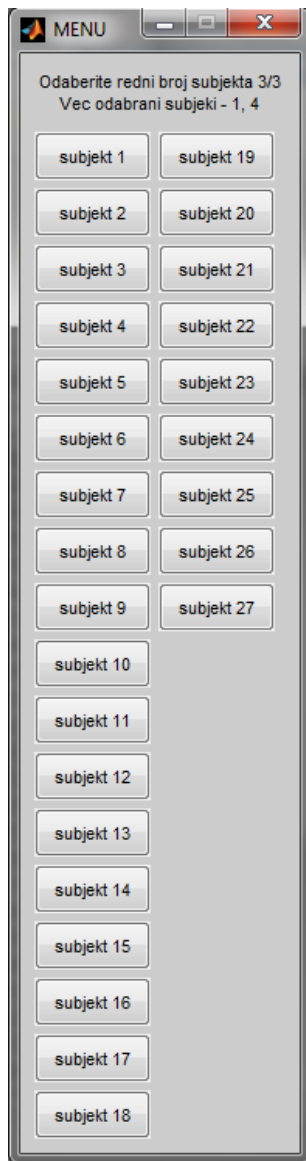
Command Window
Unesite ukupan broj subjekata [1-27]
3
Unesite ukupan broj pokreta [1-52]
fx |
  
```

Sl. A.1 Izgled MATLAB komandnog prozora prilikom odabira broja ispitanika i pokreta u projektovanoj GUI aplikaciji

Nakon odabira pokreta i ispitanika sledi učitavanje potrebnih podataka (.mat fajlovi dobijeni nakon pretprocesiranja) i formiranje odgovarajućih trodimenzionalnih matrica za pokrete i pauzu. Obe 3D matrice su istih dimenzija - $10n_s \times 10 \times n_p$, što odgovara 10 ponavljanja za svakog od n_s subjekata, 10 elektroda i n_p različitih pokreta. Analiza glavnih komponenti se sprovodi nad svakom od n_p dvodimenzionalnih matrica, kako za pokrete tako i za pauze. Računaju se kovarijacione matrice i njihve sopstvene vrednosti i sopstveni vektori. Množenjem matrica sa odbircima pokreta i sopstvenih vektora se dobijaju odbirci 10 glavnih komponenti, od kojih se za prikaz koriste samo prve dve. Finalni rezultat je dvodimenzionalni prikaz prve dve glavne komponente, crnom bojom za pauzu i različitim bojama i oznakama za različite pokrete. Slike sadrže legendu sa oznakama i rednim brojevima pokreta i naslov u kome je sadržan ukupan broj

Analiza glavnih komponenti elektromiografskih signala za klasifikaciju pokreta šake

ispitanika i pokreta, kao i konkretni redni brojevi ispitanika. Program je projektovan tako da se legenda i naslov slike automatski generišu prilikom izbora broja subjekata i pokreta.



Sl. A.2 Izgled MATLAB menija za izbor subjekata u projektovanoj GUI aplikaciji



Sl. A.3 Izgled MATLAB menija za izbor pokreta u projektovanoj GUI aplikaciji

Analiza glavnih komponenti elektromiografskih signala za klasifikaciju pokreta šake



Program je projektovan tako da izračunava i grafički prikazuje procentne udele glavnih komponenti kada su odabrane određene vrednosti ukupnog broja pokreta i ispitanika, odnosno u karakterističnim slučajevima koji ilustruju razlike koje postoje kod različitih ispitanika i pokreta. U slučaju jednog ispitanika i svih 52 pokreta se računaju procentni udeli glavnih komponenti za svaki od pokreta posebno. Te vrednosti se zatim usrednjavaju i računaju standardne devijacije za svaku glavnu komponentu. Srednje vrednosti procentnih udela se prikazuju u formi bar plot-a, sa označenim standardnim devijacijama i kumulativnom sumom. U slučaju da je izabran jedan ispitanik i tri pokreta, procentni udeli se računaju i prikazuju za svaki pokret. U slučaju da su izabrana tri ispitanika i jedan pokret, procentni udeli se računaju i prikazuju za svakog ispitanika.

Klasifikacija pokreta kvadratnim klasifikatorom se vrši u slučaju da su izabrana tri pokreta i jedan ispitanik. Za klasifikaciju po grupama, u ovom slučaju pokretima, je korišćena funkcija *classify* iz MATLAB Statistics Toolbox-a. Za tip klasifikatora je izabran kvadratni. Rezultat klasifikacije je prikazan u dvodimenzionalnom prostoru definisanom prvim dvema glavnim komponentama – deo prostora koji je klasifikovan kao određeni pokret je „osenčen“ bojom kojom su prikazani odbirci odgovarajućeg pokreta. Rezultat klasifikacije takođe obuhvata računanje ukupne greške klasifikacije.