



UNIVERZITET U NIŠU  
MAŠINSKI FAKULTET



**Marko N. Kovandžić**

**OPTIMALNO PREPOZNAVANJE I LOKALIZACIJA  
IZVORA ZVUKA PRIMENOM METODA  
VEŠTAČKE INTELIGENCIJE**

doktorska disertacija

Niš, 2020.



U N I V E R S I T Y   O F   N I Š  
F A C U L T Y   O F   M E C H A N I C A L   E N G I N E E R I N G



**Marko N. Kovandžić**

**OPTIMAL RECOGNITION AND LOCALIZATION  
OF ACOUSTIC SOURCE USING ARTIFICIAL  
INTELLIGENCE METHODS**

doctoral dissertation

Niš, 2020.

*Srbijanki i Nekojši,*

*mojim prvim učiteljima*

**Mentor:**

red. prof. dr Vlastimir Nikolić, Mašinski fakultet, Univerzitet u Nišu

**Članovi komisije za pregled, ocenu i odbranu doktorske disertacije:**

red. prof. dr Vlastimir Nikolić, Mašinski fakultet, Univerzitet u Nišu

red. prof. dr Dragan Antić, Elektronski fakultet, Univerzitet u Nišu

red. prof. dr Žarko Čojbašić, Mašinski fakultet, Univerzitet u Nišu

doc. dr Miloš Simonović, Mašinski fakultet, Univerzitet u Nišu

doc. dr Ivan Ćirić, Mašinski fakultet, Univerzitet u Nišu

**Datum odbrane:**

14.2.2020.



## PODACI O DOKTORSKOJ DISERTACIJI

*Naslov:* Optimalno prepoznavanje i lokalizacija izvora zvuka  
primenom metoda veštačke inteligencije

*Naučna oblast:* Mašinsko inženjerstvo

*Uža naučna oblast:* Automatsko upravljanje i robotika

*Broj strana:* 164

*Broj slika:* 85

*Broj tabela:* 3

*Broj literaturnih citata:* 127

*UDK broj:* 007.52:519.8]+004.8.032.26:534.88(043.3)

*CERIF klasifikacija:* Automatizacija, robotika, kontrolni inženjering

## DOCTORAL DISSERTATION DATA

*Title:* Optimal Recognition and Localization of Acoustic Source Using Artificial Intelligence Methods

*Scientific field:* Mechanical engineering

*Narrow scientific field:* Automatic control and robotics

*Number of pages:* 164

*Number of figures:* 85

*Number of tables:* 3

*Number of references:* 127

*UDK number:* 007.52:519.8]+004.8.032.26:534.88(043.3)

*CERIF classification:* T 125, Automation, robotics, control engineering

## **Zahvalnica**

Zahvaljujem nastavnicima i saradnicima Mašinskog fakulteta na saradnji i purženoj podršci za vreme pohađanja doktorskih studija. Posebnu zahvalnost dugujem svom mentoru, prof. dr Vlastimiru Nikoliću, koji me je inspirisao da preduzmem ovaj poduhvat i usmeravao tokom svakog koraka njegove realizacije.

Zahvalan sam svom dragom teči, Tomislavu Ljubisavljeviću, na rečima ohrabrenja i poklonjenoj tašni koji sam koristio svakodnevno.

## **Rezime**

Predmet disertacije je prepoznavanje i lociranje akustičnog izvora, u realnim uslovima, pomoću algoritama veštačke inteligencije. Ključni naučni cilj je razvoj optimalnog postupka opservacije akustičnih objekata uz pomoć veštačkih neuronskih mreža zbog njihove ekstremno velike operativne brzine. On treba da omogući implementaciju hibridnog sistema za akustičnu opservaciju, koji će biti u mogućnosti da prepozna zvučni izvor i prostorno ga locira u prisustvu poremećaja. Za obučavanje i verifikaciju mreža upotrebljeni su rezultati dva odvojena eksperimenta a za optimizaciju signala, u cilju povećanja robusnosti, primenjen je genetski algoritam. Rezultati naučnog istraživanja predstavljaju proširenje postojećih saznanja od značaja za dalji razvoj akustične opservacije.

## **Ključne reči**

Veštačka inteligencija, Akustična opservacija, Akustično prepoznavanje, Prepoznavanje obrazaca, Neuronske mreže, Evolucionni račun

## **Abstract**

The subject of the thesis is sound source recognition and sound source localization, in real conditions, using artificial intelligence algorithms. The main goal is optimal procedure for sound source observation using artificial neural networks for signal processing, because of their extreme processing speed. It has to provide implementation of hybrid system capable to recognize and locate sound source in the presence of disturbances. For the training and the testing of neural networks two sets of data are provided, from two different experiments, and to increase robustness genetic algorithm is applied. The results of the investigation will contribute the existing body of acoustic observation knowledge.

## **Key words**

Artificial intelligence, Acoustic perception, Acoustic recognition, Pattern recognition, Neural networks, Evolutionary algorithm

# SADRŽAJ

1. UVOD.....	1
1.1. Motiv.....	1
1.2. Predmet.....	1
1.3. Cilj .....	3
1.4. Metode .....	3
1.5. Sadržaj .....	4
2. VEŠTAČKA INTELIGENCIJA.....	7
2.1. Definicija .....	7
2.2. Simbolička veštačka inteligencija.....	11
2.2.1. Kognitivna simulacija.....	11
2.2.2. Logičko rasuđivanje .....	14
2.2.3. Reprerentacija znanja uz pomoć pravila .....	15
2.2.4. Strukturalna reprerentacija znanja.....	16
2.2.5. Princip matematičke lingvistike .....	20
2.3. Računska inteligencija .....	21
2.3.1. Konekcionistički modeli.....	22
2.3.2. Modeli bazirani na matematičkim teorijama .....	22
2.3.3. Modeli bazirani na biološkim zakonima .....	36
3. VEŠTAČKE NEURONSKE MREŽE.....	38
3.1. Model veštačkog neurona .....	40
3.1.1. Sumator.....	40
3.1.2. Aktivaciona funkcija.....	41
3.2. Topologija veštačkih neuronskih mreža .....	46
3.2.1. Jednoslojna neuronska mreža sa direktnim prostiranjem signala.....	47
3.2.2. Višeslojna neuronska mreža sa direktnim prostiranjem signala.....	47
3.2.3. Rekurentne neuronske mreže.....	48
3.2.4. Neuronske mreže sa radijalnom funkcijom .....	48
3.3. Algoritam treniranja.....	50

3.3.1. Učenje korigovanjem greške .....	51
4. EVOLUCIONI ALGORITMI .....	60
4.1. Komponente evolucionih algoritama.....	60
4.1.1. Reprerentacija jedinki .....	61
4.1.2. Inicijalizacija .....	62
4.1.3. Evaluacija .....	63
4.1.4. Modeli evolucije .....	63
4.1.5. Rekombinacija .....	65
4.1.6. Mutacija .....	66
4.1.7. Prekid algoritma .....	67
4.2. Vrste evolucionih algoritama.....	67
4.2.1. Genetski algoritam.....	67
4.2.2. Evolucione strategije .....	68
4.2.3. Evoluciono programiranje .....	68
4.2.4. Genetsko programiranje .....	69
5. PREPOZNAVANJE AKUSTIČNOG IZVORA.....	70
5.1. Mehanizam akustičnog prepoznavanja.....	70
5.1.1. Recepcija akustičnog talasa .....	70
5.1.2. Percepcija akustičnog signala .....	73
5.1.3. Kategorizacija akustičnih talasa .....	77
5.2. Eksperiment akustičnog prepoznavanja .....	79
5.2.1. Prikupljanje akustičnih uzoraka.....	79
5.2.2. Ekstrakcija osobina akustičnog signala .....	81
5.2.3. Kategorizacija akustičnih signala .....	82
5.2.4. Rezultati.....	84
5.2.5. Zaključak .....	90
6. LOKALIZACIJA AKUSTIČNOG IZVORA.....	92
6.1. Mehanizam akustične lokalizacije kod živih bića .....	92
6.1.1. Dupleks teorija.....	92

6.1.2.	Efekat ušne školjke .....	94
6.1.3.	Određivanje udaljenosti akustičnog izvora.....	95
6.2.	Mehanizam akustične lokalizacije kod veštačkih sistema.....	97
6.2.1.	Monoauralna lokalizacija .....	97
6.2.2.	Binauralna lokalizacija .....	98
6.2.3.	Prostorno filtriranje.....	98
6.2.4.	Akustična lokalizacija na osnovu vremena prostiranja talasa .....	100
6.3.	Eksperiment akustične lokalizacije u bliskom polju .....	109
6.3.1.	Prikupljanje akustičnih uzoraka.....	109
6.3.2.	Određivanje kašnjenja između signala .....	115
6.3.3.	Rešavanje sistema hiperboličkih jednačina .....	115
6.3.4.	Rezultati.....	116
6.3.5.	Zaključak .....	122
7.	PREDOBRADA I POSTOBRADA AKUSTIČNIH SIGNALA .....	124
7.1.	Metode predobrade .....	125
7.1.1.	Normalizacija.....	125
7.1.2.	Filtriranje .....	126
7.1.3.	Ekstrakcija karakteristika .....	129
7.2.	Postobrada rezultata akustične opservacije .....	130
7.2.1.	Kalmanov filter.....	130
7.3.	Eksperiment akustičnog prepoznavanja .....	133
7.3.1.	Metodologija.....	133
7.3.2.	Rezultat .....	134
7.3.3.	Zaključak .....	136
7.4.	Eksperiment lokalizacije.....	137
7.4.1.	Metodologija.....	137
7.4.2.	Rezultati.....	140
7.4.3.	Zaključak .....	146
8.	HIBRIDNI SISTEM ZA PERCEPCIJU ZVUKA.....	147



8.1. Realizacija.....	147
8.2. Zaključak .....	151
9. ZAKLJUČAK I PRAVCI DALJIH ISTRAŽIVANJA .....	152
LITERATURA .....	154

## **Lista tabela**

**Tabela 2-1** Definicije veštačke inteligencije u zavisnosti od načina rezonovanja i ponašanja..... 7

**Tabela 6-1** Težinske funkcije uopštene unakrsne korelacije ..... 108

**Tabela 6-2** Srednje kvadratna greška u zavisnosti od prostornog rasporeda mikrofona ..... 117

## Lista slika

Slika 2.1 Turing-ov test [3].....	9
Slika 2.2 Princip funkcionisanja racionalnog agenta.....	10
Slika 2.3 Metode pretraživanja prostora stanja: a) po dubini, b) po širini.....	12
Slika 2.4 Analiza sredstvo-cilj .....	14
Slika 2.5 Strategije rezonovanja: a) ulančavanje unapred, b) ulančavanje unazad .....	16
Slika 2.6 Semantička mreža sa hijerarhijskom strukturom.....	18
Slika 2.7 Reprerentacija formalnog automata: a) simbolički, b) grafički .....	21
Slika 2.8 Dvodimenzionalni prostor atributa sa tri klase obrazaca: a) regularni klasteri–metod najmanjeg rastojanja, b) iregularni klasteri–metod najbližeg suseda .....	23
Slika 2.9 Granice odluke: a) linearna, b) nelinearna.....	25
Slika 2.10 Bayes-ova granica za klase obrazaca sa različitom prethodnom distribucijom .....	27
Slika 2.11 Sukcesivno konstruisanje stabla odlučivanja i podela prostora atributa .....	28
Slika 2.12 Hijerarhijska analiza klastera.....	30
Slika 2.13 Bayes-ova mreža.....	31
Slika 2.14 Verovanje i verodostojnost.....	32
Slika 2.15 Funkcije članstva tri lingvističke vrednosti promenljive "Uzrast osobe".....	34
Slika 2.16 Grubi skupovi u prostoru atributa: a) niža granulacija, b) viša granulacija.....	36
Slika 3.1 Veštački neuron .....	38
Slika 3.2 Tipična struktura neuronske mreže [38].....	39
Slika 3.3 Blok dijagram veštačkog neurona .....	40
Slika 3.4 Aktivacione funkcije: a) odskočna, b) logička sa pragom, c) logistička, d) hiperbolička tangens .....	44
Slika 3.5 Gausove funkcija za različite vrednosti parametara .....	45
Slika 3.6 Topologija višeslojne neuronske mreže.....	46
Slika 3.7 Topologije veštačkih neuronskih mreža: a) jednoslojna, b) višeslojna, c) rekurentna ..	47
Slika 3.8 Neuronska mreža sa radijalnom funkcijom .....	50
Slika 3.9 Učenje korigovanjem greške .....	51

Slika 3.10 Razvijanje neuronske mreže u vremenu: a) rekurentna mreža, b) razvijena mreža ....	58
Slika 4.1 Blok dijagram evolucionog algoritma .....	61
Slika 5.1 Ljudsko uho [56].....	71
Slika 5.2 Dinamički mikrofons.....	72
Slika 5.3 Mikrofons sa paraboličnim reflektorom.....	73
Slika 5.4 Fletcher-Munson-ove konture jednake glasnoće [66] .....	75
Slika 5.5 Akustični signal proizveden na klaviru .....	76
Slika 5.6 Prepoznavanje kao apstrakcija u prostoru kategorija .....	78
Slika 5.7 Segmentacija akustičnog signala .....	80
Slika 5.8 Matrica konfuzije.....	84
Slika 5.9 Srednje kvadratna greška u zavisnosti od broja neurona u skrivenom sloju .....	85
Slika 5.10 Uspeh u zavisnosti od broja neurona u skrivenom sloju .....	85
Slika 5.11 Srednje kvadratna greška testiranja u zavisnosti od parametra treniranja .....	86
Slika 5.12 Srednje kvadratna greška u zavisnosti od broja epoha treniranja.....	87
Slika 5.13 Uspeh u zavisnosti od broja epoha treniranja .....	87
Slika 5.14 Srednje kvadratna greška u zavisnosti od broja uzoraka .....	88
Slika 5.15 Uspeh u zavisnosti od broja tačaka za treniranje.....	88
Slika 5.16 Srednje kvadratna greška u zavisnosti od dužine signala.....	89
Slika 5.17 Uspeh u zavisnosti od dužine signala .....	89
Slika 5.18 Matrica konfuzije u tački optimuma.....	90
Slika 6.1 Interauralna razlika akustičnog signala .....	94
Slika 6.2 Položaj ušiju u odnosu na izvor zvuka .....	96
Slika 6.3 Snimak vodenog mlaza, mašine za pranje pod visokim pritiskom, primenom akustične kamere koja funkcioniše po metodi prostornog filtriranja [100] .....	100
Slika 6.4 Akustična lokalizacija: a) daleko polje, b) blisko polje.....	101
Slika 6.5 Lokalizacija na osnovu vremena prostiranja signala .....	103
Slika 6.6 Eliptička lokalizacija u dve dimenzije .....	104
Slika 6.7 Hiperbolička lokalizacija u dve dimenzije .....	105

Slika 6.8 Grafici dva akustična signala i njihova unakrsna korelacija .....	106
Slika 6.9 Blok dijagram uopštene unakrsne korelacije.....	107
Slika 6.10 Dve varijante omnidirekcionog izvora zvuka.....	110
Slika 6.11 Akvizicija akustičnih signala namenjenih lokalizaciji u bliskom polju .....	110
Slika 6.12 Vertikalni presek eksperimentalne postavke kroz strunu .....	112
Slika 6.13 Grafički interfejs drajvera za upravljanje akvizicijom signala za potrebe akustične lokalizacije u bliskom polju .....	113
Slika 6.14 Putanja izvora zvuka u jednom kvadrantu: a) pre optimizovanja, b) nakon optimizovanja.....	114
Slika 6.15 Primer akustičnih signala istovremeno zabeleženog na 8 kanala.....	114
Slika 6.16 Srednje apsolutna greška lokalizacije u zavisnosti od broja mikrofona .....	117
Slika 6.17 Srednje apsolutna greška lokalizacije u zavisnosti od broja neurona u jedinstvenom skrivenom sloju.....	118
Slika 6.18 Srednje apsolutna greška lokalizacije u zavisnosti od broja neurona u dva skrivena sloja.....	119
Slika 6.19 Srednje apsolutna greška lokalizacije u zavisnosti od broja uzoraka za treniranje ...	120
Slika 6.20 Srednje apsolutna greška lokalizacije u zavisnosti od broja epoha treniranja.....	121
Slika 6.21 Srednje apsolutna greška lokalizacije u zavisnosti od parametara algoritma sa povratnim prostiranjem greške, koeficijenta učenja i faktora inercije.....	121
Slika 6.22 Stvarna i putanje akustičnog izvora dobijene akustičnom lokalizacijom.....	122
Slika 7.1 Tačnost akustične percepcije u zavisnosti od odnosa signala i zvuka za osobe normalnog i oštećenog sluha [120].....	124
Slika 7.2 Frekventni odziv digitalnog filtera .....	127
Slika 7.3 Amplitudno-frekventne karakteristike tipičnih vrsta filtera .....	128
Slika 7.4 Blok dijagram Kalmanovog filtera .....	132
Slika 7.5 Blok dijagram algoritma za projektovanje digitalnog filtera.....	135
Slika 7.6 Srednje kvadratna greška akustičnog prepoznavanja u zavisnosti od sadržaja poremećaja u signalu.....	135
Slika 7.7 Frekventna karakteristika digitalnog filtera u eksperimentu akustičnog prepoznavanja .....	136

Slika 7.8 Srednje kvadratna greška prepoznavanja prethodno filtriranog signala u zavisnosti od sadržaja poremećaja .....	137
Slika 7.9 Blok dijagram algoritma za projektovanje predobrade akustičnog signala namenjenog lokalizaciji u bliskom polju.....	138
Slika 7.10 Amplitudno frekventni odziv: a) poremećaj, b) signal, c) razlika.....	141
Slika 7.11 Srednja greška izračunatih kašnjenja u zavisnosti od dužine signala, za različite metode predobrade akustičnog signala .....	141
Slika 7.12 Srednja greška izračunatih kašnjenja u zavisnosti od frekvencije presecanja visokopropusnog filtera .....	142
Slika 7.13 Srednja greška izračunatih kašnjenja u zavisnosti od relativne vrednosti maksimalno dopuštenog kašnjenja .....	143
Slika 7.14 Frekventni odziv sa optimalnim efektom na smanjenje greške izračunatih kašnjenja kod uzoraka koji traju 131 ms.....	145
Slika 7.15 Srednja greška izračunatih kašnjenja u zavisnosti od broja decimalnih mesta u parametrima predobrade .....	145
Slika 8.1 Blok dijagram hibridnog sistema za percepciju akustičnih signala.....	148
Slika 8.2 Baza podataka za realizaciju strategije prepoznavanja.....	149
Slika 8.3 Slika akustičnog izvora dobijena metodom prostornog filtriranja .....	150

# 1. UVOD

Kako se širi primena upravljačkih sistema na različite oblasti života tako raste potreba za njihovom interakcijom sa okolinom. U percepciji okolne dominiraju vizuelni sistemi zbog činjenice da predstavljaju prirodan izbor čoveka u interakciji sa njom jer pružaju najveću količinu informacija. Međutim, akvizicija i obrada vizuelnih signala je skupa i kompleksna a osetljivost na poremećaje velika pa rezultati često nisu dovoljno pouzdani. Sa druge strane, zadovoljavajuća interakcija sa okolinom se ponekad može obaviti sa mnogo manje raspoloživih resursa. Zbog toga je planirano istraživanje orijentisano na detekciju akustičnih signala što može predstavljati daleko bolju alternativu sa stanovišta cene, jednostavnosti i pouzdanosti. U literaturi su odavno poznate različite metode akustične opservacije od kojih svaka zahteva specifične preuslove i u skladu sa njima pruža ograničene rezultate. U cilju dobijanja potpunije slike okoline potrebno je kombinovati više metoda akustične opservacije kao i metode filtriranja signala, namenjene smanjenju uticaja poremećaja. Svaka od metoda zahteva određen broj matematičkih operacija a njihov ukupan broj predstavlja osnovni limitirajući faktor primene.

## 1.1. Motiv

Uprkos napretku teoretskog znanja i kapaciteta jedinica za obradu podataka razvoj veštačkih sistema za percepcije zvuka je u inicijalnoj fazi u poređenju sa prirodnom sposobnosti čoveka i životinja da čuju. Najuspešnija rešenja zahtevaju najveći kapacitet izračunavanja pa se retko mogu primeniti u praksi. Time je otvoreno veliko polje za istraživanje u cilju unapređenja akustičnih sistema namenjenih opservaciji okoline. Glavna prepreka leži u potrebi kategorizacije (grupisanja) objekata u referentni okvir na osnovu zajedničkih karakteristika zvuka (visina, boja, trajanje, jačina...). Ovo je kognitivni proces koji podrazumeva apstrakciju rezultata opservacije na bazi prethodnog iskustva i teško se može obaviti klasičnim metodama matematičke analize. U istraživanju su predviđeni metodi veštačke inteligencije jer funkcionišu po ugledu na biološke sisteme.

## 1.2. Predmet

Predmet planiranih istraživanja je definisanje optimalnog postupka za konfiguraciju sistema akustične opservacije primenom metoda veštačke inteligencije. Za razliku od vizuelnih,

kod akustičnih sistema postupci lociranja i prepoznavanja izvora signala obavljaju se apsolutno nezavisno na osnovu različitih osobina zvuka.

Zbog iskonske potrebe čoveka da izdaje komande putem govora većina radova o percepciji zvuka posvećena je percepciji govora. Ovo istraživanje isključuje ljudski govor i ograničava se na ostale, daleko jednostavnije slučajeve percepcije zvuka koji pružaju manje informacija ali zahtevaju daleko manje resursa. Predviđeno je eksperimentalno ispitivanje i verifikacija postupka prepoznavanja izvora zvuka obradom signala u frekventnom domenu. Za ispitivanje su upotrebljeni uzorci koji pripadaju trima kategorijama različitog nivoa apstrakcije. Prvu, najspecifičniju, kategoriju čine akustični uzorci jednog istog cvrčka, drugu čine uzorci različitih jedinki muve, koji pripadaju različitim podgrupama i kategorijama, i treću čine uzorci različitog porekla počev ljudskog glasa, životinja, prirodnih fenomena do različitih postrojenja, vozila, mašina, muzičkih instrumenata i ostalih tehničkih uređaja. Ova grupa treba da simulira najširi spektar različitih zvukova koji mogu da ometaju prepoznavanje u realnim okolnostima. Za obradu signala primenjene su različite konfiguracije neuronskih mreža i odabrana je optimalna.

Lokalizacija izvora zvuka je daleko više obrađena u literaturi. Ona se obavlja obradom signala sakupljenih, istovremeno, sa više mikrofona na bazi činjenice da zvuk stiže do svakog od njih u različitim trenucima. U odnosu na položaj i raspored mikrofona razlikujemo dva bazična slučaja, prvi kada su senzori udaljeni od izvora i drugi kada su njegovoj okolini. Utisak je da je prvi slučaj daleko više primenjen zbog daleko jednostavnijeg matematičkog aparata. U drugom slučaju primena geometrijskih relacija i zakona o brzini prostiranja zvuka rezultuje preodređenim sistemom hiperboličnih jednačina čije rešenje predstavlja nelinearni optimizacioni problem. Istraživanjem je predviđeno testiranje neuronskih mreža različitih konfiguracija za rešavanje ovog problema. U tu svrhu konstruisan je sistem za akviziciju signala koji navodi izvor zvuka duž slučajno izabrane putanje i istovremeno sakuplja akustične uzorke sa 8 prostorno raspoređenih mikrofona. Osim optimalne konfiguracije neuronske mreže ovde se nameće problem optimalnog broja i rasporeda mikrofona.

Posebna pažnja predviđena je ispitivanju robusnosti prethodnih postupaka u odnosu na poremećaje. U cilju kompenzovanja poremećaja predviđena je optimizacija signala koja je takođe izvršena primenom veštačke inteligencije.



### **1.3. Cilj**

Ključni naučni cilj istraživanja je razvoj efikasne metode za opservaciju akustičnih objekata u prostoru. Realizacija ovog cilja predviđena je u okviru sledećih faza:

- Analiza inteligentnih algoritama za prepoznavanje akustičnih objekata
- Razvoj, konfigurisanje, testiranje i eksperimentalna verifikacija optimalnog algoritma za prepoznavanje akustičnih objekata, u realnim uslovima, na osnovu obrade frekventnog spektra signala primenom neuronske mreže
- Analiza inteligentnih algoritama za lociranje objekata na osnovu zvuka
- Razvoj, konfigurisanje, testiranje i eksperimentalna verifikacija optimalnog sistema za lokalizaciju izvora zvuka. Preduslov je da signali potiču sa niza mikrofona koji su prostorno raspoređeni u okolini izvora.
- Analiza uticaja poremećaja na robusnost opservacije akustičnog izvora i inteligentnih algoritama namenjenih prigušenju poremećaja
- Razvoj, konfigurisanje, testiranje i eksperimentalna verifikacija genetskog algoritma za optimizaciju zvučnih signala sa ciljem povećanja robusnosti akustične opservacije
- Razvoj hibridnog sistema za akustičnu opservaciju koji pruža što je moguće više informacija o posmatranom objektu.

Cilj se neće smatrati ispunjenim dok sistem ne bude mogao da funkcioniše u realnim uslovima. On mora biti robusan u odnosu na sveprisutne akustične poremećaje uprkos svojoj jednostavnosti. U svrhu povećanja robusnosti predviđena je optimizacija postupka predobrade signala.

### **1.4. Metode**

U svrhu ostvarenja ciljeva doktorske disertacije izvršena je detaljna analiza relevantnih postupaka akustične opservacije. Ova analiza uključuje različite aspekte problema počev od psiholoških preko matematičkih zaključno sa tehničkim odnosno aspektima realizacije u realnim uslovima. Izdvojene su metode koje najbolje odgovaraju pretpostavljenim uslovima. Kao trag za prepoznavanje akustičnog izvora upotrebljen je frekventni spektar signala a za njegovu lokalizaciju kašnjenje između signala sakupljenim na različitim prostornim lokacijama. Za

obradu signala, u oba slučaja, primenjena je veštačka neuronska mreža sa direktnim prostiranjem signala koja je, na osnovu raspoloživih podataka, trenirana pomoću algoritma sa povratnim prostiranjem greške i inercijom. Konfiguracija mreže je izvršena eksperimentalno a postupak predobrade signala, u cilju povećanja tačnosti rezultata, obavljen je pomoću evolucione strategije, takođe na osnovu skupa raspoloživih podataka.

Radi eksperimentalne potvrde teoretskih razmatranja preduzeta su tri eksperimenta:

- Prepoznavanje (klasifikacija) izvora zvuka, na osnovu obrade digitalnog spektra, primenom veštačke neuronske mreže sa direktnim prostiranjem signala. Za potrebe ovog eksperimenta zabeleženo je 685 uzoraka, sakupljenih u laboratoriji ili sa interneta, koji prema nivou apstrakcije pripadaju trima različitim kategorijama.
- lokalizacija zvučnog izvora u okolini, na osnovu kašnjenja među signalima zabeleženim na različitim prostornim lokacijama, primenom veštačke neuronske mreže sa direktnim prostiranjem. Za potrebe ovog eksperimenta snimljeno je preko 2000 uzoraka, uz pomoć 8 prostorno raspoređenih mikrofona, i svakom od njih pridružene su odgovarajuće prostorne koordinate dobijene laserskim merenjem.
- optimizacija signala, u slučaju eksperimenata prepoznavanja i lokalizacije, primenom evolucionog računa, na osnovu skupa raspoloživih podataka, sa ciljem minimizacije greške rezultata opservacije.

Eksperimentalni metod sistema za prepoznavanje i lociranje akustičnih izvora sa prigušenjem šuma čini oprema i softver koji mogu biti podeljeni u tri funkcionalne celine namenjene:

- akviziciji i obradi akustičnih signala
- obučavanju sistema
- testiranju sistema

## **1.5. Sadržaj**

Istraživanja, po sadržaju, obuhvataju više aktuelnih naučnoistraživačkih pravaca koja se odnose na:

- Analizu postupaka za prepoznavanje (klasifikaciju) akustičnih izvora

- Razvoj neuronske mreže za prepoznavanje akustičnog izvora u realnim uslovima
- Analiza postupaka za lokalizaciju izvora zvuka
- Razvoj platforme za akviziciju zvučnih signala u okolini izvora zvuka
- Razvoj neuronske mreže za lokalizaciju akustičnog izvora u prostoru na osnovu signala sakupljenih u njegovoj okolini
- Analiza metoda veštačke inteligencije za optimizaciju digitalnih signala
- Dizajn optimalnog filtera za kompenzaciju poremećaja pri opservaciji akustičnog signala u cilju minimizacije greške

U drugom poglavlju objašnjen je predmet izučavanja veštačke inteligencije, definisani metodi kojima se ona služi i napravljena podela na osnovu pristupa rešavanju problema. Predstavljeni su reprezentativni algoritmi veštačke inteligencije sa osvrtom na računsku inteligenciju s obzirom da je ona predmet izučavanja disertacije.

Treće poglavlje posvećeno je veštačkim neuronskim mrežama jer su one u postupcima akustične opservacije bile primenjene kao sredstvo za obradu podataka. U ovom poglavlju objašnjena je struktura neuronskih mreža, princip funkcionisanja i način njihovog obučavanja na osnovu skupa podataka za treniranje.

U četvrtom poglavlju objašnjen je evolucioni račun pošto je on bio primenjen za konfiguraciju postupaka predobrade signala. Objašnjena je podela evolucionog računa, na 4 opšte kategorije, i u skladu sa ovom podelom namena, osnovni operatori i smernice za izvođenje algoritama.

Prepoznavanje akustičnih izvora razmatrano je u petom poglavlju. Objašnjen je mehanizam prepoznavanja, osnovni perceptivni kvaliteti na osnovu kojih se ono vrši i metode za njegovo izvođenje. Drugi deo ovog poglavlja objašnjava eksperiment akustičnog prepoznavanja, primenom veštačkih neuronskih mreža, na osnovu frekventnog spektra.

Šesto poglavlje posvećeno je akustičnoj lokalizaciji. Predstavljena je metodologija i fizički fenomeni na osnovu kojih se ona vrši u prirodi i od strane veštačkih sistema. Dat je pregled metoda lokalizacije i metoda za predobradu signala namenjenih ovoj svrsi. Opisan je eksperiment akustične lokalizacije i prezentovani njegovi rezultati.

Sedmo poglavlje objašnjava uticaj poremećaja na tačnost akustične opservacije i daje predlog mera za njihovo prigušenje. Podaci sakupljeni u eksperimentima akustičnog prepoznavanja i akustične lokalizacije upotrebljeni su za izvođenje eksperimenta sa postupcima predobrade signala. Predložen je metod za konfigurisanje postupka predobrade akustičnih signala i prezentovani rezultati njegove primene.

U osmom poglavlju predložen je model hibridnog sistema za akustičnu opservaciju projektovan u skladu sa principima efikasnosti, jednostavnosti i univerzalnosti. Model ostvaruje strategiju opservacije kombinovanjem različitih metoda veštačke inteligencije u cilju prikupljana što veće količine informacija o posmatranom objektu.

Poslednje poglavlje rezervisano je za izvođenje zaključaka i predstavljanje planova daljih istraživanja na polju akustične opservacije.

## 2. VEŠTAČKA INTELIGENCIJA

Prve ideje o veštačkoj inteligenciji sreću se još u antičkoj kulturi ali je ona ostala u sferi fikcije sve do prve polovine 20. veka. Pojava računara, koji su omogućili njenu fizičku realizaciju, označila je početak procvata ove naučne discipline kome se ne nazire kraj ni do današnjih dana. Za kratko vreme veštačka inteligencija je integrisana u sve oblasti ljudskog delovanja, do te mere da su evidentne čak i brojne ekonomske, političke i sociološke posledice.

### 2.1. Definicija

Uprkos intenzivnom istraživanju i masovnoj primeni, pojam veštačke inteligencije nije jednostavno definisati. U Tabeli 1 date su različite definicije veštačke inteligencije u zavisnosti od predviđenog načina rezonovanja i ponašanja [1].

Prvi pristup, kome odgovara prva definicija u prvom redu, fokusira se na ljudsko rasuđivanje i u odnosu na njega meri uspeh veštačke inteligencije. Da bi ovo bilo moguće neophodno je razumeti mehanizam rasuđivanja kod ljudi i životinja. Ovaj proces je predmet istraživanja interdisciplinarnе oblasti nauke o umu i inteligenciji koja se naziva kognitivna nauka. Najčešće metode istraživanja kognitivne nauke su:

- **introspekcija** ili analiza sopstvenih misli
- **psihološki eksperimenti**, koji podrazumevaju posmatranje ljudi ili životinja za vreme obavljanja neke radnje
- **slikanje mozga**, različitim tehnikama [2], za vreme obavljanja funkcije

**Tabela 2-1** Definicije veštačke inteligencije u zavisnosti od načina rezonovanja i ponašanja

	<b>Ljudski</b>	<b>Racionalno</b>
<b>Rasuđivanje</b>	“[Automatizovanje] aktivnosti koje povezujemo sa ljudskim rasuđivanjem, kao što su donošenje odluka, rešavanje problema, učenje...” (Bellman, 1978)	“Proučavanje algoritama koji imaju sposobnost opažanja, rasuđivanja i delovanja”. (Winston, 1992)
<b>Delovanje</b>	“Veština pravljenja mašina za obavljanje funkcija koje, kad ih obavljaju ljudi, zahtevaju inteligenciju.” (Kurzweil, 1990)	“Računska inteligencija je nauka o projektovanju inteligentnih agenata.” (Poole, 1998)

Kognitivne funkcije nisu predmet istraživanja tehničkih nauka, s obzirom da se utvrđuju na realnim ljudima i životinjama, ali se rezultati ovih istraživanja primenjuju intenzivno u sistemima veštačke inteligencije.

Prva definicija u drugom redu (**Tabela 2-1**) odnosi se na sisteme veštačke inteligencije koji oponašaju delovanje čoveka. Za ocenu performansi ovih sistema potrebna je zadovoljavajuća operativna definicija inteligencije. Ovakvu definiciju, još 1950. god., dao je engleski matematičar Alan Turing u vidu testa koji je uz različite izmene i u različitim varijantama, ostao opšte prihvaćen sve do današnjih dana. Sistem veštačke inteligencije (kompjuter A na slici 2.1) prolazi Turing-ov test ako ispitivač (čovek C) ne pravi razliku između njegovih i odgovora dobijenih od čoveka (B). U svom osnovnom obliku Turing-ov test od sistema veštačke inteligencije zahteva sledeće sposobnosti:

- **obrada prirodnih jezika** radi ostvarenja uspešne komunikacije
- **reprezentacija znanja** u cilju njegovog skladištenja
- **automatsko rezonovanje** odnosno upotreba uskladištenog znanja za donošenje zaključaka i odgovora na postavljena pitanja
- **mašinsko učenje** u cilju adaptacije na nove okolnosti

Osnovni oblik Turing-ovog testa polazi od pretpostavke da fizička interakcija između računara i čoveka nije bitna za merenje inteligencija pa se zbog toga pitanja postavljaju pismenim putem. Totalni Turing-ov test, međutim, uključuje još dve sposobnosti

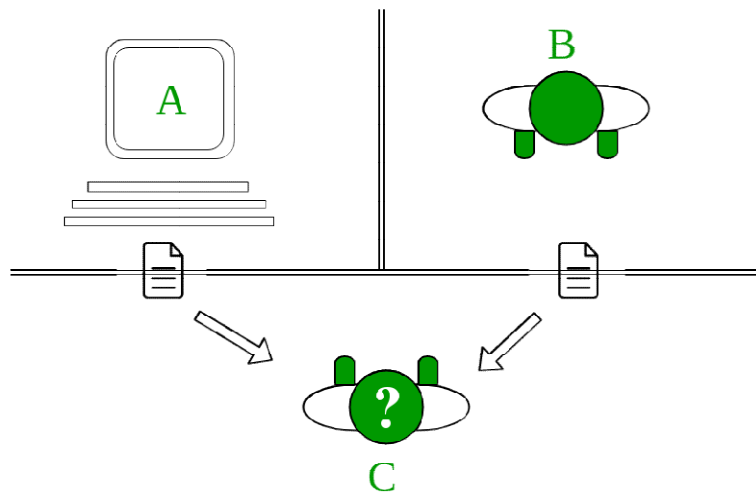
- **Percepciju okoline** na osnovu različitih signala (vizuelnih, mehaničkih, hemijskih, električnih, termičkih, magnetnih...)
- **Motoričke sposobnosti** koje omogućavaju kretanje i manipulaciju objektima u prostoru

Percepcija okoline predmet je izučavanja **računarske vizije** a motoričke sposobnosti su predmet izučavanja **robotike**, posebnih naučnih disciplina koje su nastale kao posledica ekspanzije veštačke inteligencije.

U skladu sa savremenim tendencijama u razvoju saznanja, koje podrazumevaju napuštanje modela kopiranja i imitiranja u korist kreativnosti i inovativnosti, veštačka inteligencija se sve više udaljava od čoveka kao referentnog modela ponašanja. Fokus interesovanja, naročito

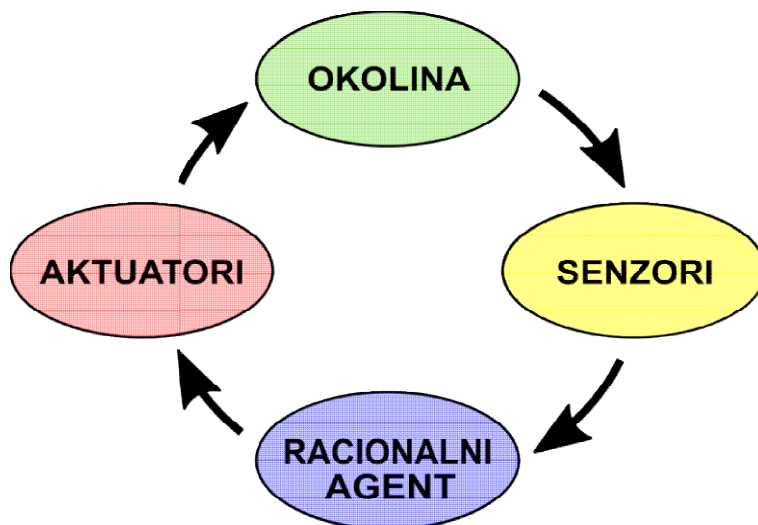
tehničkih disciplina, pomera se ka racionalnom ponašanju. Pri ostvarenju ovakvog ponašanja, pošlo se od osnova racionalnog ponašanja odnosno osnova logike koje je postavio Aristotel kreirajući svoje **hipotetičke silogizme**. To su strukture logičkih tvrdnji koje nepobitno vode ka tačnom zaključku. Aristotelova logika bazira se na tri zakona:

- **Identitet:** sve što je istinito mora se potpuno podudarati samo sa sobom
- **Kontradikcija:** nemoguće je jednom istom na isti način jedno isto dodati i ne dodati
- **Isključenja trećeg:** između dva kontradiktorna stava ne može biti trećeg



*Slika 2.1 Turing-ov test [3]*

U 19. veku postavljeni su temelji matematičke logike, mada postoje indicije da su ovi principi u Indiji bili primenjivani daleko ranije. Objavljivanjem knjige o zakonima rasuđivanja (engl. “The Laws of Thought”), 1854. god., u kojoj prezentuje svoju algebru nad logičkim iskazima, engleski matematičar Boole (George Boole, 1818-1864) označava početak informatičkog doba. Već sredinom sedamdesetih godina 20. veka sistemi veštačke inteligencije bazirani na logističkom pristupu (desna definicija u prvom redu, Tabela 2-1) bili su sposobni da reše bilo koji problem predstavljen u logičkoj notaciji (pod uslovom da rešenje postoji). Sistemi koji teže logičkom rasuđivanju daju izuzetne rezultate u principijelnom rešavanju problema. Međutim, u praktičnim uslovima, naročito kada postoji veliki broj iskaza ili kada iskazi nisu apsolutno tačni, performanse ovih sistema drastično opadaju.



*Slika 2.2 Princip funkcionisanja racionalnog agenta*

Za rešavanje problema kod kojih logičko rasuđivanje ne daje zadovoljavajuće rezultate primenjuju se racionalni agenti. Racionalni agent je bilo koji program sa sledećim karakteristikama:

- Autonomna kontrola
- Percepcija okoline
- Prilagođavanje okolini
- Delovanje za vreme produženih vremenskih intervala
- Težnja ka postavljenom cilju

Na osnovu prethodnih karakteristika jasno je da ovom vidu veštačke inteligencije odgovara druga desna definicija u drugom redu tabele 2-1.

Funkcionalna šema racionalnog agenta prikazana je na slici 2.2. On formira sliku o okolini na osnovu različitih uređaja koji se nazivaju **senzori**. Na osnovu stanja okoline, racionalni agent odlučuje kako na nju treba delovati kako bi ona menjala ka ostvarenju predviđenog cilja. Odlučivanje može ali ne mora biti zasnovano na logičkom rasuđivanju. Primer racionalnog delovanja koje ne uključuje logičko rasuđivanje je refleksna radnja koju ispoljavaju ljudi i životinje. Neposredni uticaj racionalnog agenta na okolinu ostvaruje se pomoću uređaja koji se nazivaju aktuatori.



U pogledu načina predstavljanja modela razlikuju se simbolička veštačka inteligencija, kod koje je model predstavljen pomoću simbola razumljivih ljudima i računski odnosno numerička veštačka inteligencija.

## **2.2. Simbolička veštačka inteligencija**

Simbolička veštačka inteligencija polazi od pretpostavke da se inteligentno ponašanje može postići manipulacijom nad simbolima. Ovakvu pretpostavku formalno su definisali Newell (Allen Newell, 1927-1992) i Simon (Herbert A. Simon, 1916-2001) u svojoj hipotezi o fizičkom sistemu simbola [4]. Simbolički sistemi odlikuju se sledećim karakteristikama:

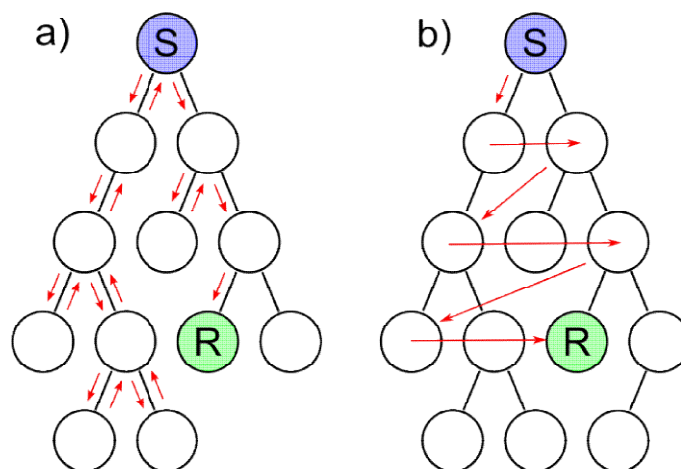
- Model inteligentnog sistema definisan je eksplicitno
- Znanje je predstavljeno simbolički
- Kognitivne operacije su predstavljene u vidu formalnih operacija nad simboličkim izrazima i strukturama koje pripadaju modelu znanja

U pogledu načina ponašanja simbolička veštačke inteligencija može biti orijentisana na rasuđivanje ili na delovanje. Kognitivna simulacija i logičko rasuđivanje pripadaju tehnikama veštačke inteligencije koje su orijentisane na rasuđivanje. Ove metode teže opštem modelu predstavljanja znanja i inteligentnih operacija. U pogledu načina rasuđivanja, kognitivna simulacija teži ljudskom a logičko rasuđivanje racionalnom modelu rasuđivanja (Tabela 2-1). U okviru metoda simboličke veštačke inteligencije orijentisanih ka delovanju primenjuju se tri načina predstavljanja znanja. To su reprezentacija znanja uz pomoć pravila, strukturalna reprezentacija znanja i pristup matematičke lingvistike.

### **2.2.1. Kognitivna simulacija**

Kognitivna simulacija primenjuje heurističke algoritme kako bi simulirala funkcija ljudskog uma. Temelje kognitivne simulacije, po mnogima i temelje veštačke inteligencije, postavili su, sredinom šezdesetih godina 20. veka, Allen i Simon, kreiranjem programa koga su nazvali “Logički teoretičar” (engl. “Logic Theorist”). Smatra se prvim programom koji je pokazao uspeha u oponašanju ljudskog rezonovanja prilikom rešavanja problema [5]. Četiri osnovna koncepta kognitivne simulacije su prostor stanja, pretraživanje prostora stanja, analiza sredstvo-cilj i redukcija problema.

**Koncept prostora stanja** podrazumeva definisanje svih mogućih stanja (kružići na slici 2.3) na putu nalaženja rešenja počev od inicijalnog stanja (kružić S), koje odgovara početku rešavanja problema, preko svih dozvoljenih prelaznih stanja (prazni kružići) zaključno sa ciljnim stanjem (kružić R), koje odgovara rešenju problema. Osim dozvoljenih stanja neophodno je definisati sve dozvoljene tranzicije (prelaze) između stanja (na slici obeležene linijama). Optimalni put od početnog stanja do rešenja bira se sa ciljem minimiziranja utrošenih resursa.



*Slika 2.3 Metode pretraživanja prostora stanja: a) po dubini, b) po širini*

**Pretraživanje prostora stanja** vrši se u odsustvu pouzdanih pravila koji bi doveli do rešenja problema. Primenuje se na strukture podataka kao što su stabla i grafovi u cilju nalaženja najboljeg stanja u odnosu na unapred zadati kriterijum. Pretraživanje počinje od korena stabla i nastavlja se prema nekom prioritetu. Ako kriterijum pretraživanja čini isključivo **funkcija uspeha**, odnosno binarna informacija o tome da li je ili nije dostignuto konačno rešenje, pretraživanje se vrši nasumično. Ovaj vid pretraživanja prostora stanja naziva se **neinformisano pretraživanje**. Primer je pogađanje zadate šifre (lozinke). U odsustvu **funkcije koštanja**, odnosno informacije o kvalitetu trenutnog stanja, prioritet pretraživanja određuje se u odnosu na strukturu prostora stanja. U zavisnosti od prioriteta pretraživanja metode neinformisanog pretraživanja su:

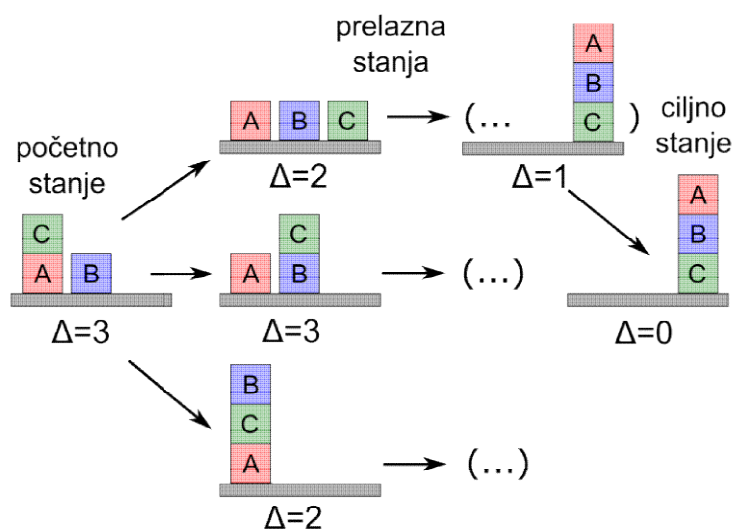
- **Pretraživanje u dubinu** (engl. Depth-first search) po prioritetu potpunog ispitivanja jedne grane pre nego što se pređe na sledeću [6]. Tok pretraživanja u dubinu predstavljen je crvenim strelicama na slici 2.3 a.
- **Pretraživanje u širinu** (engl. Breadth-first search) po prioritetu potpunog ispitivanja jednog nivoa strukture prostora stanja pre nego što se pređe na sledeći [7]. Tok pretraživanja u širinu predstavljen je crvenim strelicama na slici 2.3 b.
- **Pretraživanje po kriterijumu najmanjih troškova** [8] svakom stanju pridružuje se funkcija koštanja koja se, u nedostatku prave informacije koštanju, izračunava na osnovu njegovog položaja u strukturi prostora stanja. Najbolje rešenje bira se sa ciljem minimizacije ove funkcije.

Ako kriterijum pretraživanja pruža informaciju o blizini rešenja, pretraživanje može biti usmereno na regije sa najvećim potencijalom za njegovo nalaženje. Ovakav vid pretraživanja naziva se **informisano pretraživanje**. U temperaturnom polju, na primer, verovatno je da prostorna lokacija sa maksimalnom temperaturom bude pronađena u okolini lokacije sa trenutno najvišom izmerenom temperaturom. Primeri metoda informisanog pretraživanja su:

- **Pretraživanje po kriterijumu pohlepe (gladi)** (engl. Greedy best-first search) se vrši po prioritetu minimizacije heurističke funkcije koštanja. Ova funkcija izračunava troškove puta od trenutnog stanja do cilja [9]. Funkcija koštanja zavisi od prirode problema.
- **A\* pretraživanje** (engl. A\* search) se vrši po prioritetu minimizacije funkcije koštanja koja [10] se dobija kao zbir troškova putanje do trenutne pozicije i očekivanog puta do rešenja. Trošak puta od trenutnog stanja do konačnog rešenja ocenjuje se pomoću heurističke funkcije koja zavisi od prirode problema.

**Analiza Sredstvo-Cilj** (engl. Means-Ends Analysis) se primenjuje ako je na raspolaganju funkcija za ocenu kvaliteta stanja. Prema ovom konceptu, za promenu iz jednog u drugo stanje potrebno je određeno sredstvo. Između različitih sredstava bira se ono koje će najviše povećati kvalitet stanja odnosno smanjiti rastojanje do konačnog cilja. Slika 2.4 ilustruje standardni problem veštačke inteligencije koji se sastoji u premeštanju blokova [11]. Potrebno je počev od početnog stanja, premeštanjem po jednog bloka u svakom koraku, doći do rešenja problema

odnosno ciljnog stanja. Ograničenja su da nije dozvoljeno podići dva bloka ili izvući blok koji stoji ispod drugog bloka. Analiza sredstvo-cilj vrši se u odnosu na ciljno stanje, u ovom slučaju definisano sa tri tvrđenja: A je na B, B je na C i C je na stolu. Funkcija odstupanja se definiše kao broj netačnih među navedenim tvrđenjima. Primera radi za početno stanje  $\Delta = 3$  jer ni jedno od tvrđenja ne stoji. Polazeći od nekog stanja, biraju se rešenja koja vode rešenju problema odnosno ona sa manjom funkcijom odstupanja. U primeru sa slike to su prvo i poslednje prelazno stanje, koje imaju odstupanje  $\Delta = 2$ . Postupak se izvršava iterativno do postizanja konačnog rešenja kome odgovara funkcija odstupanja  $\Delta = 0$ .



Slika 2.4 Analiza sredstvo-cilj

**Redukcija problema** se sastoji u podeli kompleksnog problema na niz jednostavnijih koji se rešavaju pojedinačno. Svakom od pojedinačnih problema odgovara poseban prostor stanja, posebne funkcije kvaliteta i posebni operatori [12]. Skup svih parcijalnih rešenja zapravo čini globalno rešenje problema.

### 2.2.2. Logičko rasuđivanje

Ovaj vid simboličke veštačke inteligencije bazira se na tvrđenju da inteligentni sistem može biti kreiran pomoću formalizovanih modela logičkog razmišljanja (John McCarthy, 1950). Ovaj pristup od programera očekuje da definiše (deklariše) šta su karakteristike rešenja i kako do njega doći. Zbog toga pripada kategoriji **deklarativnog programiranja**. Samo rešavanje

problema prepušteno je **univerzalnom programu** za rešavanje problema (engl. General Problem Solver). Postoje dva pristupa logičkom rasuđivanju :

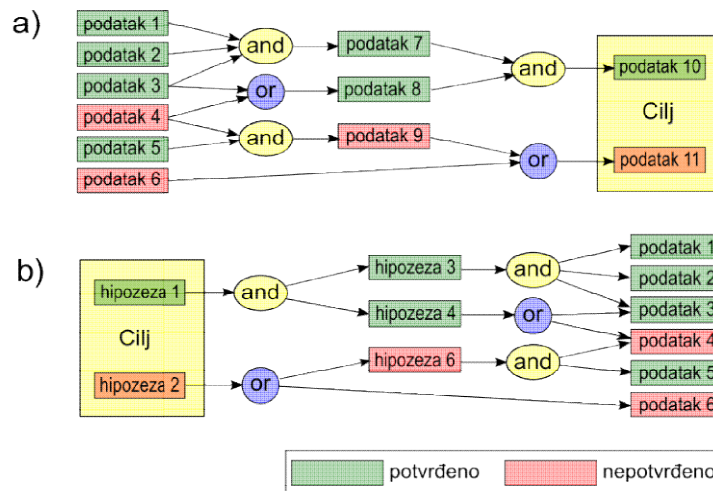
- **Logičko programiranje**, zahteva opisivanje željenih karakteristika rešenja pomoću logičkih teorema (hipoteza) [13]. Na osnovu ovih hipoteza, logičkih pravila i činjenica koje su smeštene u bazi znanja, univerzalni program donosi zaključke odnosno obavlja proces rasuđivanja.
- **Funkcionalno programiranje**, zahteva opisivanje željenih karakteristika rešenja u vidu kompleksne funkcije [14]. Ovo moguće zahvaljujući primeni **lambda računa**, formalnog sistema za predstavljanje računskih funkcija pomoću visoko formalizovanih izraza. Univerzalni program mora biti sposoban da vrši izračunavanje ovih izraza u skladu sa pravilima lambda računa.

### 2.2.3. Reprezentacija znanja uz pomoć pravila

1972. god. Newell i Simon su predložili model produkcionog sistema (engl. Production System Model). Ovaj sistem sastoji se od niza pravila ponašanja i mehanizma koji obezbeđuje da se ova pravila slede u cilju ispoljavanja inteligentnog ponašanja. Pravila ponašanja, sa funkcijom okidača, smeštena su u **produkcionoj** (dugotrajnoj) memoriji. **Okidač** je definisan uslovom i odgovarajućom radnjom koja treba da usledi u slučaju ispunjenja uslova. Podaci odnosno činjenice o okruženju čuvaju se u **radnoj** (kratkotrajnoj) memoriji. Treći element modela je **interpretator pravila** koji obavlja neprekidno testiranje činjenica iz radne memorije na osnovu pravila smeštenih u dugotrajnoj memoriji [15]. Svako ispunjenje kriterijuma inicira odgovarajuću radnju. Ove radnje mogu biti raznolike, počev od donošenja zaključka i njegovog smeštanja u radnu memoriju do kontrole različitih naprava. Praksa je pokazala da je formulisanje pravila u vidu okidača znatno prihvatljivije za širi krug eksperata koji učestvuju u kreiranju baza znanja od formalnih logičkih izraza. Postoje dve dominantne strategije rezonovanja u sistemima baziranim na pravilima.

**Ulančavanje unapred** (engl. Forward chaining) ili **pretraživanje vođeno činjenicama** (engl. Data Driven Search) je strategija **rezonovanja** koja predviđa kretanje od uzročnog ka posledičnom. Zbog toga se naziva još i reaktivna strategija rezonovanja. Ova strategija polazi od zadatih podataka i, u zavisnosti od ispunjenosti uslova, napreduje prema zaključku. Zaključak se

skladišti u kratkotrajnoj memoriji čime se povećava baza znanja [16]. Proces se ciklično ponavlja do ostvarenja konačnog cilja. Strategija ima sličnosti sa pretraživanjem po širini. Pogodna je za rešavanje problema sa malo podataka a puno prihvatljivih rešenja kao što su planiranje, nadgledanje i upravljanje.



Slika 2.5 Strategije rezonovanja: a) ulančavanje unapred, b) ulančavanje unazad

**Ulančavanje unazad** (engl. Backward chaining) ili **pretraživanje vođeno ciljem** je strategija (racionalnog) **rezonovanja** koja predviđa kretanje od posledičnog ka uzročnom. Ona polazi od rešenja, kao zaključka, tražeći pravila i činjenice koje ima za posledicu taj zaključak. Pravila čiji je uslov zadovoljen dodaju se listi ciljeva za koje se traže vrednosti a dokazani ciljevi se skidaju sa ove liste [17]. Proces se ciklično ponavlja dok lista ciljeva ne bude iscrpljena. Ova strategija ima sličnosti sa metodom neinformisanog pretraživanja po dubini. Pogodna je za rešavanje problema sa puno raspoloživih podataka a malo prihvatljivih rešenja, kao što je dijagnostika.

#### 2.2.4. Strukturalna reprezentacija znanja

Strukturalna reprezentacija znanja pokazuje kako su povezane činjenice sa domenom konstruisanjem takozvanih **ontologija**. To su formalne specifikacije (konceptualizacije) određenog domena u cilju rešavanja različitih problema iz istog domena [18]. Ovakve reprezentacije omogućavaju prevođenje deklarativnog znanja (“znati da”) u proceduralno znanje (“znati kako”). Ona olakšava primenu proceduralnog znanja i pruža konceptualnu osnovu za

uvođenje proceduralnog znanja višeg reda (“znati zašto”). Problemi se rešavaju uz pomoć opštih metoda rezonovanja, najčešće korišćenjem **deskriptivne logike** [17]. Šema rezonovanja i domensko znanje moraju biti nezavisni da bi automatsko rezonovanje bilo moguće. Osnovne karakteristike strukturalnog modela su:

- **Reprezentativnost** - sposobnost reprezentovanja svih potrebnih tipova informacija u okviru domena
- **Inferijalna adekvatnost** – sposobnost tumačenja i primene informacija koje dolaze sa novim informacijama
- **Inferijalna efikasnost** – sposobnost pripajanja dodatnih informacija u strukturu znanja
- **Akviziciona efikasnost** – lako prikupljanje informacija

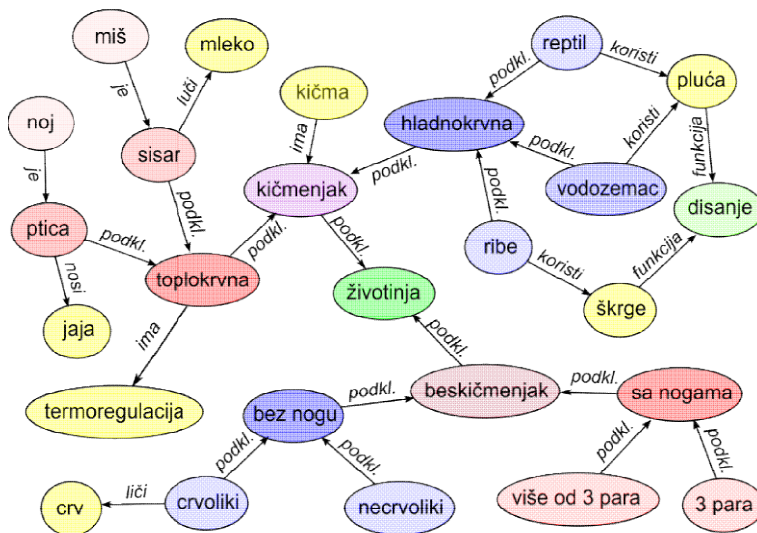
Prve korake na polju strukturalne reprezentacije znanja učinio je Schank (Carl Schank, rođen 1946) objavljivanjem svoje teorije konceptualne zavisnosti 1969. god. Ova teorija obezbedila je pogodan matematički aparat za automatsku semantičku analizu. Uprkos činjenici da je razvijena za analizu prirodnih jezika metoda je doživela primenu u ostalim poljima veštačke inteligencije.

#### 2.2.4.1. Semantičke mreže

Semantičke mreže, kao ontologije, osim koncepata sadrže instance (primere) ovih koncepata. U ovakvom slučaju koncept se naziva klasom a njegove instance **objektima** koji pripadaju toj klasi. Novi koncept definiše se pomoću jednostavnijih **elementarnih** koncepata koji se nazivaju **atomički koncepti**. Prema principu **nasleđivanja**, koji je karakteristika ontologija, izvedeni koncept nasleđuje karakteristike starog koncepta, iz koga je izveden. Primer semantičke mreže predstavljen je na slici 2.6. Koncepti odnosno objekti predstavljeni su imenovanim čvorovima (elipse) a relacije između njih imenovanim vezama (linije). Veze se, kod semantičkih mreža, drugačije nazivaju **ulogama**. Slično kao kod koncepata, složene uloge mogu se definisati na osnovu jednostavnijih elementarnih uloga pa i na osnovu koncepata. Kod semantičkih mreža istaknute su dve uloge:

- *je* – označava pripadnost klasi
- *je podklasa* – označava pripadnost podklasi

Prve semantičke mreže predložili su Allan M. Collins i Ross Quillian, 1969. god. kao rezultat istraživanja o razumevanju prirodnog jezika. Međutim, vrlo brzo su našle široku primenu, naročito u prirodnim naukama za formiranje taksonomija koje imaju za cilj sistematizaciju znanja.



Slika 2.6 Semantička mreža sa hijerarhijskom strukturom

#### 2.2.4.2. Okviri

Okviri predstavljaju suštinsko proširenje semantičkih mreža. Kod ovog koncepta strukturalnog predstavljanja znanja čvor ima kompleksnu unutrašnju strukturu koja omogućava njegovo detaljno opisivanje. Ovakav čvor nazvan je okvir odnosno **okvir klase** ako predstavlja klasu ili **okvir objekta** u slučaju da predstavlja objekt. Struktura okvira definisana je po strogim principima koji omogućavaju automatsku obradu i analizu podataka. Oni se sastoje od proreza (engl. Slot) koji služe za opisivanje osobina koncepta (ili objekta). Dominantne karakteristike okvira su:

- **Vrednost** – trenutna vrednost proreza
- **Opseg** – merni opseg ili lista vrednosti
- **Default** – vrednost proreza ako prva od osobina (Vrednost) nije dostupna
- **Pripada klasi** – odgovara ulozu “je” kod semantičkih mreža
- **Pripadnost podklasi** – odgovara ulozu “je podklasa” kod semantičkih mreža



Osnovni mehanizam rasuđivanja kod okvira je nasleđivanje. On omogućava da izmenom osobine klase utičemo na sve objekte i podklase koje su iz nje izvedene [19]. Na ovaj način obezbeđena je efikasna distribucija znanja. Kod okvira, mehanizam je dodatno unapređen mogućnošću da jedan objekt (ili klasa) bude izveden iz više klasa. Ovakva mogućnost naziva se višestruko nasleđivanje (engl. multiple inheritance). Osim nasleđivanja, kao mehanizam rasuđivanja primenjuju se takozvane **demonske** procedure. Ove leže, zajedno sa osobinama objekta (ili klase) u prorezima nekog okvira i čekaju da budu pozvane (okinite) odgovarajućim događajem. Spektar događaja može biti najrazličitiji a dominantni slučajevi su:

- **Potreba** (engl. if-needed) – pri čitanju vrednosti ako je ona nepoznata
- **Dodavanje** (engl. if-added) – pri dodavanju nove vrednosti
- **Podešavanje** (engl. if-updated) – pri promeni vrednosti
- **Uklanjanje** (engl. if-removed) – pri uklanjanju vrednosti
- **Čitanje** (engl. if-read) – pri čitanju vrednosti
- **Novi** (engl. if-new) – pri dodavanju novog okvira

Nedostatak okvira je odsustvo eksplicitnih informacija o kontroli procesa. Zbog toga oni zahtevaju kompleksno programsko okruženje kako bi obezbedili zadovoljavajuće rezultate. Pri tome uvek pretil opasnost da kompleksno okruženje učini obradu informacija uz pomoć okvira neprihvatljivo sporim.

#### 2.2.4.3. Skripte

Skripte su strukturalne reprezentacije koje opisuju događaj određenog tipa na osnovu stereotipa pri čemu uzimaju u obzir i kontekst. Formalno se mogu definisati na osnovu sledećih pojmova:

- **Agenti** (engl. Agents) su objekti koji utiču na druge i primaju uticaje drugih objekata
- **Rekviziti** (engl. Props) su stvari koje se pojavljuju
- **Akcije** (engl. Actions) su elementarni događaji koji se upotrebljavaju za konstrukciju čitavog događaja
- **Propozicije** (engl. Preconditions) su preduslovi koji važe pre početka rasuđivanja
- **Rezultati** su uslovi koji važe na kraju rasuđivanja

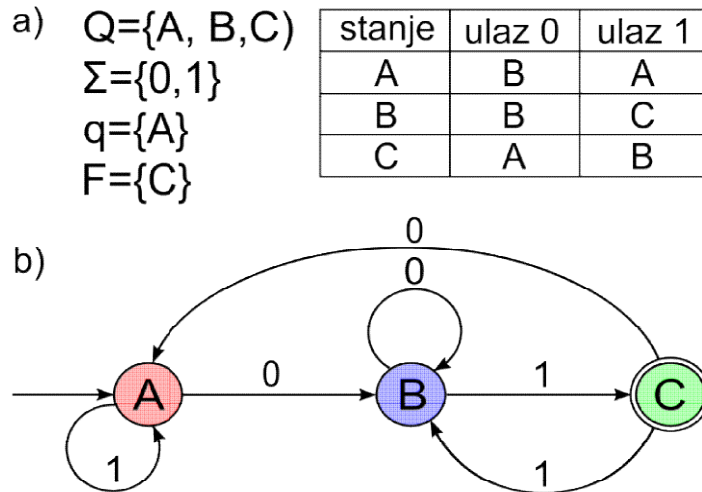
Kod skripti se za analizu nekog događaja koristi opšti obrazac o takvom tipu događaja. Opšti obrazac je rezultat ranijih iskustava o sličnim događajima, odnosno događajima istog tipa. Na osnovu ranijih iskustava moguće je predvideti tok ili ishod nekog događaja. Zbog toga je mehanizam skripti idealan za donošenje odluka u cilju ostvarenja zadatog cilja.

### 2.2.5. Princip matematičke lingvistike

Ovaj pristup teži predstavljanju znanja putem jezika simulirajući način na koji to čovek čini koristeći se jezikom. Osnovni lingvistički zadatak koji je postavljen pred veštačkom inteligencijom bio je provera sintakse prirodnog jezika. S obzirom da je prirodni jezik zapravo neograničeni skup rečenica, koje su konstruisane na osnovu odgovarajućih gramatičkih pravila, ovaj zadatak nije jednostavno ispuniti. Prvo rešenje pružio je američki naučnik Noam Chomsky definisanjem **generativnih gramatika** koje su namenjene sintaktičkoj analizi prirodnih jezika. Umesto eksplicitne reprezentacije, on je predvideo predstavljanje jezika u vidu pravila koje je nazvao **produkcijama**. Primenom produkcija moguće je generisati sve dozvoljene kombinacije reči odnosno sve smislene rečenice odgovarajućeg jezika. Za generisanja rečenica primenjeni su **formalni automati** odnosno apstraktna računaska sredstva koja automatski slede zadatu sekvencu operacija [20]. Zadatak formalnog automata bio je da otkrije da li zadati izraz pripada nekom jeziku odnosno da li je on u skladu sa njegovim pravilima. Formalni automati su omogućili kontrolisanje neograničenog jezika koristeći se konačnim brojem produkcija. Formalni automat, sa konačnim brojem stanja, odnosno **konačni automat** ili **mašina sa konačnim brojem stanja** je definisan sledećom uređenom petorkom:

- $Q$  – konačan skup stanja
- $\Sigma$  – konačan skupa znakova (alfabet)
- $\delta$  – funkcija prelaza
- $q_0$  - početno stanje ( $q_0 \in Q$ )
- $F$  – skup konačnih stanja  $F \subseteq Q$

Nasuprot njemu, automat sa beskonačnim brojem stanja naziva se **beskonačni automat** ili **mašina sa beskonačnim brojem stanja**. Slika 2.7 a daje primer definicije jednog formalnog automata gde je funkcija prelaza ( $\delta$ ) definisana tabelom [21]. Alternativni metod reprezentacije pomoću signalnog grafa dat je na slici 2.7b.



Slika 2.7 Reprezentacija formalnog automata: a) simbolički, b) grafički

Pristup matematičke lingvistike pokazao je skromne rezultate u slučaju prirodnih jezika ali se model pokazao vrlo korisnim pri analizi **formalizovanih jezika**. U cilju povećanja deskriptivnog kapaciteta generativna gramatika je doživljavala modifikacije kao što su dodavanje atributa komponentama jezika odnosno atributizovane (engl. attributed) gramatike i višedimenzionalne generativne (engl. multi-dimensional) gramatike.

### 2.3. Računska inteligencija

Računska inteligencija se može definisati kao “proučavanje adaptivnih mehanizama koji omogućavaju ili olakšavaju inteligentno ponašanje u kompleksnom i promenljivom okruženju.” [22] Predstavlja komplementarni pristup veštačkoj inteligenciji u odnosu na simboličku veštačku inteligenciju [17]. Metode računske inteligencije se odlikuju sledećim opštim karakteristikama:

- Model inteligentnog sistema je definisan najčešće implicitno
- Znanje je predstavljeno numerički
- Obrada znanja bazirana je na numeričkim operacijama

Međutim, pojedine metode računske inteligencije, u manjoj ili većoj meri, odstupaju od navedenih pravila, zbog čega u naučnim krugovima postoje oprečna mišljenja o tome koji algoritmi pripadaju računskoj inteligenciji a koje ne. Među metodama veštačke inteligencije

moгу se izdvojiti tri opšte kategorije: konekcionišćki modeli, modeli bazirani na matematišćkim teorijama i modeli bazirani na biološkim zakonima.

### 2.3.1. Konekcionišćki modeli

Prema konekcionišćkom pristupu, u mehanizmu rasuđivanja leži povezivanje mentalnih stanja koje se obavlja na osnovu prethodnog iskustva odnosno na osnovu stimulansa. Veza između stanja jaća ako se stimulans često ponavlja odnosno slabi u slućaju da izostaje. Kod raćunske inteligencije slićan mehanizam postignut je uz pomoć konekcionišćkih mreža. Kod ovih mreža znanje je distribuirano (raspodeljeno), na implicitan naćin, među elementarnim jedinicama, taćnije među parametrima (pojaćanjima) veza između gradivnih jedinica. Stanje se definiše kao **nastajući proces** u mrežama elementarnih jedinica pri ćemu se pod nastajućim procesom podrazumeva bilo koji proces koji se ne mođe objasniti na osnovu elementarnih potprocesa. Tipićan predstavnik konekcionišćkih modela su vešćaćke neuronske mreže. U predmetom istraživlju percepcije zvuka, ove mreže su izabrane kao alat za obradu akustićnih signala. Zbog toga je neuronskim mrežama posvećeno posebno poglavlje ove disertacije (Poglavljje 3).

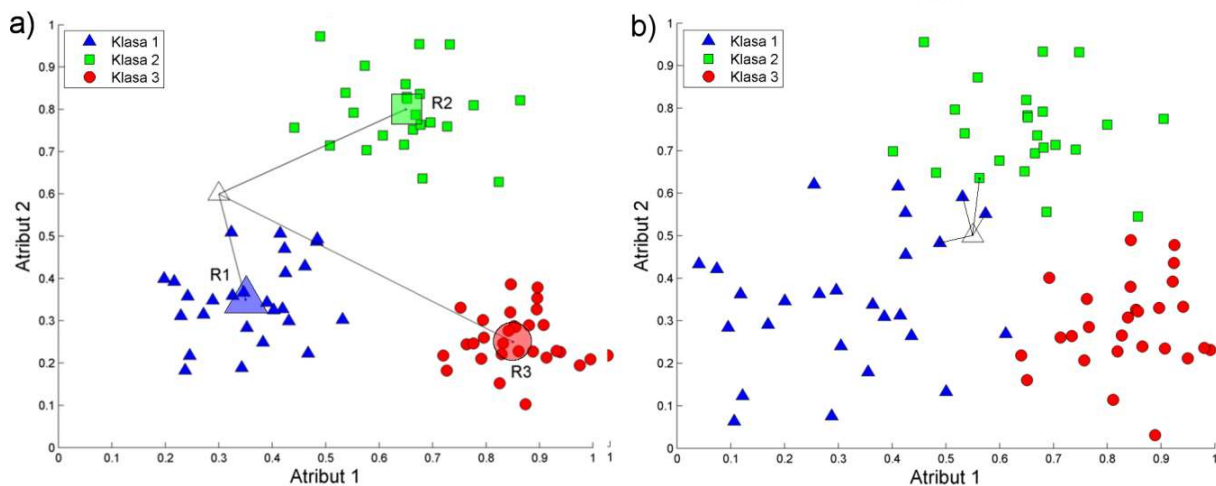
### 2.3.2. Modeli bazirani na matematišćkim teorijama

Modele bazirane na matematišćkim teorijama mođeemo podeliti u dve fundamentalne grupe. Prvu grupu metoda ćine prepoznavanje obrazaca i analiza klastera. Ove dve srodne tehnike razlikuju se u tome što prva pripada **nadgledanom** a druga **nenadgledanom** mašinskom ućenju. Nagledano mašinsko ućenje obavlja se podešavljanjem parametara algoritma na osnovu podataka za treniranje odnosno skupa ulazno izlaznih velićina. Nenadgledano mašinsko ućenje predstavlja modeliranje skrivene strukture ili distribucije podataka u cilju izvođenja zakljućaka o njima. Druga grupa metoda tiće se notacije nesavršenog znanja i rezonovanja na osnovu nesavršenog znanja. Nesavršenost znanja mođe poticati od razlićitih uzroka i u odnosu na njih mođe biti:

- **Neizvesno** – kao posledica neizvesnosti podataka
- **Neprecizno** – kao rezultat nepreciznosti mernih instrumenata
- **Nekompletno** – kada nisu poznate sve ćinjenice

### 2.3.2.1. Prepoznavanje obrazaca

Prepoznavanje obrazaca je jedan od bazičnih zadataka koji se postavljaju pred modelima veštačke inteligencije [17]. Sastoji se u svrstavanju objekta u neku od unapred definisanih kategorija na osnovu vrednosti njegovih atributa (osobina). Osim fizičkog objekta obrazac može biti bilo koji fenomen, struktura, proces itd. Način predstavljanja obrasca može biti pomoću **vektora atributa** ili strukturalno (Poglavlje 2.2.1.4.). Prostor koji sadrži vektore atributa naziva se **prostor atributa**. Ovakav prostor, sa dve dimenzije i tri klase obrazaca u njemu, predstavljen je na slici 2.8. Ako su obrasci odvojeni u klase očigledno, kao na slici 2.8 a, kaže se da formiraju **regularne klasterne**.



*Slika 2.8 Dvodimenzionalni prostor atributa sa tri klase obrazaca: a) regularni klasteri–metod najmanjeg rastojanja, b) iregularni klasteri–metod najbližeg suseda*

Klasifikator se konstruiše na osnovu podataka za treniranje (učenje) odnosno skupa obrazaca (uzoraka) koji su pravilno klasifikovani. Da bi bio klasifikovan u neku od kategorija nepoznati obrazac ne mora imati potpuno iste karakteristike sa nekim od obrazaca te kategorije. Dovoljno je da njegove osobine budu slične odnosno približno iste sa osobinama obrazaca te kategorije koji pripadaju skupu obrazaca za treniranje. Prepoznavanje obrazaca praktično se vrši uz pomoć nekog od naprednih matematičkih modela kao što su Bayes-ov model verovatnoće, diskriminantne funkcije i modeli najmanjeg rastojanja.

Postupku prepoznavanja obrazaca najčešće prethode faze predobrada signala, ekstrakcije atributa i selekcije atributa. **Faza predobrade** signala uključuje **uklanjanje šuma** (poremećaja), uz pomoć različitih metoda filtriranja, kao i normalizaciju atributa kako bi imali uporedive opsege. **Ekstrakcija atributa** podrazumeva smanjenje njihovog broja kombinovanjem i transformisanjem starih u nove attribute. Razlog je smanjenje broja matematičkih operacija, samim tim smanjenje potrebnog vremena i greške klasifikacije [17]. Iz sličnog razloga na kraju se vrši selekcija atributa sa najvećom **diskriminativnom moći** odnosno onih čija upotreba dovodi do najmanjih grešaka u procesu klasifikacije.

#### 2.3.2.1.1. Klasifikator najbližeg suseda

Ovaj klasifikator inspirisan je kognitivnim mehanizmom prepoznavanja kod ljudi i životinja po kome se nepoznati obrazac svrstava u kategoriju obrazaca u kojoj je pronađen najbliži njemu. Sličnost među obrascima ocenjuje se na osnovu međusobnog rastojanja u višedimenzionalnom prostoru atributa a rastojanje se izračunava primenom neke od vektorskih normi, najčešće Euklidovom.

Klasifikator najbližeg suseda pogodan je za slučajeve kada kategorije objekata nisu odvojene u regularne klastere već postoji njihovo preklapanje i rasipanje (slika 2.8 b). Osnovni nedostatak ove metode je neophodnost upoređivanja nepoznatog obrasca sa svim elementima iz skupa obrazaca za treniranje, što zahteva veliki broj matematičkih operacija [23]. Pored toga, ova metoda ispoljava veliku osetljivost na greške. Dovoljno je da najbliži sused sadrži greške u atributima da rezultat klasifikacije bude pogrešan.

#### 2.3.2.1.2. Klasifikator k-najbližih suseda

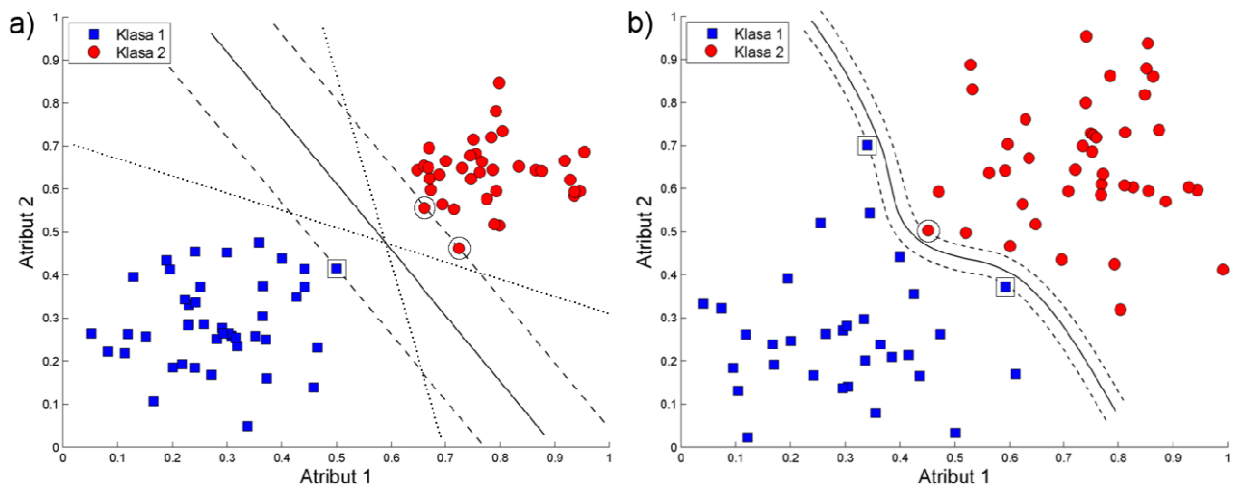
Jedan od načina da se umanju mogućnost greške klasifikatora prvog suseda, je da se umesto jednog identifikuju nekoliko najbližih suseda. Ovakav pristup poznat je pod nazivom metod k-najbližih suseda. Nepoznati obrazac se dodeljuje onoj klasi obrazaca koja dominira, odnosno ima najviše predstavnika, među  $k$  identifikovanih suseda [24]. Za  $k$  se najčešće bira neparan broj kako bi se eliminisala mogućnost da dve kategorije imaju jednak broj predstavnika među  $k$  najbližih suseda.

### 2.3.2.1.3. Klasifikator minimalnog rastojanja

Klasifikator minimalnog rastojanja nastao je iz potrebe da se izbegne upoređivanje nepoznatog obrasca sa svim elementima iz skupa za treniranje. Umesto toga, za svaku od klasa konstruiše se po jedan reprezentativni obrazac (slika 2.8 a). Pod uslovom da su obrasci grupisani u regularne klasterne, atributi reprezentativnog obrazca klase dobijaju se uprosečavanjem atributa pripadajućih elemenata. Nepoznati obrazac svrstava se u onu kategoriju obrazaca sa čijim reprezentativnim obrascem ima najviše sličnosti. Sličnost između obrazaca ocenjuje na isti način kao kod prethodnih metoda. Primenom reprezentativnih obrazaca drastično je smanjen potreban broj matematičkih operacija odnosno povećana efikasnost klasifikacije.

### 2.3.2.1.4. Klasifikatori bazirani na granici odluke

Metoda se bazira na konstruisanju granica u prostoru atributa koje odvajaju elemente različitih klasa. Ako obrasci formiraju regularne klasterne granica se može definisati uz pomoć linearne funkcije. Ovakve klase nazivaju se linearno odvojive a granica između njih ima oblik hiperravnini u višedimenzionalnom prostoru atributa. Slučaj linearne diskriminativne funkcije u dvodimenzionalnom prostoru atributa prikazan je punom linijom na slici 2.9 a. Ako se nepoznati obrazac nalazi ispod diskriminativne linije pripada klasi 1 (kvadrati) dok u protivnom pripada klasi 2 (krugovi). U opštem slučaju, kada postoje više od dve klase obrazaca prostor atributa deli se pomoću više granica pri čemu svaka od njih međusobno odvaja par klastera obrazaca.



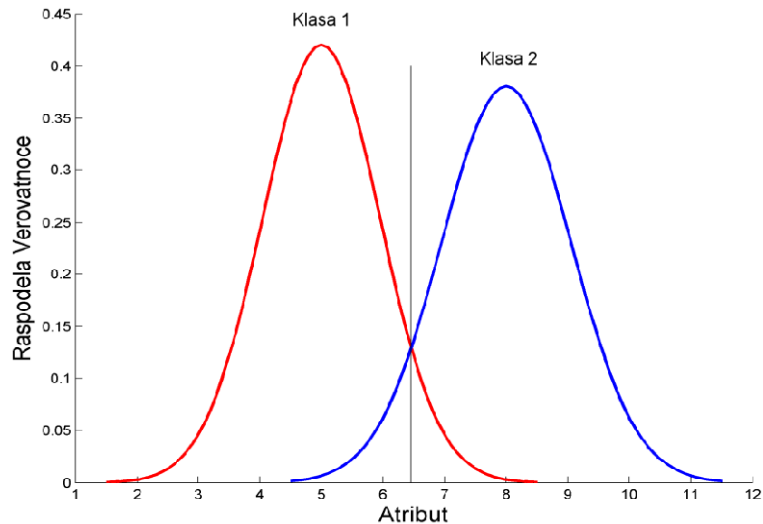
Slika 2.9 Granice odluke: a) linearna, b) nelinearna

Kada elemente različitih klasa obrazaca nije moguće odvojiti linearnom funkcijom definiše se nelinearni klasifikator. Za ovu namenu primenjuje se neka od nelinearnih diskriminativnih funkcija. Nelinearna diskriminativna funkcija, za slučaj dvodimenzionalnog prostora atributa, predstavljena je punom linijom na slici 2.9 b. Konstrukcija nelinearnih klasifikatora je daleko kompleksnija u odnosu na konstrukciju linearnih što ograničava njihovu praktičnu primenu. Alternativni pristup može biti primena **splajnova** odnosno većeg broja funkcija nižeg stepena koje definišu diskriminativnu granicu deo po deo.

#### 2.3.2.1.5. Statističko prepoznavanje obrazaca

Statističko prepoznavanja obrazaca bazirano je na Bayes-ovom statističkom modelu, odnosno na stepenu verovanja u ishod nekog događaja. Rezultat prepoznavanja je verovatnoća dodeljivanja nepoznatog obrasca određenoj klasi obrazaca [25]. Slika 2.10 prikazuje raspodelu verovatnoće za dve klase sa normalnom distribucijom i različitim prethodnim raspodelama verovatnoće u jednodimenzionalnom prostoru atributa. Raspodela verovatnoće klase obrazaca je funkcija koja za zadatu vrednost atributa daje verovatnoću njenog pojavljivanja. U primeru sa slike većina obrazaca klase 1 ima vrednost atributa između 2 i 8 dok većina obrazaca klase 2 ima vrednost atributa između 5 i 11. Najveću verovatnoću pojavljivanja u klasi 1 ima vrednost atributa 5 a u klasi 2 vrednost atributa 8. Grafici raspodele verovatnoće seku se za vrednost atributa oko 6,5. Ova granična vrednost atributa je na slici označena vertikalnom linijom. Za manje vrednosti atributa veća je verovatnoća da obrazac pripada klasi 1 odnosno manja verovatnoća da pripada klasi 2 i obrnuto. Ne postoje principijelne razlike u određivanju verovatnoće ishoda kada postoje više od dve klase obrazaca. Bayes-ov klasifikator dodeljuje nepoznati obrazac onoj klasi obrazaca koja za zadatu vrednost atributa ima najveću raspodelu verovatnoće. U opštem slučaju, kada postoje više od jednog atributa, primenjuje se naivni Bayes-ov klasifikator. Pod pretpostavkom da su atributi međusobno nezavisni, izračunava se verovatnoća javljanja obrasca u odnosu na svaki atribut posebno. Raspodela verovatnoće višedimenzionalnog vektora atributa dobija se kao proizvod elementarnih verovatnoća koje odgovaraju pojedinačnim atributima. Proces se može dalje uopštiti uvođenjem funkcije koštanja greške, koja služi da se napravi razlika u značaju među atributima.





*Slika 2.10 Bayes-ova granica za klase obrazaca sa razlicitom prethodnom distribucijom*

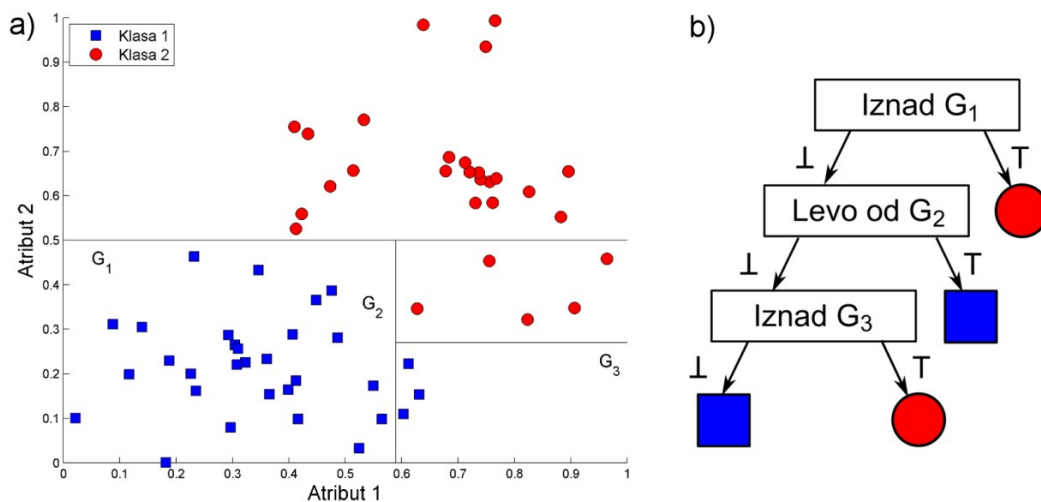
#### 2.3.2.1.6. Mašina potpornih vektora

Linearna diskriminativna funkcija u opštem slučaju nije jednoznačno određena. Na slici 2.9a su, osim neprekidne, date dve tačkaste linije koje potpuno odvajaju klase obrazaca. Očigledno je da ove linije predstavljaju lošije rešenje u odnosu na punu liniju jer se neki od obrazaca nalaze njenoj neposrednoj blizini. Neznatno pomeranje ovih obrazaca učinilo bi diskriminativnu granicu nevalidnom. Puna linija je dovoljno udaljena od elemenata oba klastera što je čini manje osetljivom na greške odnosno na pomeranje obrazaca u prostoru atributa. U cilju nalaženja optimalne diskriminativne granice, koja je maksimalno udeljena od elemenata oba klastera, definisana je mašina potpornih vektora. Postupak se sastoji u nalaženju dve međusobno paralelne hiper površine koje su podjednako udaljene od diskriminativne granice. Na slici 2.9 a i b ove hiper površine predstavljene su isprekidanim linijama. Svaka od hiperpovršina sadrži elemente jednog klastera koji su najviše istureni (štrče) ka drugom klasteru, na slici okruženi linijom. Diskriminativna granica je paralelna i podjednako udaljena od obe hiperpovršine [26].

#### 2.3.2.1.7. Stablo odlučivanja

Za razliku od prethodnih vrsta klasifikatora, kod kojih se klasifikacija nad svim klasama i atributima obavlja u jednom koraku, ovaj metod predviđa klasifikaciju u više faza. U svakoj od faza definiše se po jedna granica odlučivanja. Granice moraju biti paralelne sa nekom od osa koordinatnog sistema kojim je definisan prostor atributa. Svakoј granici odlučivanja odgovara

jedan korak u stablu odlučivanja. Na slici 2.11 predstavljen je dvodimenzionalni prostor atributa sa dve klase obrazaca. U svrhu podele prostora atributa, upotrebljene su 3 granice odlučivanja kojima odgovaraju 3 koraka odnosno uslova u stablu odlučivanja. Prvi korak pokazuje da svaki obrazac iznad granice  $G_1$  pripada klasi 2 odnosno klasi crvenih kružića. Ako uslov nije ispunjen prelazi se na sledeći korak u stablu odlučivanja kome odgovara sledeća granica koja deli prostor atributa. Proces se ponavlja dok u prostoru atributa ne budu iscrpljene sve zone koje sadrže više od jedne klase obrazaca.



Slika 2.11 Sukcesivno konstruisanje stabla odlučivanja i podela prostora atributa

### 2.3.2.2. Analiza klastera

Analiza klastera je komplementarni problem veštačke inteligencije u odnosu na problem prepoznavanja obrazaca [17]. Ovde ne postoji skup podataka za treniranje sa unapred poznatom klasifikacijom. Zadatak se sastoji u grupisanju skupa nepoznatih obrazaca u klastere, koji predstavljaju klase, na osnovu principa sličnosti. Sličnost se ocenjuje na isti način kao kod prepoznavanja obrazaca, na osnovu rastojanja između njihovih stanja u višedimenzionalnom prostoru atributa. Pošto se rastojanje može ocenjivati na različite načine izbor odgovarajuće metrike ima ključan uticaj na dobijeni rezultat. Problem analize klastera se može podeliti u dve opšte kategorije:

- **Metodi bazirani na podeli** kod kojih je unapred definisan broj klastera
- **Hijerarhijski modeli** kod kojih broj klastera nije unapred poznat

#### 2.3.2.2.1. Klasterizacija metodom k-srednjih vrednosti

U različitim varijacijama, ovo je jedna od najpopularnijih metoda analize klastera baziranih na podeli [27]. Podela obrazaca vrši se primenom centroida klastera. Centroid predstavlja srednju vrednost pozicija svih elemenata jednog klastera. Pod pretpostavkom da je skup obrazaca potrebno grupisati u  $k$  klastera, postupak se obavlja u sledećim koracima

1. Selektovati  $k$  inicijalnih centroida, slučajnim izborom  $k$  obrazaca ili izborom  $k$  proizvoljnih tačaka unutar prostora atributa
2. Pridružiti svaki obrazac nekom od inicijalnih centroida na osnovu principa sličnosti
3. Za klastere kreirane u prethodnom koraku konstruisati nove centroide
4. Ponavljati korake 2 i 3 dok se pozicija centroida ne stabilizuje

Smatra se da su klasteri stabilizovani kada se raspored obrazaca ne menja između uzastopnih iteracija ili su promene manje od dozvoljenog praga odnosno kada se pozicija centroida ne menja ili menja manje od dozvoljenog praga.

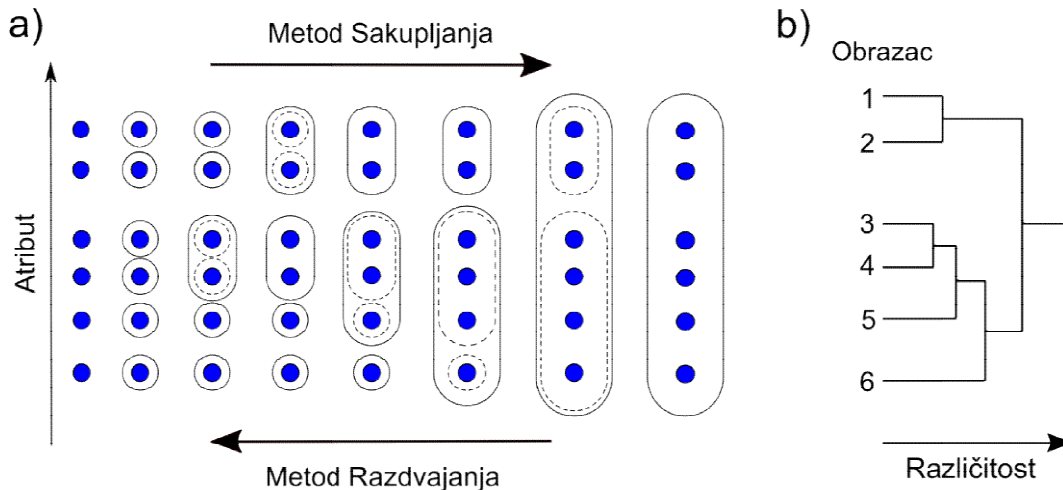
#### 2.3.2.2.2. Hijerarhijska analiza klastera

Postoje dva pristupa hijerarhijskoj analizi klastera. Jedna je spajanje postojećih klastera u veće klastera a drugi je razdvajanje većih klastera na manje. Pod pretpostavkom da se skup uzoraka sastoji od  $M$  obrazaca proces grupisanja obrazaca u klastera na osnovu principa spajanja izvodi se u sledećim koracima:

1. Definisati  $M$  inicijalnih klastera koji sadrže samo po jedan obrazac i izračunati njihova međusobna rastojanja u višedimenzionalnom prostoru atributa
2. Izabrati najbliži odnosno najbliži par klastera i spojiti ih u jedan
3. Izračunati rastojanja između novoformiranih i ostatka klastera
4. Ponavljati korake 2 i 3 do konačnog kreiranja klastera koji sadrži svih  $M$  uzoraka obrazaca

Postoje različite metode određivanja rastojanja između klastera, kada klaster sadrži više od jednog elementa. Najčešće se upotrebljavaju sledeće:

- **Metod jednostrukog povezivanja** – rastojanje između dva klastera izračunava se kao rastojanje između dva najbliža elementa
- **Metod kompletnog povezivanja** – rastojanja između dva klastera izračunava se kao rastojanje između dva najudaljenija obrasca
- **Metod centroida** – rastojanje između dva klastera izračunava se kao rastojanje između njihovih centroida

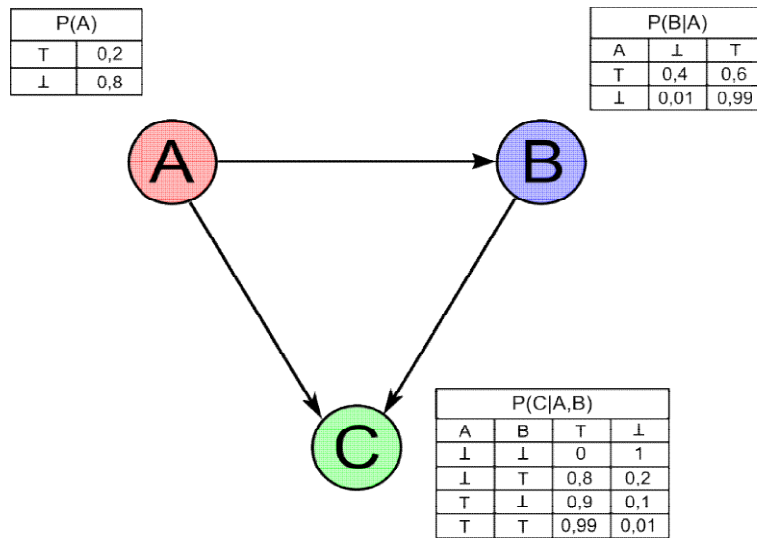


Slika 2.12 Hijerarhijska analiza klastera

Koraci hijerarhijske analize klastera metodom jednostrukog povezivanja u jednodimenzionalnom prostoru atributa prikazani su na slici 2.12 a. Sa leva na desno analiza klastera izvršena je po principu spajanja odnosno sa desna na levo po principu razdvajanja. Slika 2.12 b daje prikaz analize klastera, karakterističnu za hijerarhijsku analizu, koja se naziva **dendogram**. Na vertikalnoj osi nanose se obrasci a na horizontalnoj mera različitosti među njima [28].

### 2.3.2.3. Bayes-ove mreže

Bayes-ove mreže su fundamentalni metod veštačke inteligencije koji operiše sa teorijom verovatnoće. Pripada kategoriji probabilističkih grafičkih metoda. U matematičkom smislu Bayes-ova mreža predstavlja **direktni aciklični graf** koji reprezentuje združenu raspodelu verovatnoće nad skupom slučajnih promenljivih. Primer Bayes-ove mreže sa tri slučajne promenljive dat je na slici 2.13.



Slika 2.13 Bayes-ova mreža

U strukturi mreže razlikuje se kvantitativni i kvalitativni deo. Kvalitativni deo čine čvorovi obeleženi krugovima, koji predstavljaju slučajne promenljive, zajedno sa vezama među njima, koje su predstavljene strelicama. Strelica polazi iz čvora koji vrši uticaj, nazvan roditeljem, i završava se u čvoru koji trpi uticaj, nazvan detetom. Produženjem ove terminologije na sve prethodne čvorove dobijaju se preci a na sve naredne čvorove dobijaju se potomci. Čvor koji nema roditelje naziva se koren a čvor koji nema dece naziva se list itd. Kvantitativni parametri definisani su u skladu sa osobinom Markova što znači da raspodela uslovne verovatnoće svakog čvora zavisi isključivo od njegovih roditelja. Za diskretne slučajne promenljive uslovna verovatnoća se predstavlja u vidu tabele koje daju listing svih ostvarljivih vrednosti čvora za svaku od kombinacija vrednosti njegovih roditelja. Združena raspodela kolekcije promenljivih može biti određena jednoznačno na osnovu lokalnih tabela uslovnih verovatnoća pojedinih čvorova. Na raspolaganju su dva tipa rasuđivanja [29]

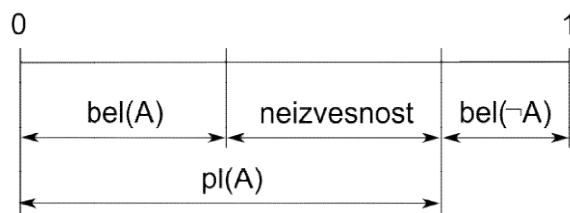
- **Prediktivno rasuđivanje** (odozgo na dole) - na osnovu činjenica dobijenih od roditelja
- **Dijagnostičko rasuđivanje** (odozdo na gore) - na osnovu činjenica dobijenih od dece

Kod većine praktičnih primena Bayes-ova mreža nije unapred poznata. Ona treba da bude konstruisana kroz postupak nazvan treniranjem ili učenjem Bayes-ove mreže. Treniranje

podrazumeva upotrebu raspoloživih podataka i prethodnih informacija (ekspertske znanje i neformalni odnosi) za određivanje topologije grafa i parametra združene raspodele verovatnoće. Određivanje strukture smatra se kompleksnijim delom problema u odnosu na određivanje parametara.

#### 2.3.2.4. Teorija Dempster-Shafer-a

Teoriju rasuđivanja, namenjenu radu sa nekompletnim znanjem, poznatu još kao **teorija funkcije verovanja** ili **matematička teorija o dokazima**, razvio je Arthur P. Dempster a kasnije doradio Glenn Shafer, obojica američki matematičari. U nedostatku kompletnih činjenica o nekom događaju logično je pretpostaviti da rezultat zaključivanja bude manje kvalitetan nego kada su sve činjenice dostupne. Da bi se nekompletnost saznanja izrazila kvantitativno uveden je nov koncept verovatnoće i u skladu sa njim nov model rasuđivanja. Za razliku od Bayes-ovog pristupa, koji računa verovatnoću tačnosti neke pretpostavke (hipoteze), novi koncept daje verovatnoću da raspoložive informacije podržavaju verovanje da je određena hipoteza tačna. Drugim rečima, potrebno je poznavati razliku između neizvesnosti znanja koje posedujemo i našega neznanja, gde se neznanje definiše kao svesnost o ograničenosti sopstvenog znanja [17]. Ove veličine su formalno uvedene u vidu koncepata **donje** i **gornje verovatnoće**, slika 2.14. Donja verovatnoća ili **vera** (engl. belief), na slici označena sa  $bel(A)$ , predstavlja meru snage dokaza u korist hipoteze  $A$ . Gornja verovatnoća ili **verodostojnost** (engl. plausibility), na slici označena sa  $pl(A)$ , predstavlja meru u kojoj dokazi koji obaraju hipotezu ostavljaju mogućnost za nju [30]. Obe veličine mogu imati vrednost između 0 i 1 a vrednost između njih naziva se neizvesnost. Osnovna prednost modela je promena vrednosti navedenih veličina sa sakupljanjem novog znanja koji treba da odrazi redukciju neznanja o događaju. Ova redukcija se odražava kvantitativno kroz redukciju intervala neizvesnosti [31].



Slika 2.14 Verovanje i verodostojnost

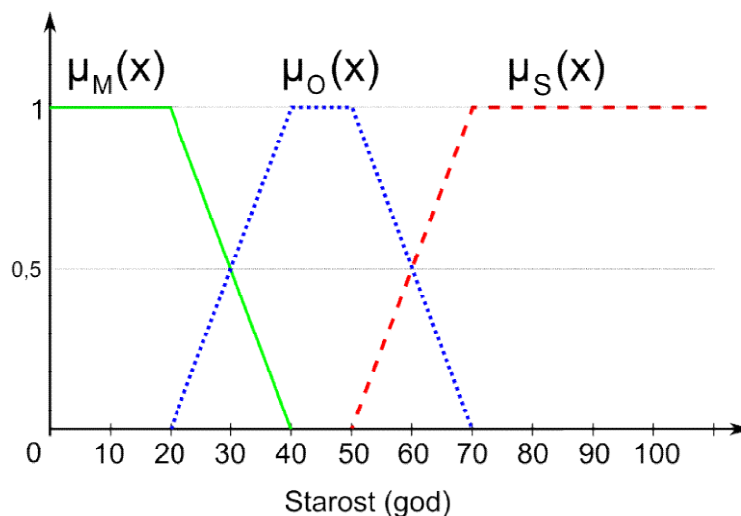
### 2.3.2.5. Nemonotono rasuđivanje

Klasična logika je **monotona**, što znači da dodavanje novog dokaza ne može uticati na smanjenje skupa mogućih posledica. Međutim, prilikom modeliranja realnog sveta retko su na raspolaganju apsolutno tačne činjenice. U takvim situacijama pretpostavke se prave na bazi **normalnih** odnosno **očekivanih** rezultata, primer: *sisar ne leti osim ako nije slepi miš*. Ovakve pretpostavke su tačne u najvećem broju slučajeva ali u izuzetnim situacijama se mogu ispostaviti kao neočekivane odnosno **abnormalne**. Mehanizam rasuđivanja treba da obezbedi prepoznavanje abnormalne situacije i u skladu sa njom odgovarajuću reviziju ranije donesenih zaključaka. Ovakav mehanizam naziva se **nemonotono rasuđivanje**. Osnovna karakteristika nemonotonog rasuđivanja je promena logike kao rezultat novih saznanja. Jedan od formalnih načina za predstavljanje promene u logici je **situacioni račun** [32]. Ovaj račun predstavlja promenu u vidu situacija koje nastaju kao posledica niza akcija, primer: *miš je zagadio hranu ako je izgrickao pakovanje*. Efekat akcije na situaciju je relativno lako izračunati promenom situacionog računa ali je takođe potrebno odrediti šta se nije promenilo. Da se ne bi pristupilo taksativnom navođenju velikog broja nepromenjenih činjenica uveden je koncept **okvira**. Okviri su pravila koja, prilikom razmatranja efekata neke akcije, reprezentuju veliku količinu nepromenjenih činjenica, primer: *hrana je ostala zdrava osim ako miš nije izgrickao pakovanje*. Kod nemonotonog rasuđivanja dominiraju dva pristupa: logika fiksne tačke i logika poželjnih modela. **Logika fiksne tačke** teži generisanju maksimalnog skupa prihvatljivih verovanja, koje se nazivaju ekstenzije ili ekspanzije, na osnovu datih uslova [33]. Nemonotonost ovog pristupa proizilazi iz skupa verovanja koji se konstantno menja sa novim saznanjem. **Logika poželjnih modela** bazira se na nemonotonosti rasuđivanja. Među modelima bira se skup **reprezentativnih** (poželjnih) modela i u odnosu na njih definiše nemonotono rasuđivanje [34]. Nemonotonost se ostvaruje ažuriranjem skupa poželjnih modela sa prikupljanjem novog saznanja.

### 2.3.2.6. Logika čupavih skupova

Logika čupavih (engl. **fuzzy**) skupova koristi se za rezonovanje sa dvosmislenim konceptima. U odnosu na klasičnu logiku kod koje objekt može pripadati ili nepripadati nekom skupu, kod logike fazi skupova je uveden koncept delimične pripadnosti. Delimična pripadnost nekom skupu se kvantitativno izražava preko **stepena članstva** odnosno **stepena tačnosti**, čija vrednost pripada intervalu  $(0,1)$ . Logika fazi skupova dozvoljava primenu **lingvističkih**

**promenljivih** pri čemu ona može imati proizvoljan broj lingvističkih vrednosti. Na primer, za lingvističku promenljivu „uzrast osobe” mogu se definisati sledeće lingvističke vrednosti: mlada, odrasla i stara. Pošto se radi o nejasnim odnosno nepreciznim (engl. crisp) notacijama, stepen njihovog članstva u zavisnosti od vrednosti atributa predstavlja se uz pomoć funkcija članstva,  $\mu(x)$ . Funkcije članstva za tri lingvističke vrednosti promenljive “uzrast osobe” date su na slici 2.15. Prva funkcija članstva,  $\mu_M$  odnosi se na lingvističku vrednost mlada. Ona pokazuje da se osoba smatra apsolutno mladom do 20 godina, nakon čega stepen članstva ove vrednosti opada sve do vrednosti 0 koju dostiže sa 40 godina. Između 20 i 40 godina osoba je delimično, manje ili više, mlada a nakon 40 za nju se ne može reći da je mlada. Slična objašnjenje dobija se analizom ostale dve funkcije članstva koje se odnose na odraslu osobu,  $\mu_O$ , i staru osobu,  $\mu_S$ . Logika čupavih skupova je razvijena do te mere da obuhvata većinu operacija karakterističnih za klasičnu teoriju skupova kao što su unija, presek, itd [35]. Iz nje su proistekli neki novi koncepti kao što su: čupavi brojevi i operacije sa njima, čupave relacije, čupava gramatika i sl.



Slika 2.15 Funkcije članstva tri lingvističke vrednosti promenljive "Uzrast osobe"

Definisanjem pravila (2.2.1.3) nad čupavim skupovima dobijaju se **čupavi sistemi bazirani na pravilima** [36]. Ovi sistemi rezonuju na osnovu fazi pravila tako što primaju podatke i generišu rezultate u numeričkom obliku. Zbog toga moraju biti sposobni da konvertuju numeričke u čupave vrednosti (**fazifikacija**) odnosno čupave u numeričke (**defazifikacija**).

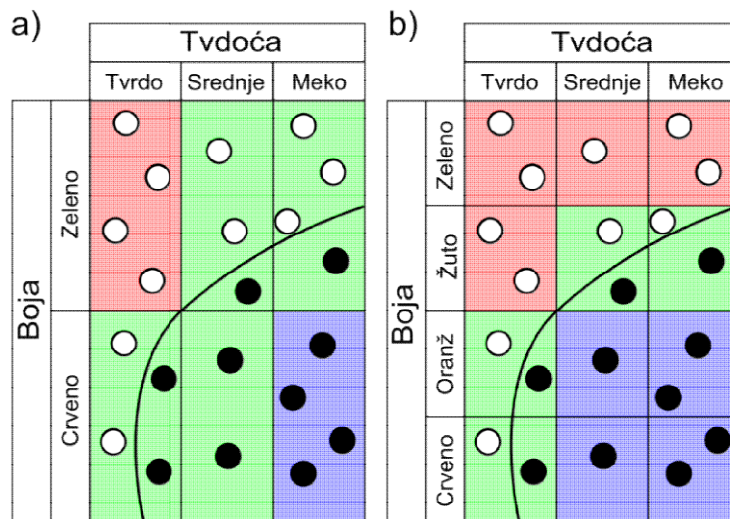


### 2.3.2.7. Logika grubih skupova

Logika grubih (engl. rough) skupova predstavlja komplementarni pristup modelovanja nejasnih koncepata u odnosu na čupave skupove [17]. Koristi se za rezonovanje sa nepreciznim podacima, međutim, da bi problem bio rešen potrebno je da nivo preciznosti bude **adekvatan**. Adekvatnost se definiše kao osobina sistema da pravi razliku među objektima različitih kategorija odnosno ne pravi razliku između objekata iste kategorije. Dve reprezentacije istog skupa, koji čine plodovi paradajza, prikazane su na slici 2.16. Ovaj skup se u teoriji grubih skupova naziva **univerzumom diskursa** i sadrži sve objekte koji su predmet analize. U konkretnom slučaju, analizira se da li je plod paradajza spreman za branje ili nije. U skladu sa problemom neprecizna notacija se može definisati kao “Zreo paradajz”. Plodovi koji zadovoljavaju notaciju označeni su crnim a koji ne zadovoljavaju belim kružićima. Između dve kategorije objekata stoji granična linija. Lista mogućih vrednosti atributa naziva se **domen atributa**. Na osnovu vrednosti atributa univerzum diskursa je podeljen na podzone koje se nazivaju **granule znanja**. Za objekte koji pripadaju istoj granuli kaže se da su **neprimetni** (engl. indiscernible) u odnosu na dati skup atributa. Od ovog skupa zavisi nivo preciznosti elementarnih notacija odnosno **granulacija univerzuma diskursa**. U primeru sa slike, slučaj (a) ima manju granulaciju jer ima veće granule znanja od primera (b). To je posledica manje **rezolucije** u definisanju atributa “Boja“. Pojedini objekti koji su bili neprimetni za skup atributa iz slučaja (a) postali su primetni za skup atributa iz slučaja (b). Osim povećanjem rezolucije atributa, granulacija univerzuma diskursa može se povećati uvođenjem novih atributa.

Definisanje nejasne notacije u teoriji grubih skupova, vrši se preko dve precizne notacije: donja aproksimacija i gornja aproksimacija. **Donju aproksimaciju** čine objekti koji se nalaze u **pozitivnom regionu**, na slici označenom plavom bojom. Pozitivni reon čine granule znanja koje sa sigurnošću mogu biti pridružene notaciji. **Negativni region** čine granule znanja, na slici crvene boje, koje sa sigurnošću ne pripadaju traženom konceptu. Razlika između negativnog i pozitivnog regiona je region granice, na slici označen zelenom bojom, za koji se ne može sa sigurnošću reći da li pripada pozitivnom ili negativnom regionu. **Gornju aproksimaciju** nejasne notacije čine objekti koji se nalaze u pozitivnom regionu i regionu granice. Za ove objekte se smatra da možda pripadaju regionu notacije. **Koeficijent tačnosti aproksimacije** određuje se kao količnik donje i gornje aproksimacije. Ovaj koeficijent direktno zavisi od granulacije

univerzuma diskursa. Bolja granulacija utiče na smanjenje regiona granice odnosno na povećanje koeficijenta tačnosti aproksimacije, na slici 0,2 za slučaj (a) i 0,5 za slučaj (b).



Slika 2.16 Grubi skupovi u prostoru atributa: a) niža granulacija, b) viša granulacija

### 2.3.3. Modeli bazirani na biološkim zakonima

Modeli bazirani na biološkim zakonima koriste se za pretraživanje prostora stanja koji sadrži potencijalno rešenje. Pri tome se primenjuju metodi koji su dobijeni simulacijom bioloških modela. **Evolucioni račun**, kao dominantni predstavnik modela veštačke inteligencije baziranih na biološkim zakonima, predstavlja matematički model Darwin-ove teorije evolucije. Pošto je evolucioni račun primenjen u istraživanju,

njemu je posvećeno posebno poglavlje ove disertacije (Poglavlje 4). Ostali metodi, koji pripadaju ovoj kategoriji veštačke inteligencije, nastali su modeliranjem **roja (jata)** odnosno **imunog sistema** živih organizama.

#### 2.3.3.1. Inteligencija roja

Ovaj metod računске inteligencije nastao je modeliranjem ponašanja životinja u velikoj grupi odnosno jatu ili roju. Drugim rečima, sistem je modeliran kao samoorganizujuća populacija autonomnih jedinki koje interaguju sa okolinom i međusobno. Svaka jedinka transformiše opservaciju u odgovarajuću akciju koja se vrši sa ciljem ostvarenja zadatka. Saradnja odnosno interakcija među jedinkama vrši se na osnovu ustanovljenih pravila. **Pravilo rastojanja** služi da

održi zahtevano rastojanje među jedinkama kako bi se izbeglo grupisanje. Razlog leži u činjenici da nije racionalno obrađivati više istovetnih rešenja. **Pravilo kohezije** pomera svaku jedinku ka centru mase čitavog jata kako bi se izbegla njegova fragmentacija. **Pravilo poravnanja** usmerava jedinke ka prosečnom cilju čitavoga jata. Postoje veliki broj algoritama nastalih modeliranjem ponašanja jata.

**Optimizacija kolonije mrava** je reprezentativni primer inteligencije roja kod koga modeli veštačkih mrava pretražuju prostor stanja u potrazi za potencijalnim rešenjem. “Mravi” se kreću na osnovu “tragova feromona” koji su jači na obećavajućim lokacijama koje se često posećuju a slabiji na mestima sa slabijom posetom. Rezultat je koncentrisanje kolonije oko mesta sa većim potencijalom za pronalaženje adekvatnog rešenja.

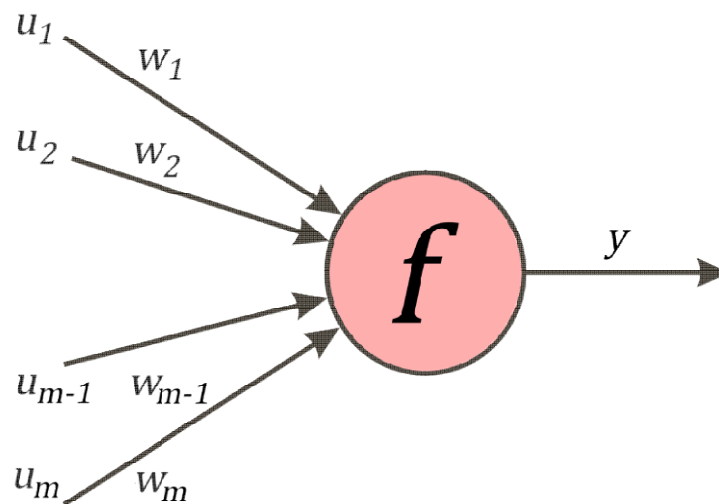
**Optimizacija roja čestica** je još jedan metod inteligencije roja koji služi za pretraživanje višedimenzionalnog prostora stanja. Kod ovog algoritma, rešenje se traži pomoću roja čestica koje se kreću u prostoru potencijalnih rešenja. Za razliku od prethodnog metoda, ovde se roj kreće u pravcu lidera odnosno jedinki sa najvećim potencijalom za nalaženje rešenja.

#### 2.3.3.2. Veštački imuni sistem

Ovaj metod računске inteligencije nastao je modeliranjem biološkog imunog sistema. Kako se imunologija razvija tako se i metode veštačke inteligencije nastale modeliranjem biološkog imunog sistema menjaju i razvijaju. Jedna od prvih strategija bila je **metoda negativne selekcije** koja se bazira na prepoznavanju razlike između **normalnog** odnosno **sopstvenog** slučaja i **patološkog** odnosno **stranog** slučaja. Pretraživanje prostora stanja vrši se uz pomoć **detektora** odnosno slučajnih uzoraka. Slučajevi koji razlikuju od uobičajenih detektuju se kao **anomalije**. Anomalije se podvrgavaju mutacijama nakon čega se vrši ponovna evaluacija. Na ovaj način algoritam uči odnosno prilagođava se da prepozna konkretnu vrstu anomalije. U fazi eksploatacije, kada se testira nepoznati slučaj, zapravo se određuje njegovo rastojanje u prostoru stanja od najbližeg detektora. Ako je ovo rastojanje manje od unapred zadatog praga slučaj je normalan dok se u protivnom klasifikuje kao patološki.

### 3. VEŠTAČKE NEURONSKE MREŽE

Model veštačkih neuronskih mreža inspirisan je funkcionisanjem biološkog nervnog sistema. Ljudski mozak, na primer, uprkos manjem računskom kapacitetu obavlja zadatke percepcije okoline brzinom koja je nedostižna za konvencionalne digitalne računare. Razlog leži u činjenici da čitav mozak radi neprestano dok je dominantan deo digitalnog računara zapravo pasivno skladište podataka [37]. Očigledno je da mozak postiže bolje performanse zahvaljujući sposobnosti paralelne obrade informacija. Druga fundamentalna prednost mozga je sposobnost učenja odnosno izgradnje pravila na osnovu sopstvenih iskustava. Sposobnost učenja leži u osnovi adaptivnosti živih bića na dejstvo spoljašnjih uticaja. Veštačke neuronske mreže su osmišljene sa ciljem da iskoriste prednosti principa na kojima se bazira funkcionisanje biološkog nervnog sistema.

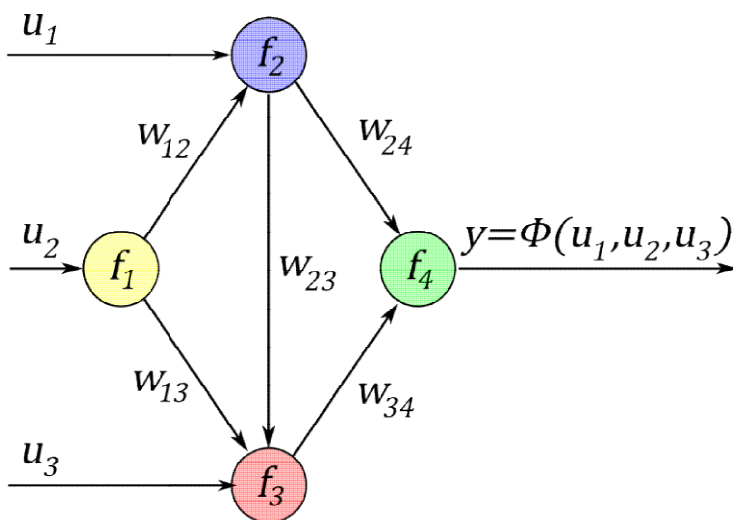


Slika 3.1 Veštački neuron

Osnovna gradivna jedinica veštačkih neuronskih mreža je veštački neuron, predstavljen grafički na slici 3.1. On je rezultat matematičkog modeliranja biološkog neurona odnosno nervne ćelije koja je osnovni gradivni element nervnog sistema kod živih bića. Veštački neuron prima informacije iz okoline, ili od drugih neurona iz mreže, uz pomoć svojih ulaznih terminala,  $u_1, u_2, \dots, u_m$ , gde je sa  $m$  označen broj terminala. Svakom od ulaznih signala pridruženo je odgovarajuće pojačanje odnosno težinski koeficijent,  $w_1, w_2, \dots, w_m$ . Funkcija težinskog koeficijenta je inhibiranja signala, u slučaju negativne vrednosti, odnosno pobuđivanja signala,

ako je njegova vrednost pozitivna. Podešavanjem težinskih koeficijenata vrši se obučavanje odnosno treniranje neurona kako bi on mogao da obavlja predviđenu funkciju. Na osnovu ovako podešenih ulaza veštački neuron proizvodi izlazni signal,  $y$ , u skladu sa karakteristikama svoje prenosne funkcije,  $f$ . U opštem slučaju, prenosna funkcija veštačkog neurona može biti svaka diferencijabilna funkcija ali neke od njih se redovno koriste u praksi jer su ispoljile izuzetno povoljne karakteristike.

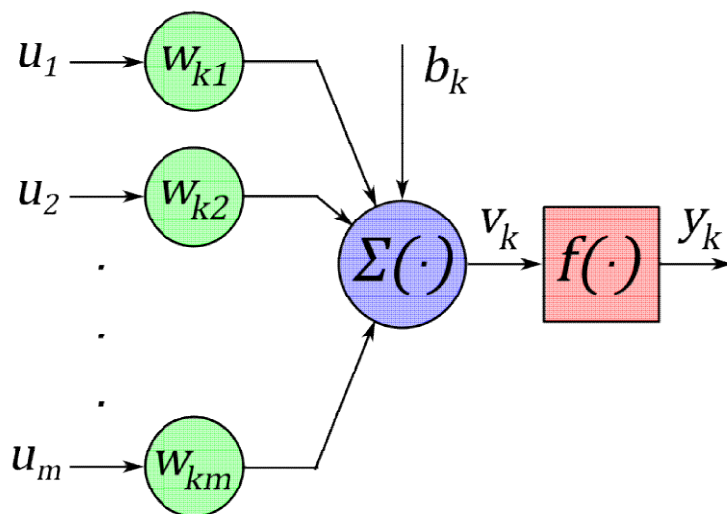
Na slici 3.2 prikazana je tipična struktura veštačke neuronske mreže na kojoj su neuroni predstavljeni čvorovima, u obliku kružića koji su obeleženi odgovarajućim prenosnim funkcijama, i veza među njima, koje su predstavljene strelicama sa odgovarajućim faktorima pojačanja. Često su prenosne funkcije svih čvorova, u okviru jedne neuronske mreže, iste pa se njihova oznaka može izostaviti. Neuronska mreža, sa slike, može biti shvaćena kao funkcija  $\Phi$  triju nezavisnih promenljivih  $u_1, u_2$  i  $u_3$ . Funkcija mreže,  $\Phi$ , dobija se kombinovanjem prenosnih funkcija svakog od neurona,  $f_1, f_2, f_3$  i  $f_4$  uz pomoć veza koje među njima postoje. Snaga veze koja vodi od neurona  $f_i$  do neurona  $f_j$  određena je odgovarajućim faktorom pojačanja,  $w_{ij}$ . Prema tome, izlaz iz neuronske mreže u opštem slučaju zavisi od tri faktora: **modela neurona** odnosno njihovih prenosnih funkcija, **topologije mreže**, odnosno rasporeda neurona i veza među njima i težinskih koeficijenata odnosno **algoritma treniranja** koji se primenjuje za njihovo podešavanje.



Slika 3.2 Tipična struktura neuronske mreže [38]

### 3.1. Model veštačkog neurona

Na osnovu “grubog” modela veštačkog neurona, prikazanog na slici 3.1, može se razviti detaljniji model, slika 3.3, koji dozvoljava izvođenje matematičkih relacija. U kontekstu čitave mreže neuron je obeležen rednim brojem  $k$  a njegovi ulazi odnosno **sinapse**, sa  $u_i, i = 1, \dots, m$ , gde je  $m$  broj ulaznih signala. Ulazni signal može imati bilo koju realnu vrednost, bilo pozitivnu ili negativnu. Pojačanje ulaznog signala  $j$  na ulazu u neuron  $k$  obeleženo je sa  $w_{kj}$ . Uobičajeno je da prvi indeks označava redni broj neurona a drugi redni broj ulaza na koji se pojačanje odnosi. Sistematsko označavanje parametara mreže olakšava izvođenje algoritama za treniranje kao i rukovanje mrežom u fazi eksploatacije. Osim tipičnih ulaza, model neurona uključuje poseban signal nazvan bias ili prednapon, označen sa  $b_k$ . Alternativni prikaz biasa je kao redovnog ulaznog signala sa konstantnom jediničnom vrednošću,  $u_o = 1$ , i pojačanjem  $w_{ko} = b_k$ . U matematičkom smislu oba pristupa svode se na isto. Nakon pojačavanja, signali se dovode na ulaz u veštački neuron gde se vrši njihova obrada pomoću dva funkcionalna elementa.



Slika 3.3 Blok dijagram veštačkog neurona

#### 3.1.1. Sumator

Prvi funkcionalni element veštačkog neurona je **sumator** (engl. adder), na slici obeležen operatorom  $\Sigma(\cdot)$ . Sumator ima funkciju zbrajanja dejstava svih ulaza (uključujući i bias). Na osnovu blok dijagrama, izlaz iz sumatora,  $v_k$ , sa  $m$  ulaznih signala određen je jednačinom

$$v_k = \sum_{i=1}^m w_{ki} u_i + b_k \quad (3.1)$$

Retko se za spajanje ulaznih signala koriste neke druge funkcije kao što je proizvod

$$v_k = \prod_{i=1}^m w_{ki} u_i + b_k \quad (3.2)$$

Na osnovu prethodne dve jednačine se vidi da je funkcija biasa,  $b_k$ , u zavisnosti od znaka, da poveća ili smanji opšti nivo izlaza iz sumatora. Na ovaj način bias utiče na položaj tačke aktivacije veštačkog neurona [38,39].

### 3.1.2. Aktivaciona funkcija

Aktivaciona funkcija ili **prenosna funkcija**, na blok dijagramu označena operatorom  $f(\cdot)$ , određuje vrednost izlaza neurona,  $y_k$ , u zavisnosti od vrednosti izlaza sumatora  $v_k$ . Pod pretpostavkom da je izlaz iz sumatora definisan jednačinom (3.1) izlaz iz neurona određen je sledećom jednačinom

$$y_k = f(v_k) = f\left(\sum_{i=1}^m w_{ki} u_i + b_k\right) = f(\mathbf{W}_k \mathbf{U} + b_k) \quad (3.3)$$

gde je sa  $\mathbf{W}_k = [w_{k1} \ w_{k2} \ \dots \ w_{km}]$  označena matrica vrsta koju čine pojačanja neurona  $k$  a sa  $\mathbf{U} = [u_1 \ u_2 \ \dots \ u_m]^T$  vektor ulaza. Već je rečeno da je uslov da bi funkcija bila upotrebljena kao aktivaciona da bude diferencijabilna na čitavom intervalu definisanosti. Međutim, u praksi su se profilisale određene funkcije sa pa posebno povoljnim karakteristikama koje olakšavaju postupak treniranja i eksploatacije neuronske mreže. Izbor odgovarajuće aktivacione funkcije uslovljen je prirodom problema koji mreža treba da reši odnosno njenom namenom.

#### 3.1.2.1. Odskočna funkcija

Odskočna ili Heaviside-ova (Oliver Heaviside, 1850-1925) funkcija, slika 3.4 a, trenutno menja svoje stanje, sa 0 na 1, u momentu kad ulazna promenljiva dostigne nulu. Matematičke definicija ove funkcije data je izrazom

$$f(v) = \begin{cases} 1 & \forall v \geq 0 \\ 0 & \forall v < 0 \end{cases} \quad (3.4)$$

Ovakvu karakteristiku nemoguće je ostvariti, u bukvalnom smislu, ali je brzina savremenih računara toliko velika da se trajanje promene između njene dve vrednosti odskočne funkcije može zanemariti. Neuron sa odskočnom aktivacionom funkcijom i binarnim ulazima čine posebnu grupu veštačkih neurona koji su nazvani McCulloch–Pitt-ove kapije (engl. McCulloch–Pitts gates) po svojim tvorcima (Warren Sturgis McCulloch, 1898 –1969 i Walter Harry Pitt, Jr., 1923 –1969). Zbog toga što imaju samo dva stanja izlaza, ovi neuroni pogodni su za modeliranje logičkih funkcija. Dokazano je da se primenom pojedinačnog neurona sa odskočnom aktivacionom funkcijom moguće realizovati bilo koju elementarnu logičku funkciju (negaciju, konjunkciju i disjunkciju) sa proizvoljnim brojem elemenata [39]. Pošto se svaka logička funkcija može izraziti preko elementarnih logičkih funkcija znači da se primenom veštačkih neurona sa odskočnom aktivacionom funkcijom može ostvariti bilo koja logička funkcija.

### 3.1.2.2. Logička funkcija sa pragom

Kod logičkih funkcija sa pragom (engl. piecewise-linear function) postoji interval prelaska sa jedne na drugu graničnu vrednost. slika 3.4 b prikazuje logičku funkcije sa pragom i kodomenom  $[0 \ 1]$ . Matematička definicija iste funkcije data je jednačinom

$$f(v) = \begin{cases} 1, & v \geq \frac{1}{2k} \\ kv + \frac{1}{2}, & -\frac{1}{2k} > v > \frac{1}{2k} \\ 0, & v \leq -\frac{1}{2k} \end{cases} \quad (3.5)$$

gde je  $k$  koeficijent pravca prelaznog dela funkcije. Prethodni izraz istovetan je sa izlazom karakteristikom nelinearnog pojačivača pa se on najčešće koristi za fizičku realizaciju logičke aktivacione funkcije sa pragom. Ova funkcija je pogodna za modeliranje promene stanja u realnim sistemima jer se promena između graničnih stanja ne dešava trenutno. Prelazni opseg se definiše shodno potrebama ali se najčešće koristi simetrični opseg ulaza, od  $-\frac{1}{2k}$  do  $\frac{1}{2k}$  sa opsegom izlaza  $[0 \ 1]$  ili  $[-1 \ 1]$ .



### 3.1.2.3. Sigmoidne funkcije

Ova klasa aktivacionih funkcija podrazumeva sve striktno rastuće funkcije koje imaju oblik latiničnog slova “s”. Sigmoidne funkcije preslikavaju neograničeni ulaz u obostrano ograničeni izlaz. Primeri ovih funkcija su logistička, arkus tangensna ili hiperbolička tangens funkcija. Često se pod sigmoidnom podrazumeva samo **logistička** funkcija jer se ona ubedljivo najčešće koristi kao aktivaciona funkcija u praktičnim realizacijama neuronskih mreža. Logistička sigmoidna funkcija definisana je izrazom

$$f(v) = \frac{1}{1 + e^{-\alpha v}} \quad (3.6)$$

a njen grafik za različite vrednosti parametra nagiba,  $\alpha$ , prikazan je na Slici 3.4 c. Crne strelice označavaju tendenciju porasta nagiba. Ova funkcija ima povoljnu osobinu koja proističe iz izraza

$$f'(v) = \alpha f(v)(1 - f(v)) \quad (3.7)$$

da se izračunavanje njenog parcijalnog izvoda svodi na izračunavanje polinoma drugog stepena. Ova osobina olakšava primenu algoritama za treniranje koji uključuju izračunavanje parcijalnog izvoda aktivacione funkcije. Vrednost ove funkcije u koordinatnom početku je, na primer,  $f(0) = \frac{1}{2}$  bez obzira na parametar nagiba,  $\alpha$ . Na osnovu jednačine (3.7) vrednost nagiba logističke sigmoidne funkcije u ovoj tački je

$$f'(0) = \frac{\alpha}{2} \quad (3.8)$$

Za velike vrednosti parametra nagiba,  $\alpha \rightarrow \infty$ , oblik logističke sigmoidne funkcije teži obliku odskočne funkcije. Alternativa logističkoj funkciji, sa sličnim pozitivnim karakteristikama, je hiperbolička tangens funkcija, definisana izrazom

$$f(v) = \tanh(\alpha v) = \frac{e^{\alpha v} - e^{-\alpha v}}{e^{\alpha v} + e^{-\alpha v}} \quad (3.9)$$

Grafik ove funkcije, za različite vrednosti parametra  $\alpha$ , prikazan je na slici 3.4 d. I na ovom grafiku, crne strelice označavaju tendenciju povećanja nagiba. Hiperbolička tangens funkcija ima

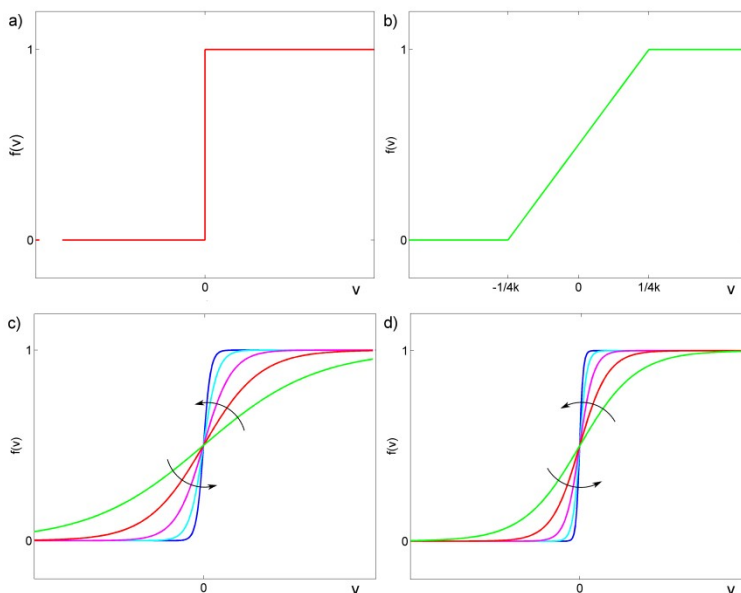
veoma sličan oblik logističkoj aktivacionoj funkciji. Diferenciranjem ove funkcije (3.9) dobija se izraz

$$f'(v) = \alpha(1 - f(v)^2) \quad (3.10)$$

pa se nalaženje izvoda i ovde svodi na izračunavanje polinoma drugog stepena. Vrednost tangens hiperboličke funkcije u koordinatnom početku je  $f(0) = 0$  pa je na osnovu (3.10)

$$f'(0) = \alpha \quad (3.11)$$

Nagib ove funkcije u koordinatnom početku veći je za duplo od nagiba logističke funkcije što znači da i ova funkcija teži odskočnoj za  $\alpha \rightarrow \infty$ .



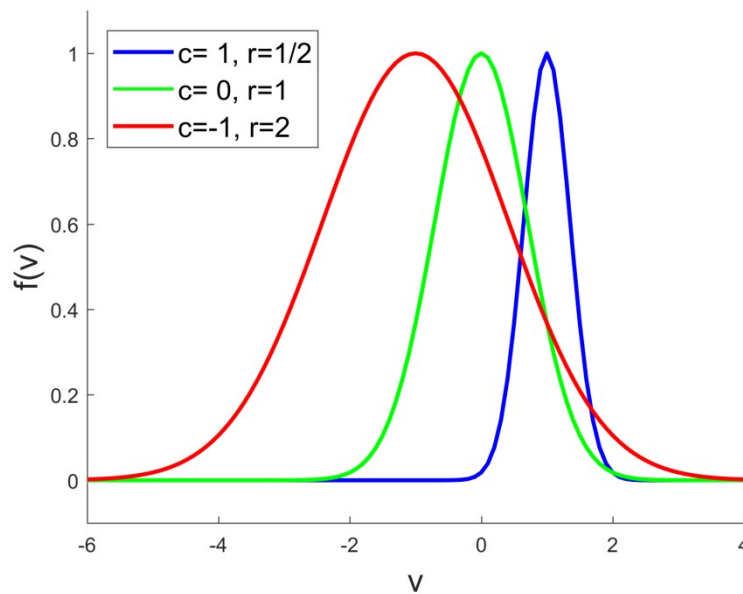
Slika 3.4 Aktivacione funkcije: a) odskočna, b) logička sa pragom, c) logistička, d) hiperbolička tangens

#### 3.1.2.4. Radijalne funkcije

Radijalne funkcije predstavljaju posebnu klasu funkcija sa osobinom da im odziv monotono opada ili monotono raste sa rastojanjem od centralne tačke. Oblik radijalne funkcije i položaj njene centralne tačke određeni su parametrima funkcije. Tipičan primer radijalne funkcije je Gausova funkcija

$$f(v) = e^{-\frac{(v-c)^2}{r^2}} \quad (3.12)$$

gde je  $c$  koordinata centralne tačke a  $r$  radijus funkcije. Grafici ove funkcije za različite vrednosti parametara  $c$  i  $r$  dati su na slici 3.5. Odziv Gausove funkcije značajan je u okolini centralne tačke a neznatan ka periferiji. Zbog ovakve karakteristike, Gausove funkcije su pogodne za formiranje baze elementarnih funkcija koje služe za aproksimaciju proizvoljne funkcije. Baza elementarnih funkcija je skup Gausovih funkcija sa različitim položajima centralnih tačaka. Aproksimacija zadate funkcije dobija se u vidu linearne kombinacije elementarnih funkcija.



Slika 3.5 Gausove funkcija za različite vrednosti parametara

### 3.1.2.5. Stohastički funkcija

Za razliku od prethodnih modela, sa precizno definisanim relacijama između ulaza i izlaza, postoje neuronske mreže sa stohastičkim modelom. Kod ovakvih mreža aktivaciona funkcija ima samo dva stanja a prelazak između jednog u drugo stanje određen je probabilistički

$$f(v) = \begin{cases} 1, & \text{sa verovatnoćom } P(v) \\ 0, & \text{sa verovatnoćom } 1 - P(v) \end{cases} \quad (3.13)$$

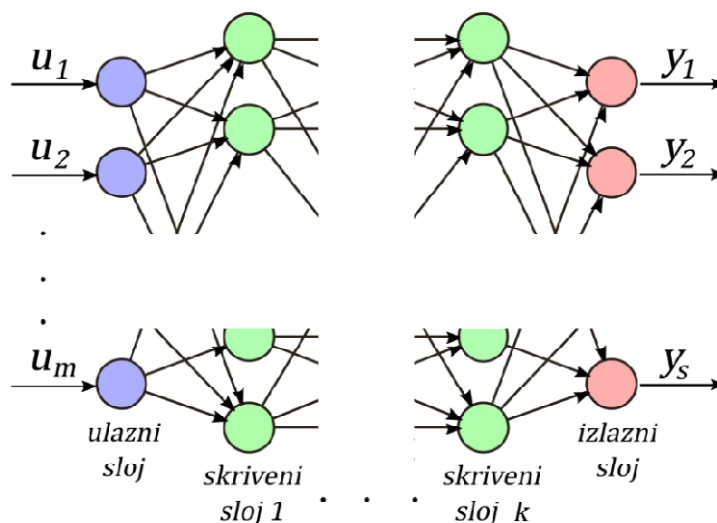
Za funkciju verovatnoće  $P(v)$  najčešće se koristi sigmoidna funkcija definisana izrazom

$$f(v) = \frac{1}{1 + e^{-\frac{v}{T}}} \quad (3.14)$$

gde je  $T$  parametar nazvan **pseudotemperatura** koji se koristi za kontrolu nivoa poremećaja a samim tim i neizvesnosti izlaznog stanja neurona. Za visoke vrednosti pseudotemperature,  $T \rightarrow \infty$ , stohastički neuron teži determinističkom modelu. Alternativni pristup stohastičkim neuronskim mrežama bio bi primena stohastičkih pojačanja između neurona sa determinističkim aktivacionim funkcijama.

### 3.2. Topologija veštačkih neuronskih mreža

Način povezivanja neurona u neuronsku mrežu naziva se topologijom neuronske mreže. Umesto topologije često se upotrebljavaju termini arhitektura ili struktura mada oni mogu označavati fizičku realizaciju neuronske mreže. Topologija ima presudnu ulogu na performanse neuronske mreže kako u fazi treniranja tako i u fazi eksploatacije. Veštački neuroni su najčešće prostorno organizovani u grupe nazvane **slojevima** pri čemu se prvi sloj naziva **ulaznim** a poslednji **izlaznim** slojem neurona. Između ulaznog i izlaznog može se naći proizvoljan broj slojeva, koji se nazivaju **skrivenim** slojevima neurona. Na slici 3.6 predstavljena je tipična neuronska mreža kod koje je svaki neuron jednog sloja povezan sa svim neuronima koji pripadaju sledećem sloju.

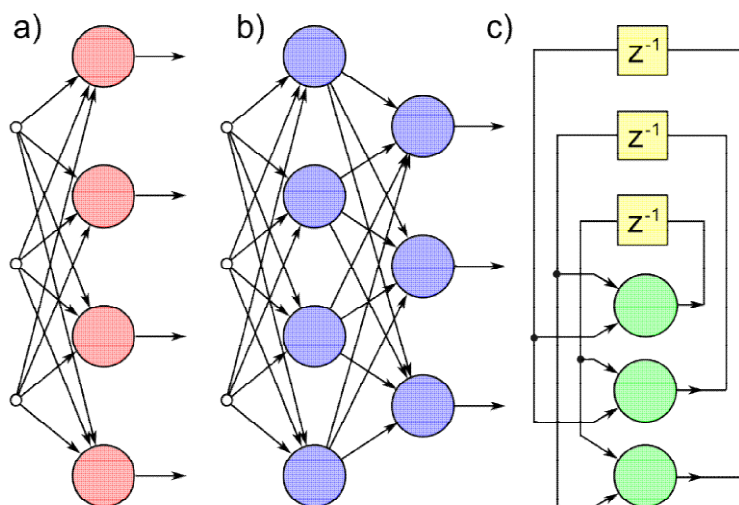


Slika 3.6 Topologija višeslojne neuronske mreže

U opštem slučaju moguća je veza neurona sa elementima udaljenih slojeva, koji u strukturi neuronske mreže dolaze posle sledećeg ili sa prethodnim slojevima neurona. U pogledu topologije neuronske mreže se dele na tri osnovne kategorije.

### 3.2.1. Jednoslojna neuronska mreža sa direktnim prostiranjem signala

Ovo je najjednostavniji oblik veštačkih neuronskih mreža jer sadrži samo jedan sloj neurona (slika 3.7 a). Ovaj sloj neurona, obeleženih velikim kružićima, smatra se izlaznim slojem. Sami ulazni terminali, obeleženi malim kružićima, se u literaturi tradicionalno nazivaju ulaznim slojem, mada se na njima ne vrši nikakvo izračunavanje. Ako se prihvati takva terminologija onda jednoslojna neuronska mreža ima ulazni sloj, koga čine ulazni terminali, i izlazni sloj koga čine neuroni. Sve veze se kreću od ulaznog ka izlaznom sloju što znači da je struktura ove mreže strogo aciklična.



Slika 3.7 Topologije veštačkih neuronskih mreža: a) jednoslojna, b) višeslojna, c) rekurentna

### 3.2.2. Višeslojna neuronska mreža sa direktnim prostiranjem signala

Kod sledeće topologije neuronskih mreža, takođe se direktnim prostiranjem signala, osim ulaznog i izlaznog sloja postoji proizvoljan broj skrivenih slojeva. Na slici 3.7 b prikazan je primer ovakve mreže sa jednim skrivenim slojem. Dodavanje novih slojeva neurona omogućava ostvarenje kompleksnije funkcije mreže odnosno funkcije višeg reda nego je to moguće kod jednoslojne neuronske mreže [39]. Tipično za višeslojne neuronske mreže je postojanje veza

samo između susednih slojeva neurona. Ovakvo povezivanje naziva se još i **međuslojno** povezivanje. Takođe je uobičajeno da izlaz svakog neurona bude povezan sa svim neuronima susednog sloja. U takvom slučaju povezivanje se naziva **potpuno** a u protivnom **delimično** povezivanje. Ako su neuroni međuslojno i potpuno povezani topologija mreže može se definisati nizom koji označava broj neurona u svakom od slojeva. Za primer na slici 3.7 b to bi bilo 3-4-3 jer ima 3 ulazna terminala, 4 neurona u skrivenom sloju i 3 izlazna neurona.

### 3.2.3. Rekurentne neuronske mreže

U prethodnom poglavlju je analizirana veštačka neuronska mreža sa direktnim prostiranjem signala (engl. feed-forward), od ulaza ka izlazu. Kod ovakvih mreža dimenzija ulaznog i izlaznog signala je unapred definisana. Nije potrebno voditi računa o sinhronizaciji neurona jer se pretpostavlja da se sva izračunavanja dešavaju istovremeno. Međutim, mreže sa direktnim prostiranjem nisu sposobne za obradu signala promenljive dužine, na primer sabiranje dva binarna broja dovođenjem na ulaz jednog po jednog bita. Da bi se ostvarila obrada signala promenljive dužine, osim vrednosti tekućih ulaza mreža mora da uzme u obzir vrednosti prethodnih ulaza. Ovo se postiže uvođenjem vremenskog faktora odnosno **kašnjenja** u model veštačkog neurona. Ako na ulaz dospe signal u momentu  $t$ , izlaz iz neurona se dobija u momentu  $t + 1$ . Mreža koja osim elemenata kašnjenja sadrži bilo kakav ciklus, odnosno povratnu spregu, naziva se **rekurentna** neuronska mreža. Na datom primeru rekurentne mreže, slika 3.7 c. izlaz svakog neurona dovodi se na ulaz svih neurona osim sopstvenog. Ovakva arhitektura dobila je naziv Hopfield-ova mreža, po američkom naučniku (John Hopfield) koji ju je prvi predložio. Bez obzira da li se radi o samo povezivanju ili ne, ciklusi u rekurentnoj mreži igraju ulogu skladišta informacija radi njihovog korišćenja u budućnosti. Uvođenje povratnih sprega u topologiju mreže dramatično menja njen kapacitet učenja i performanse. U datom primeru neuronske mreže za obradu diskretnih signala, elementi jediničnog kašnjenja označeni su sa  $z^{-1}$ . Rekurentne mreže se isto tako primenjuju za obradu kontinualnih signala.

### 3.2.4. Neuronske mreže sa radijalnom funkcijom

U principu, svaka neuronska mreža koja sadrži radijalne aktivacione funkcije može se smatrati neuronskom mrežom sa radijalnom bazom. Međutim, u praksi se pod neuronskom

mrežom sa radijalnom funkcijom najčešće smatra posebna topologija mreže sa linearnim matematičkim modelom.

$$f(u) = \sum_{i=1}^m w_i g_i(u) \quad (3.15)$$

gde su  $w_i$  pojačanja a  $g_i$ , radijalne aktivacione funkcije. Da bi model bio linearan potrebno je da radijalne funkcije budu nepromenljive, odnosno sa konstantnim parametrima, i smeštene u jedinstvenom skrivenom sloju. Osim skrivenog sloja, ove neuronske mreže poseduju izlazni sloj neurona sa linearnom aktivacionom funkcijom, najčešće jediničnog pojačanja, i podrazumevani ulazni sloj koji sadrži terminale. Blok dijagram neuronske mreže sa radijalnom aktivacionom funkcijom prikazan je na slici 3.8. Ulaz u nelinearnu aktivacionu funkciju dat je jednačinom

$$v_i = \|u \times \bar{W}_i - C_i\|, \quad i = 1, \dots, m \quad (3.16)$$

gde je  $\bar{W}_i$  vektor težinskih koeficijenata skrivenog neurona  $i$ ,  $C_i$  je vektor položaja njegove centralne tačke a  $m$  broj neurona u skrivenom sloju. Operator  $\|\cdot\|$  je operator konfluencije [40] koji se u slučaju mreža sa radijalnom funkcijom svodi na Euklidsku normu. U praksi su pojačanja ulaza u skriveni sloj najčešće jednaka jedinici. Izlaz skrivenog sloja, za slučaj Gausove aktivacione funkcije, dat je jednačinom

$$\bar{y}_i = e^{-\frac{v_i^2}{r_i^2}}, \quad i = 1, \dots, m \quad (3.17)$$

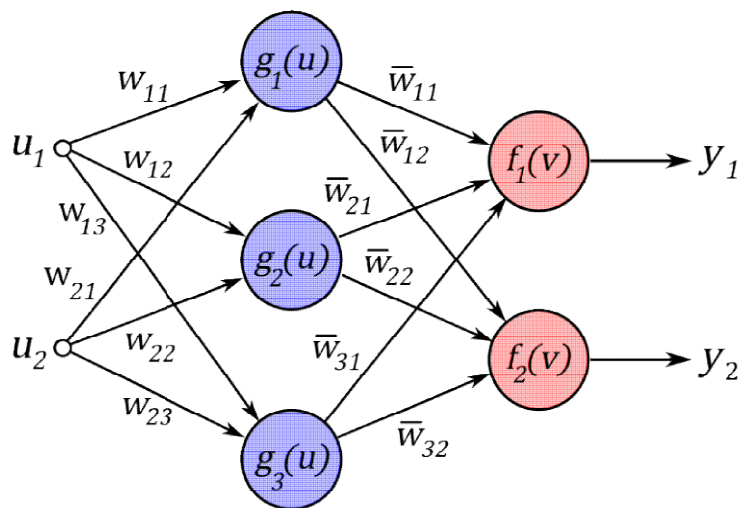
Izlazni sloj služi za sumiranje izlaza skrivenog sloja. Za slučaj linearne aktivacione funkcije izlaznog sloja sa jediničnim pojačanjem vektor izlaza mreže dat je izrazom

$$Y = W\bar{Y}, \quad i = 1, \dots, m \quad (3.18)$$

gde je  $W$  vektor pojačanja ulaza u izlazni sloj a  $\bar{Y}$  vektor izlaza iz skrivenog sloja, čije su koordinate određene na osnovu jednačine (3.17).

Neuronske mreže sa radijalnim aktivacionim funkcijama koriste se za aproksimaciju proizvoljne kontinualne nelinearne funkcije. Vrednost aproksimacije određena je položajem središnje tačke neurona radijusom aktivacionih funkcija i vrednostima težinskih koeficijenata izlaznog sloja. Vrednosti ovih parametara određuju se uz pomoć različitih algoritama za

treniranje mreže. U slučaju jednostavnijih nelinearnosti, parametri aktivacionih funkcija se unapred fiksiraju, na osnovu utvrđenog rasporeda, a treniranje se primenjuju sa ciljem određivanja vrednosti težinskih koeficijenata izlaznog sloja,  $W$ . Raspored središnjih tačaka neurona značajno utiče na ponašanje mreže. Najčešće je broj središta neurona jednak broju uzoraka odnosno podataka koji se aproksimiraju. U slučaju da je broj središta manji od broja podataka ona se raspoređuju na osnovu različitih strategija, na primer slučajnim ili jednolikim rasporedom. Primenom ovih mreža moguće je aproksimirati proizvoljnu neprekidnu funkciju. Za to je, međutim, potreban dovoljno veliki broj neurona koji mrežu može učiniti previše glomaznom. Funkcionalna aproksimacija dobija se smanjivanjem broja neurona kroz postupak nelinearne optimizacije [40].



Slika 3.8 Neuronska mreža sa radijalnom funkcijom

### 3.3. Algoritam treniranja

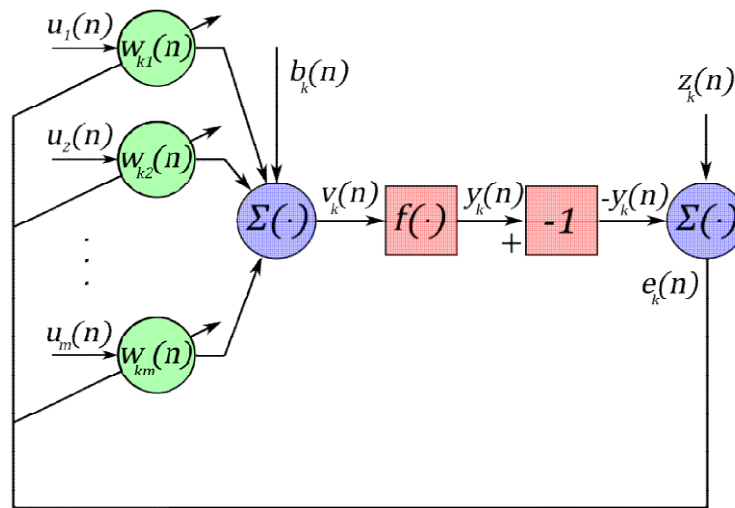
Fundamentalna osobina neuronske mreže je sposobnost unapređenja svojih performansi kroz proces učenja odnosno treniranja. Treniranje se izvodi sa ciljem određivanja parametara mreže. Najčešće su to samo težinski koeficijenti veza između neurona, zajedno sa biasom, ali mogu biti i neki drugi parametri, na primer parametri aktivacionih funkcija. U ređim slučajevima, sam izbor aktivacionih funkcija ili topologije mreže mogu takođe biti predmet treniranja. Podešavanje parametara mreže vrši se, na osnovu stimulansa iz okoline, kroz niz iterativnih koraka. U idealnom slučaju, svaki korak dovodi do pozitivne adaptacije mreže u odnosu na



okolinu. Skup pravila definisanih sa ciljem treniranja mreže naziva se algoritam treniranja ili algoritam učenja. Postoje veliki broj različitih algoritama za treniranje veštačkih neuronskih mreža svaki od njih sa svojstvenim prednostima i manama. U osnovi ovih razlika stoji način podešavanja parametara mreže kao i način predstavljanja (modeliranja) okoline.

### 3.3.1. Učenje korigovanjem greške

Slika 3.9 ilustruje princip učenja veštačke neuronske mreže metodom korigovanja greške na primeru jednog neurona.



Slika 3.9 Učenje korigovanjem greške

Cilj procedure učenja neurona  $k$  je podešavanje njegovih težinskih koeficijenata,  $w_{ki}$ ,  $i = 1, \dots, m$ . Ovo podešavanje se obavlja sukcesivno, u vidu iterativnog postupka, zbog čega su promenjive veličine označene kao funkcije diskretnog parametra  $n$ . Ovaj parametar predstavlja redni broj koraka u okviru iterativne procedure treniranja mreže. Signal greške,  $e_k(n)$ , u koraku  $n$ , dobija se kao razlika između željene,  $z_k(n)$ , i ostvarene vrednosti izlaza,  $y_k(n)$ , neurona

$$e_k(n) = z_k(n) - y_k(n) \quad (3.19)$$

Strategija podešavanja težinskih koeficijenata treba da obezbedi smanjenje greške odnosno približavanje izlaznog signala željenim vrednostima. U ovu svrhu može poslužiti funkcija koštanja definisana izrazom

$$E_k(n) = \frac{1}{2} e_k^2(n) \quad (3.20)$$

Ova funkcija, poznata još pod nazivom **indeks performansi** [38], predstavlja energiju greške neurona  $k$  u iterativnom koraku  $n$ . Minimizacijom energije greške (3.20) dobija se pravilo za podešavanje težinskih koeficijenata poznato kao delta pravilo ili pravilo opadajućeg gradijenta

$$\Delta w_{ki}(n) = \eta e_k(n) f'_k(v_k(n)) u_i(n) \quad (3.21)$$

gde je  $f' = \frac{\partial f(v)}{\partial(v)}$  izvod aktivacione funkcije po argumentu  $v$ ,  $v_k(n)$  je argument aktivacione funkcije, u momentu  $n$ , određen jednačinom (3.1) a  $\eta$  je pozitivna konstanta nazvana **koeficijent učenja**. Da bi korekcija pojačanja,  $\Delta w_{ki}$ , bila određena potrebno je da aktivaciona funkcija  $f$  bude diferencijabilna na čitavom domenu. Osim toga njeno izračunavanje treba da bude dovoljno jednostavno da ne oduzima previše računskih resursa. Zbog toga se u praksi biraju specijalne aktivacione funkcije sa povoljnim karakteristikama (Paragraf 3.1.2). Na osnovu (3.21) dobija se jednačina za izračunavanje pojačanja ulaza  $i$  u neuron  $k$  koja se odnosi na naredni korak algoritma odnosno korak  $n + 1$

$$w_{ki}(n + 1) = w_{ki}(n) + \Delta w_{ki}(n) \quad (3.22)$$

gde  $w_{ki}(n)$  pojačanje istog ulaza u koraku  $n$ .

#### 3.3.1.1. Algoritam sa povratnim prostiranjem greške

Dominantni predstavnik algoritama na bazi korekcije greške je algoritam sa povratnim prostiranjem greške (engl. Backpropagation algorithm). Ovaj algoritam bazira se na minimizaciji ukupne energije greške (3.20) svih neurona jednog sloja

$$E(n) = \sum_{k \in S} E_k(n) = \sum_{k \in S} e_k^2(n) \quad (3.23)$$

gde je sa  $S$  označen skup svih neurona u okviru jednog sloja. Minimalna vrednost ukupne energije greške dobija se sumiranjem izraza (3.21), za sve neurone, što svedoči o lokanom karakteru algoritma sa povratnim prostiranjem greške [38].

U pogledu izračunavanja greške izlaza iz neurona imamo dva različita slučaja. Ako se neuron nalazi u izlaznom sloju onda je poznata njegova željena vrednost, na osnovu podataka za treniranje mreže, te se greška izračunava direktno uz pomoć izraza (3.19). Međutim, ukoliko je neuron smešten u skrivenom sloju njegov željni izlaz nije eksplicitno definisan. U tom slučaju pristupa se izračunavanju greške rekursivnim postupkom, na osnovu grešaka izlaznih veličina neurona sa kojima predmetni neuron direktno povezan. Izračunavanjem gradijenta ukupne energije greške u koraku  $n$  (3.23) dobija se

$$\delta_k(n) = -\frac{\partial E(n)}{\partial v_k(n)} = e_k(n)f'_k(v_k(n)) \quad (3.24)$$

na osnovu čega se izraz (3.21) može preformulisati u oblik

$$\Delta w_{ki}(n) = \eta \delta_k(n) u_i(n) \quad (3.25)$$

Gradijent totalne energije skrivenog sloja može se i drugačije odrediti na osnovu pravila složenog izvoda

$$\delta_k(n) = -\frac{\partial E(n)}{\partial v_k(n)} = \sum_{j \in T} \frac{\partial E(n)}{\partial \bar{v}_j(n)} \frac{\partial \bar{v}_j(n)}{\partial v_k(n)} = \sum_{j \in T} \bar{\delta}_k(n) \frac{\partial \bar{v}_j(n)}{\partial v_k(n)} \quad (3.26)$$

gde je sa  $\bar{\delta}_k(n)$  označen gradijent ukupne energije greške narednog sloja, sa  $\bar{v}_j(n)$  ulaz u aktivacionu funkciju neurona narednog sloja sa redni brojem  $j$  a sa  $T$  skup svih neurona u narednom sloju. Na osnovu izraza (3.1), (3.3) kao i činjenice da je ulaz u naredni sloj,  $\bar{u}_i(n)$  jednak izlazu iz tekućeg sloja,  $y_i$ , sledi

$$\bar{v}_j(n) = \sum_{i \in T} \bar{w}_{ji}(n) y_i(n) + \bar{b}_j(n) = \sum_{i \in T} \bar{w}_{ji}(n) f_i(v_i(n)) + \bar{b}_j(n) \quad (3.27)$$

gde su sa  $\bar{w}_{ji}(n)$  označen težinski koeficijent između neurona  $j$  skrivenog sloja i neurona  $i$  narednog sloja u koraku  $n$  i sa  $\bar{b}_j(n)$  odgovarajući biasi. Diferenciranjem prethodne jednačine po  $v_k(n)$  dobija se

$$\frac{\partial \bar{v}_j(n)}{\partial v_k(n)} = \bar{w}_{jk}(n) f'_k(v_k(n)) \quad (3.28)$$

pa se na osnovu (3.25), (3.26) i prethodne jednačine dobija konačna jednačina za korekciju težinskog koeficijenta  $k$  u skrivenom sloju

$$\Delta w_{ki}(n) = \eta f'_k(v_k(n)) u_i(n) \sum_{j \in T} \bar{\delta}_j(n) \bar{w}_{jk}(n) \quad (3.29)$$

Pošto se korekcija uvek vrši na osnovu vrednosti gradijenta u narednom sloju algoritam je moguće vršiti samo počev od izlaznog pa sukcesivno ka ulaznom sloju mreže. Zbog toga je ovaj postupak treniranja dobio naziv algoritam sa povratnim prostiranjem greške.

Algoritam se obavlja iterativno pri čemu se svaka iteracija sastoji iz dve faze **direktne** i **povrate**. U direktnoj fazi vrši se sukcesivno prostiranje signala između slojeva, počev od ulaza ka izlazu, pri čemu je izlaz iz prethodnog jednak ulazu u naredni sloj. Obrada signala u okviru neurona odvija se na osnovu jednačina (3.1) i (3.3). U okviru direktne faze iterativnog koraka težinski koeficijenti neuronske mreže ostaju nepromenjeni. Cilj ove faze je isključivo određivanje greške neuronske mreže na izlaznom sloju neurona. U toku povratne faze algoritma vrši se sukcesivno podešavanje težinskih koeficijenata neuronske mreže počev od izlaznog ka ulaznom sloju neurona. Podešavanje težinskih koeficijenata izlaznog sloja vrši se direktno na osnovu signala greške (3.21). Izraz (3.24) omogućava izračunavanje gradijenta ukupne energije greške izlaznog sloja a onda se na osnovu ovog gradijenta, sukcesivno unazad, izračunavaju gradijenti (3.26) odnosno korekcije težinskih koeficijenata (3.29) skrivenih slojeva.

Skup podataka za treniranje neurona  $k$  čine parovi ulazno-izlaznih veličina

$$S_k = \left\{ \left( u_k^{(1)}, z_k^{(1)} \right), \dots, \left( u_k^{(p)}, z_k^{(p)} \right) \right\} \quad (3.30)$$

gde je sa  $p$  označen broj podataka (uređenih parova) za treniranje, sa  $u_k^{(i)}, i = 1, \dots, p$ , ulazne vrednosti a sa  $z_k^{(i)}, i = 1, \dots, p$  njima odgovarajuće izlazne odnosno željene vrednosti. Jedna ciklus treniranja u kome se iskoriste svi ulazno-izlazne veličine iz skupa  $S_k$  naziva se epoha treniranja. U pogledu upotrebe skupa podataka za treniranje (3.29) moguća su dva načina izvođenja jedne epohe. Kod prethodno opisanog načina, nazvanim **sekvencijalnim**, svaki iterativni korak odnosi se na jedan podatak za treniranje. Nakon prostiranja svake ulazne veličine u direktnoj fazi iterativnog koraka, na osnovu jednačina (3.1) i (3.3), vrši se podešavanje

težinskih koeficijenata mreže, korišćenjem jednačina (3.21) za izlazni odnosno (3.29) za skrivene slojeve neurona. Drugi način izvođenja algoritma, nazvan **serijski** (engl. batch), podrazumeva izvođenje direktne faze za celu seriju podataka za treniranje u toku jednog iterativnog koraka. Korekcija težinskih koeficijenata dobija se kao prosečna vrednost korekcije koja odgovara celokupnom skupu podataka za treniranje,  $S_k$ . Sa stanovišta izvođenja sekvencijalni način je jednostavniji jer zahteva daleko manje memorijskog prostora. Kao posledica slučajnog redosleda podataka za treniranje, vrednost težinskih koeficijenata kod sekvencijalnog načina ima oscilatorni karakter. Zbog toga je manja verovatnoća da algoritam bude zarobljen u zoni lokalnog minimuma. Sa druge strane teže je obezbediti konvergenciju i uslove za prekid algoritma.

Sa stanovišta teorije upravljanja predstavljeni sistem (Slika 3.9) je sa povratnom spregom što znači da njegova stabilnost zavisi od parametara elemenata u sprezi. Ako se ostali parametri, kao što su parametri aktivacione funkcije, fiksiraju, stabilnost algoritma na bazi korigovanja greške obezbeđuje se adekvatnim izborom koeficijenta učenja,  $\eta$ . Veći koeficijent učenja obezbeđuje brže približavanje željenim vrednostima ali povećava opasnost da vrednosti dobiju nestabilni odnosno oscilatorni karakter. Sa druge strane manje vrednosti ovog parametra obezbeđuju stabilnije kretanje vrednosti težinskih koeficijenata ali to ima za cenu sporije treniranje mreže. Koeficijent učenja ima značajan uticaj i na ostale performanse algoritma kao što je tačnost [38]. U praksi se koriste različite modifikacije algoritma sa povratnim prostiranjem greške. U nekima od njih se primenjuje različit koeficijent učenja za različite slojeve neurona ili čak za svaki neuron posebno. U cilju postizanja željenog odnosa između brzine učenja i stabilnosti ponekad se startuje sa većom vrednošću koeficijenta učenja nakon čega se ona sukcesivno redukuje kako se algoritam približava željenim vrednostima izlaza. Pojedine varijante algoritma podrazumevaju podešavanje svih težinskih koeficijenata mreže a kod nekih je deo težinskih koeficijenata fiksiran a ostali se normalno podešavaju u skladu sa opisanom procedurom.

Ne postoji opšte pravilo za prekid algoritma već se ono bira u skladu sa odgovarajućom primenom. Ciljne vrednosti težinskih koeficijenata treba da obezbede minimalnu vrednost totalne greške (3.23). Gradijent totalne greške  $\delta(n)$  po vektoru težinskih koeficijenata  $W(n)$  u

tački minimuma jednak je nuli. Prema tome jedan od kriterijuma za prekid algoritma može biti da Euklidska norma gradijenta totalne greške padne ispod granične vrednosti

$$\|\delta(n)\|_2 < \varepsilon \quad (3.31)$$

Nedostatak je ovog kriterijuma je što zahteva izračunavanje gradijenta totalne greške  $\delta(n)$ . Pošto totalna energija greške treba da bude stacionarna u tački minimum, kriterijum za prekid algoritma može biti da razlika između njene vrednosti u dve uzastopne iteracije bude manja od granične vrednosti

$$|E(n) - E(n - 1)| < \varepsilon \quad (3.32)$$

U cilju sprečavanja da zaglavi u beskonačnom ciklusu, usled nemogućnosti da ispuni neki od prethodna dva kriterijuma, uvode se dodatni kriterijumi za prekid algoritma. To može biti maksimalno dozvoljeno vreme za izvođenje algoritma ili maksimalan dozvoljeni broj epoha posle čega se algoritam prekida bez obzira što nisu ispunjeni kvalitativni kriterijumi.

### 3.3.1.2. Algoritam sa povratnim prostiranjem greške i inercijom

Da bi se povećala brzina učenja a istovremeno izbegla opasnost od nestabilnosti delta pravilu (3.21) odnosno (3.25) dodaje se faktor inercije. Ovaj faktor proporcionalan je promeni istog težinskog koeficijenta u prethodnom koraku algoritma

$$\Delta w_{ki}(n) = \eta \delta_k(n) u_i(n) + \beta \Delta w_{ki}(n - 1) \quad (3.33)$$

gde je  $\beta$  pozitivna konstanta koja se naziva **faktor inercije** (engl. momentum constant). Dobijena diferencna jednačina prvog reda (3.33) naziva se uopšteno delta pravilo. Sukcesivnom primenom ovog pravila od početka algoritma,  $t = 0$ , zaključno sa poslednjim korakom,  $t = n$ , dobija se rešenje u vidu vremenskog reda

$$\Delta w_{ki}(n) = \eta \sum_{t=0}^n \beta^{n-t} \delta_k(t) u_i(t) \quad (3.34)$$

Uslov da bi prethodni red bio konvergentan je da apsolutna vrednost koeficijenta inercije bude u intervalu  $|\beta| \in [0,1]$  odnosno vrednost koeficijenta inercije u intervalu  $\beta \in [-1,1]$ . U praksi se

koriste isključivo pozitivne vrednosti koeficijenta inercije  $\beta \in [0,1]$ . Ako su uzastopne vrednosti gradijenta pozitivne, korekcija (3.33) težinskog koeficijenta ima velike vrednosti odnosno član inercije teži da ubrza približavanje željenim vrednostima. U obrnutom slučaju, ako vrednost gradijenta osciluje u znaku, amplituda korekcije opada odnosno inercija ima stabilizirajući efekat na vrednost težinskih koeficijenata. Osim toga, uvođenje člana inercije sprečava zaglavljivanje algoritma u zoni plitkog lokalnog minimum totalne energije greške [39].

### 3.3.1.3. Algoritam sa povratim prostiranjem greške kroz vreme

Algoritam sa povratnim prostiranjem greške nije moguće direktno primeniti za treniranje rekurentnih neuronskih mreža. Na sreću, svaku rekurentnu neuronsku mrežu moguće je transformisati u ekvivalentnu mrežu sa direktnim prostiranjem signala nakon čega se algoritam sa povratnim prostiranjem greške može normalno primeniti za njeno treniranje. Na slici 3.10 a predstavljena je rekurentna neuronska mreža sa jednim skrivenim slojem neurona, označenim sa S, i izlaznim slojem označenim sa T. Skriveni sloj neurona nalazi se u povratnoj sprezi sa jediničnim operatorom kašnjenja pa se dinamika modela može predstaviti parom nelinearnom matričnih jednačinama

$$x(n) = f_S \left( W_S(x(n) + x(n-1)) \right) \quad (3.35)$$

$$y(n) = f_T(W_T x(n)) \quad (3.36)$$

gde su  $x(n)$  vektor izlaza skrivenog sloja u momentu  $t = n$ ,  $f_S$  aktivaciona funkcija skrivenog sloja,  $f_T$  aktivaciona funkcija izlaznog sloja,  $W_S$  matrica težinskih koeficijenata skrivenog i  $W_T$  matrica težinskih koeficijenata izlaznog sloja neurona. Na osnovu (3.35), trenutna vrednost izlaza iz skrivenog sloja,  $x(n)$ , zavisi od trenutnog ulaza,  $u(n)$ , i izlaza iz istog sloja u prethodnom trenutku,  $x(n-1)$ . Ako povratnu spregu zamenimo istom mrežom, sa stanjem u prethodnom trenutku,  $t = n-1$ , izlaz se neće promeniti. Ponavljanjem iste procedure počev od trenutnog stanja,  $t = n$ , pa unazad sve do početnog trenutka,  $t = 0$ , rekurentna mreža se potpuno transformiše u mrežu sa direktnim prostiranjem signala, (Slika 3.10 b), pogodnu za primenu algoritma sa povratnim prostiranjem greške. Čitav postupak razvijanja rekurentne mreže kroz vreme i primene algoritma sa povratnim prostiranjem greške naziva se povratno prostiranje greške kroz vreme.

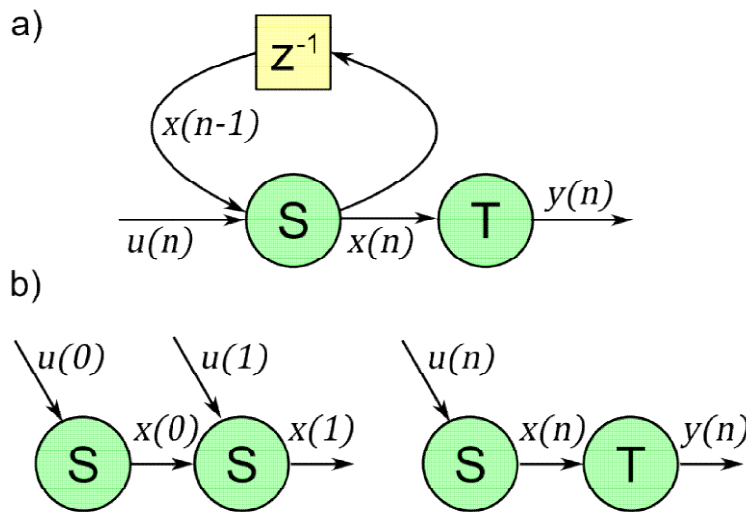
Svaki od iterativnih koraka ovog algoritma izvodi se u 3 faze. U prvoj, direktnoj fazi, vrši se prostiranje signala od ulaza ka izlazu, počev od momenta startovanja mreže,  $t = 0$ , pa sve do momenta opservacije,  $t = n$ . U ovu svrhu koriste se jednačine (3.35) i (3.36). U drugoj fazi vrši se povratno prostiranje greške počev od izlaznog ka ulaznom sloju. Ova greška se u praksi prenosi propagiranjem gradijenta funkcije totalne greške,  $\delta_k(n)$ , na osnovu jednačine (3.24) za izlazni odnosno

$$\delta_k(n) = f'_k(v_k(n)) \sum_{j \in T} \bar{\delta}_j(n) \bar{w}_{jk}(n) \quad (3.37)$$

za skriveni sloj neurona. Jednačina (3.37) izvedena je na osnovu jednačina (3.25) i (3.33) što znači da su i ovde  $\bar{\delta}_j(n)$  i  $\bar{w}_{jk}(n)$  parametri prethodnog sloja neurona. Pojačanje signala koji potiče iz neurona prethodnog sloja pod rednim brojem  $k$  i stiže na ulaz  $i$  vrši se na osnovu izraza

$$w_{ki}(n+1) = w_{ki}(n) + \eta \sum_{t=0}^n \underline{\delta}_k(t) x_i(t) \quad (3.38)$$

gde je sa  $\eta$  označen koeficijent učenja, sa  $x(t)$  predmetni signal a sa  $\underline{\delta}_k(t)$  njegov gradijent totalne greške [40]. U skladu sa (3.24) donja crta gradijenta  $\underline{\delta}_k(t)$  označava da se zapravo radi o gradijentu totalne greške izlaza iz neurona koji pripada prethodnom sloju neurona.



Slika 3.10 Razvijanje neuronske mreže u vremenu: a) rekurentna mreža, b) razvijena mreža



#### 3.3.1.4. Skraćeni algoritam sa povratnim prostiranjem greške kroz vreme

Korišćenje algoritma sa povratnim prostiranjem greške kroz vreme nije praktično za on line učenje, kada broj kopija u razvijenoj mreži (slika 3.10) postaje preveliki. Osim što računski zahtevi postaju enormni pojave gubljenja gradijenta [41] čini dublje propagiranje greške beznačajnim. Da bi se ovo izbeglo vrši se ograničavanje dubine razvijanja mreže kroz vreme na tačno određeni broj koraka,  $d$ . Mreža se razvija počev od trenutka opservacije,  $t = n$ , pa unazad kroz vreme do momenta  $t = n - d$ . Uticaj stanja i ulaza koji prethode ovom vremenskom intervalu se zanemaruje.

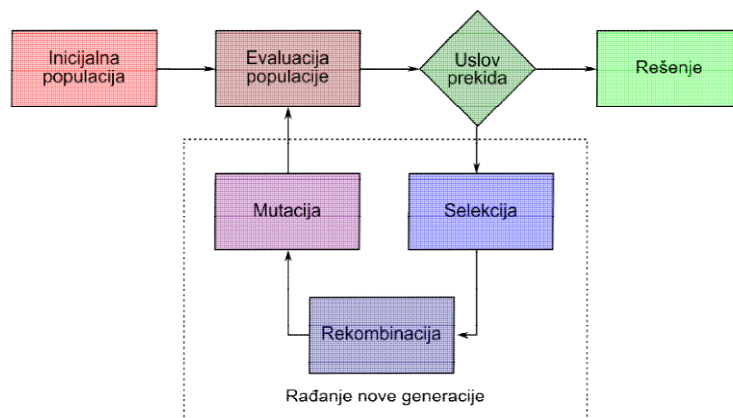
## 4. EVOLUCIONI ALGORITMI

Evolucioni algoritmi su najopštiji metodi za rešavanje problema optimizacije (engl. general problem solvers) jer mogu biti primenjeni u najrazličitijim situacijama. U odsustvu adekvatnog rešenja, baziranom na znanju iz predmetnog domena, gotovo uvek se kao alternativa može primeniti evolucionni račun. Zbog toga se za njega kaže da je nezavistan od domena (engl. domain independent). Evolucionni račun se bazira na međusobno isprepletanim procesima **istraživanja** i **eksploatacije** [42] prostora pretraživanja. Prvi predstavlja potragu za novim, potpuno neistraženim, regionima dok je drugi pretraživanje okoline trenutnih rešenja. Da bi evolucionni račun bio uspešan mora da poseduje dobar balans između ove dve aktivnosti.

### 4.1. Komponente evolucionih algoritama

U osnovi evolucionog algoritma leži generisanje velikog broja nasumičnih rešenja. Nasumična rešenja se pojedinačno nazivaju **jedinkama** a skupa **populacijom**. Algoritam se u opštem slučaju odvija prema blok dijagramu prikazanom na slici 4.1. Inicijalna populacija jedinki generiše se na slučajan način. Jedinke se rangiraju na osnovu odgovarajuće funkcije koja predstavlja matematičku formulaciju postavljenog cilja. Među populacijom se biraju samo najbolje jedinke, sa najpovoljnijom vrednošću funkcije stanja, za generisanje potomstva odnosno nove **generacije** rešenja. Procedura izbora najboljih jedinki, po analognom procesu kod evolucije živih bića, naziva se **selekcija**. Na osnovu izabranih jedinki, primenom genetskih operatora varijacije dobija se nova populacija odnosno sledeća **generacija** jedinki. Operator varijacije nad pojedinačnim jedinkama naziva se **unarna varijacija** ili **mutacija**. Varijacija koja se obavlja nad dve jedinke naziva se **binarna varijacija** ili **rekombinacija** (engl. crossover) [43]. U nekim algoritmima primenjuje se rekombinacija nad više od dve jedinke uprkos tome što analogija ne postoji kod prirodnog razmnožavanja. Stvaranje novih generacija obavlja se iterativno, ponavljanjem operatora selekcije, rekombinacije i mutacije. Rekombinacija i mutacija kao operatori varijacije obavljaju funkciju istraživanja dok selekcija obavlja funkciju eksploatacije prostora pretraživanja. Algoritam se prekida kada neka od jedinki ne postane optimalno rešenje problema na osnovu unapred zadatih kriterijuma za prekid algoritma. Optimalno rešenje može biti **globalno**, ako je najbolje u čitavom prostoru stanja, ili **lokalno**, ako je najbolje samo u određenom regionu prostora stanja. Zahvaljujući probabilističkoj prirodi,

evolucioni algoritam ima sposobnost da skače iz zone jednog u zonu drugog lokalnog minimuma čime se povećava verovatnoća nalaženja globalnog rešenja [44].



Slika 4.1 Blok dijagram evolucionog algoritma

#### 4.1.1. Reprezentacija jedinki

Reprezentacija ili kodiranje jedinki podrazumeva uspostavljanje jednoznačne relacija između predstave jedinke u originalnom kontekstu, koja se naziva **fenotip**, i njene slike u kontekstu izvođenja evolucionog algoritma, odnosno **genotipa**. Reprezentacija je ključna za izvođenje genetskog algoritma jer od nje zavisi dalji tok algoritma odnosno princip funkcionisanja genetskih operatora. Uprkos brojnim pokušajima da se formuliše, nije poznato opšte pravilo za reprezentaciju problema koje garantovano dovodi do efikasnog izvođenja evolucionog algoritma. Osim u standardnim slučajevima, nalaženje adekvatne reprezentacije predstavlja kreativni čin koji se osim sopstvenih i tuđih iskustava oslanja na intuiciju i maštu.

**Niz bitova** je najčešću i najjednostavniji metod reprezentacije jedinki sa stanovišta izvođenja algoritma. U osnovnom obliku predstavlja niz logičkih vrednosti ali takođe može predstavljati celobrojne vrednosti zapisane u formatu sa binarnom osnovom. U drugom slučaju javlja se problem da različiti geni odnosno različite pozicije u nizu bitova imaju različitu težinu. Problem se uspešno prevazilazi primenom Gray-evog (Frank Gray, 1887-1969) kodiranja kod koga se dve uzastopne binarne vrednosti uvek razlikuju samo za jedan bit. Ovim je obezbeđeno da promena jednog bita, kao rezultat mutacije na primer, ne dovodi do dramatične promene vrednosti funkcije stanja.

**String** ili niz alfanumeričkih vrednosti se, takođe, često koristi za reprezentaciju jedinki. Ovaj način reprezentacije poseduje gotovo sve pozitivne karakteristike kao i prethodni.

**Celobrojni vektor** se koristi za rešavanje funkcija sa celobrojnim promenljivama. Razlikuju se slučajevi kad je lista mogućih vrednosti gena neograničena, odnosno može biti bilo koja celobrojna vrednost, ili kada je ograničena na određeni broj vrednosti. Celobrojni vektor pokazuje svoje prednosti kada je lista vrednosti gena neograničena dok se za ograničeni najčešće primenjuje binarna reprezentacija.

**Realni vektor** primenjuje se nalaženja optimuma funkcije realnih promenljivih. Ako funkcija ima analitički oblik, posebno ako je diferencijabilna, bolje je primeniti analitičke metode za optimizaciju. Kada to nije slučaj evolucionim algoritam sa reprezentacijom u vidu realnih brojeva pruža alternativu.

**Permutacije** su pogodne za reprezentaciju problema koji imaju za cilj pronalaženje optimalnog poretka. Kod jedne varijante ovakvih problema bitan je samo poredak elemenata u permutaciji. Primer je problem dodeljivanja poslova odgovarajućim operatorima sa ciljem minimizacije ukupnih troškova izvođenja posla. Bitno je da elementi (posao-operator) budu optimalno upareni ali njihov redosled nije od značaja. Drugu kategoriju čine problemi kod kojih je bitno susedstvo između elemenata, kao što je nalaženje optimalnog puta između skupa tačaka.

#### 4.1.2. Inicijalizacija

Proces inicijalizacije podrazumeva generisanje početne populacije jedinki. Ove jedinke predstavljaju polazne tačke pretraživanja koje, kroz iterativu proceduru evolucionog algoritma, treba da evoluiraju u smeru povećanja totalne vrednosti funkcije stanja. Inicijalizacija se najčešće vrši nasumičnim izborom pozicija iz domena prostora pretraživanja. U izvesnim slučajevima, kada je na raspolaganju adekvatno znanje iz domena problema ili neka od jedinki sa visokom vrednošću funkcije stanja, korisno je primeniti način inicijalizacije koji će obezbediti veću totalnu vrednost funkcije stanja inicijalne populacije. Na ovaj način smanjuje se broj generacija potrebnih za dostizanje konačnog rešenja. Pri tome treba biti oprezan da ne budu svi elementi inicijalne populacije u povoljnim zonama prostora pretraživanja kako bi se održao diverzitet inicijalne populacije i time smanjila mogućnost od prerane konvergencije algoritma ka tački lokalnog minimuma.

### 4.1.3. Evaluacija

Evaluacija populacije predstavlja dodeljivanje jedinkama odgovarajuće mere kvaliteta sa ciljem usmeravanje algoritma ka nalaženju rešenja kao i uspostavljanja kriterijuma za njegov prekid. Evaluacija se obavlja primenom **funkcije evaluacije** koja još poznata pod nazivima **funkcija stanja, funkcija zdravlja ili funkcija cilja**, poslednje samo ako se odnosi na rešavanje problema optimizacije. Nema opšte prihvaćenog pravila za izbor adekvatne funkcije evaluacije. Ona treba da odslikava zahteve koji su postavljeni pred algoritmom odnosno da izmeri koliko je on udaljen od postavljenog cilja. Ipak, neke funkcije su se u praksi ispostavile kao karakteristične za određene klase problema. Na primer, za probleme klasifikacije sa nadgledanim učenjem, karakteristične su različite mere odstupanja kao što su Euklidsko rastojanje ili Menhetn (engl. Manhattan) rastojanje. Manhattan rastojanje predstavlja sumu dužina horizontalne i vertikalne projekcije rastojanja između dve tačke. Kod problema optimizacije kao funkcija evaluacije često se koriste bazične funkcije kao što je suma parametara vezanih za problem.

### 4.1.4. Modeli evolucije

Pod modelom evolucije podrazumeva se način selekcije roditelja zajedno sa načinom vraćanja potomaka u populaciju.

**Metoda jednog turnira** se izvodi podelom čitave populacije na male grupe . Na osnovu dve najbolje jedinke, primenom operatora varijacije, dobijaju se dva potomka koji u populaciji zamenjuju dve najgore jedinke. Prednost metode jednog turnira je da svi članovi grupe osim poslednja dva preživljavaju proces selekcije što garantuje povećanje opšteg nivoa funkcije stanja. Druga prednost je da najbolja dva člana grupe imaju samo po jednog potomka bezobzira kolika je njihova vrednost funkcije stanja. Ovo garantuje da neće doći do prevremenog gubitka diverziteta (raznolikosti) među jedinkama koja može uzrokovati da algoritam bude zarobljen u zoni lokalnog optimuma. Algoritmi kod kojih je garantovano preživljavanje jedinki sa najboljom vrednošću funkcije stanja nazivaju se **elitistički**. Kod metode jednog turnira garantovano je preživljavanje samo dve najbolje jedinke u okviru čitave populacije. Ostale jedinke preživljavaju samo ako imaju sreću da u njihovoj grupi ne bude dve jedinke sa povoljnijom vrednošću funkcije stanja.

**Metoda duplog turnira** veličine  $n$  podrazumeva biranje najboljeg roditelja među  $n$  nasumično izabranih jedinki. Drugi roditelj se bira među drugih  $n$  nasumično izabranih jedinki. Jedna varijanta ovog algoritma (**sa zamenom**) dozvoljava da isti roditelj bude izabran dva puta dok druga varijanta (**bez zamene**) isključuje prvog roditelja iz procedure izbora drugog.

**Selekcija metodom ruleta** bira roditelje sa verovatnoćom proporcionalnom njihovoj funkciji stanja. Ako jedinka  $i$  ima vrednost funkcije stanja  $f_i$  verovatnoća njenog izbora za roditelja je  $f_i/F$  gde je  $F$  totalna vrednost funkcije stanja elemenata populacije.

**Selekcija na osnovu ranga** funkcioniše na sličnom principu kao i selekcija metodom ruleta. Na osnovu vrednosti funkcije stanja jedinke se rangiraju tako da najgora jedinka dobije rang 1 a ostale redom za jedan rang veći. Verovatnoća da jedinka  $i$  sa rangom  $r_i$  bude izabrana za roditelja je  $r_i/R$  gde je  $R$  suma rangova svih elemenata populacije.

Da bi, nakon kreiranja potomaka populacija zadržala isti broj jedinki potrebno je da isti broj potomaka bude uklonjen iz nje. U pogledu uklanjanja roditelja takođe postoje različiti pristupi. Ako potomci zamenjuju roditelje koji su izabrani potpuno nasumično radi se o **nasumičnoj zameni**. Kod **zamene metodom ruleta** verovatnoća da roditelj bude zamenjen obrnuto je proporcionalna vrednosti njegove funkcije stanja. Na taj način roditelji sa najgorim karakteristikama imaju najveću verovatnoću da budu zamenjeni. **Zamena metodom ranga** izvodi se rangiranjem jedinki po redosledu, obrnutom od selekcije na osnovu ranga, tako da najbolja jedinka dobije rang 1 ostale redom za jedan rang veći. Verovatnoća da roditelj bude uklonjen proporcionalna je tako dobijenom rangom. Ako se menjaju roditelji sa apsolutno najgorom vrednošću funkcije stanja radi se o **apsolutnoj zameni** na osnovu vrednosti funkcije stanja. **Nasumična zamena elite** podrazumeva upoređivanje potomka sa nasumično izabranom jedinkom populacije i zamenu samo u slučaju da od nje ima bolju vrednost funkcije stanja. Zamena kod koje potomci direktno menjaju svoje roditelje u populaciji naziva se **zamena lokalne elite**. Kod ove metode zamena može biti безусловna ili se biraju samo oni potomci koji imaju bolju funkciju stanja od roditelja.

Osim načina, potrebno je odrediti i broj jedinki koji treba da bude zamenjen u okviru jednog ciklusa stvaranja nove generacije. Primenuju se sve varijante zamene počev od **stacionarnog algoritma**, kod koga se svaki pojedinačni čin zamene roditelja potomcima, čak i

kada zamena ne uspe, smatra generacijom pa do **generacijskog algoritma** kod koga se u okviru jednog ciklusa menja čitava populacija.

#### 4.1.5. Rekombinacija

Operator rekombinacije generiše jednog ili više potomaka na osnovu najčešće dva ili više roditelja. Rekombinacija se obavlja ugradnjom genetske strukture roditelja u genetsku strukturu potomaka sa ciljem očuvanja i unapređenja povoljnih karakteristika. Ako rekombinacija dva identična roditelja daje potomke identične ovim roditeljima operator se smatra konzervativnim. Izbor operatora rekombinacije uslovljen je prirodom problema optimizacije odnosno načinom reprezentacije.

**Rekombinacija sa jednom tačkom** je bazični metod rekombinacije koji simulira analogni proces kod živih bića. Lokacija tačke rekombinacije bira se nasumično nakon čega se u toj tački dele hromozomi oba roditelja. Jedan od potomaka dobija prednji deo hromozoma od jednog a drugi deo hromozoma od drugog roditelja. Drugi potomak dobija ostatak gena od oba roditelja. Osnovni nedostatak ovog pristupa je da se tačke sa bliskim lokacijama u hromozomu odvajaju sa manjom verovatnoćom od onih koje su udaljene.

**Rekombinacija sa dve (više) tačaka** kreirana je sa ciljem ublažavanja prethodno opisanog nedostatka. Tačke rekombinacije biraju se nasumično nakon čega se u odnosu na njih dele hromozomi i razmenjuju njihovi delovi. Potomak dobija redom neparne delove od jednog a parne od drugog roditelja. Kod nekih izvođenja operatora rekombinacije broj tačaka se, prilikom svakog izvođenja, bira nasumično.

**Uniformna rekombinacija** postiže se tako što se svaki gen tretira kao poseban deo hromozoma. Prilikom formiranja gena potomka za svaki od gena se nasumično odlučuje da li će biti nasleđen od jednog ili od drugog roditelja. Drugi potomak dobija komplementarni ostatak gena. Uniformna rekombinacija u opštem slučaju obezbeđuje dostizanje optimalnog rešenja u manjem broju generacija od prethodnog opisanih metoda. Nedostatak ovog pristupa leži u velikim računskih zahtevima zbog velikog broja potrebnih slučajnih promenljivih.

**Adaptivni operator rekombinacije** pokušava da odgovori činjenici da različitim fazama izvođenja algoritma mogu da odgovaraju različiti operatori rekombinacije. Da bi ovo bilo

postignuto svakoj jedinki, pored sopstvenog gena, pridružuje se **obrazac rekombinacije**. To je niz binarnih cifara, iste dužine kao i dužina gena, koji određuje koji geni će biti nasleđeni od jednog a koji od drugog roditelja. Rekombinacija se obavlja na osnovu obrasca rekombinacije jednog od roditelja. Nakon rekombinacije izvornih gena, među roditeljima se vrši varijacija samih obrazaca rekombinacije. U tu svrhu, sa ciljem dobijanja obrazaca rekombinacije potomaka, koriste se posebni operatori rekombinacije odnosno mutacije. Koegzistirajući, na ovaj način, sa sopstvenim jedinkama obrasci za rekombinaciju tragaju za operatorom rekombinacije koji najviše odgovara trenutnoj fazi rešavanja problema.

**Aritmetička rekombinacija** se primenjuje za rekombinaciju realnih vektora. Na osnovu realnih vrednosti odgovarajućih gena roditelja,  $x$  i  $y$ , dobija se vrednost gena potomka po formuli  $z = \alpha x + (1 - \alpha)y$  gde je parametar  $\alpha$  iz intervala  $[0, 1]$ . Nedostatak ove metode je da se u svakoj populaciji smanjuje opseg vrednosti gena.

#### 4.1.6. Mutacija

Mutacija je unarni operator varijacije koji obavlja relativno malu promenu jedinke. Najčešće se vrši zamenom jednog gena drugim, koji je izabran nasumično. Time se pospešuje lokalna pretraga i postepeno otvara mogućnosti za nove puteve istraživanja prostora pretraživanja. Izbor operatora mutacije takođe je uslovljen reprezentacijom jer u zavisnosti od nje mogu biti definisani različiti operatori za vršenje malih izmena jedinki. Neki od najčešćih operatora mutacije su [43]:

- **Mutacija jedne tačke** – zamena gena na nasumično izabranoj poziciji
- **Mutacija više tačaka** – zamena gena na više nasumično izabranih pozicija
- **Probabilistička mutacija faktora  $\alpha$**  – zamena svakog od gena pojedinačno sa verovatnoćom izvođenja  $\alpha$
- **Lamarck-ova mutacija dubine  $k$**  – obavlja sve mutacije sa  $k$  i manje tačaka i primenjuje onu koja rezultuje najvećom vrednošću funkcije stanja
- **Uniformna mutacija nad realnom reprezentacijom** gena – dodavanje, nasumično izabranom genu, slučajnog broja sa uniformnom raspodelom iz intervala  $[-\epsilon, \epsilon]$ ,  $\epsilon \in \mathbb{R}$
- **Gausova mutacija nad realnom reprezentacijom** gena - dodavanje, nasumično izabranom genu, slučajnog broja sa normalnom raspodelom



Složene reprezentacije mogu zahtevati više različitih operatora mutacije. Za svaki od njih definiše se verovatnoća da bude primenjen koja se naziva **stopa mutacije** (engl. mutation rate) operatora.

#### 4.1.7. Prekid algoritma

Ako problem ima poznato optimalno rešenje uslov za prekid algoritma je dostizanje ovog rešenja uz odgovarajuću toleranciju  $\epsilon > 0$ . Međutim, evolucioni algoritam ima stohastičku prirodu što znači da nema garancija da će prethodni uslov biti ispunjen. Zbog toga se kriterijumi za prekid dopunjuju dodatnim uslovima koji će obezbediti prekid algoritma. Neke od opcija su:

- Minimalno unapređenje funkcije stanja u toku jedne generacije
- Minimalni diverzitet populacije
- Maksimalni dozvoljeni broj generacija
- Maksimalno vreme izvođenja algoritma
- Maksimalni broj izračunavanja funkcije stanja

## 4.2. Vrste evolucionih algoritama

U pogledu upotrebe genetskih operatora razlikuju se četiri vrste evolucionih algoritama: **genetski algoritmi**, **evolucionne strategije**, **evoluciono programiranje** i **genetsko programiranje**.

### 4.2.1. Genetski algoritam

Genetski algoritam je najviše primenjena varijanta evolucionih algoritma pa se često primenjuje kao sinonim za ovaj tip računске inteligencije. Strogo definisano, genetski algoritam podrazumeva diskretnu reprezentaciju jedinki, u vidu binarnih ili alfanumeričkih nizova. U skladu sa reprezentacijom na raspolaganju je veliki broj genetskih operatora (Poglavlje 4.1). Operatori namenjeni diskretnoj reprezentaciji su generalno jednostavniji u odnosu na ostale. Karakteristika genetskih algoritama je da su mutacije relativno retke u odnosu na rekombinacije ali se ne smeju izostaviti kako bi se očuvao diverzitet populacije. Ako se parametri varijacije, kao što su nivo mutacija i obrazac rekombinacije, uključe u evolutivni proces genetski algoritam postaje adaptivan. Ovakav algoritam reaguje na različite faze sopstvenog izvođenja čime

obezbeđuje bržu i izvesniju konvergenciju. Adaptivni genetski algoritam se u literature može sresti pod nazivom **hibridni** ili **meta-genetski** algoritam.

#### 4.2.2. Evolucione strategije

Za razliku od genetskih algoritama, kod evolucionih strategija jedinka je predstavljene vektorom položaja u prostoru pretraživanja sa realnim koordinatama. S obzirom da vektori sadrže analogne veličine ovaj metod pogodan je za rešavanje numeričkih problema optimizacije. Operatori varijacije nad realnim vektorima daleko su kompleksniji od operatora namenjenih stringovima. Intenzitet njihovog dejstva kontroliše se podešavanjem velikog broja parametra. Pošto je ove parametre teško definisati unapred oni su najčešće sastavni deo genetskog materijala jedinki i čitavog evolutivnog procesa. Podešavanje parametara varijacije obavlja se primenom posebnih operatora. Koraci algoritma slični su onima koji se izvode kod genetskih algoritama ali sa izvesnim modifikacijama. Prvo, za rađanje nove generacije najčešće se upotrebljava čitava populacija roditelja pri čemu se jedan roditelj može upotrebiti više puta. Ređe, kada se roditelji biraju na osnovu vrednosti funkcije stanja, primenjuju se dva pristupa: **kona selekcija** i **plus selekcija** [45]. Kod prve se za formiranje nove generacije biraju samo predstavnici poslednje generacije dok se kod druge koriste dve poslednje generacije. Za razliku od genetskih algoritama, kod kojih potomak ima dva roditelja, kod evolucionih strategija potomak može imati više od jednog roditelja.

#### 4.2.3. Evoluciono programiranje

Evoluciono programiranje nastalo je modeliranjem teorije evolucije ali na drugačijem nivou od prethodnih algoritama evolucionog računa. U skladu sa mehanizmom funkcionisanja algoritma ovde je jedinku adekvatnije nazvati vrstom a populaciju biocenozom (zajednicom) različitih vrsta [17]. Kako između različitih vrsta nema razmnožavanja kod evolucionog programiranja nema operatora rekombinacije. Evolucija jedinki (vrsta) kroz generacije vrši se isključivo na osnovu mehanizma mutacije. Algoritam počinje generisanjem inicijalne biocenoze različitih vrsta analogno inicijalnoj populaciji jedinki kod prethodnih metoda evolucionog računa. Generisanje naredne generacije obavlja se mutacijama nad svim članovima inicijalne generacije. Mutacije se obavljaju na slučajan način u skladu sa nekim od zakona distribucije tako da se radikalne promene dešavaju sa malom a minorne promene sa velikom verovatnoćom. U

sledećoj fazi algoritma vrši se evaluacija jedinki na osnovu vrednosti funkcije stanja. Evaluacija je osnova za proces selekcije koji se obavlja sa ciljem biranja roditeljske generacije. I ovde se selekcija može vršiti primenom različitih strategija [46]. Dalje se algoritam ciklično izvršava do ispunjenja uslova za prekidanje isto kao kod prethodnih metoda evolucionog računa. Za evoluciono programiranje je karakteristično da nije ograničeno formom reprezentacije rešenja. U skladu sa prirodom problema moguće je primeniti različite načine reprezentacije. Zbog toga je evoluciono programiranje pogodno za rešavanje širokog spektra praktičnih problema. Za razliku od prethodnih metoda evolucionog računa gde je rezultat algoritma rešenje problema kod evolucionog programiranja je cilj kreiranje sistema odnosno automata koji je sposoban za rešavanja određene klase problema [17].

#### 4.2.4. Genetsko programiranje

Genetsko programiranje koristi iste operatore kao i genetski algoritam ali je predmet obrade drugačiji. Za razliku od genetskog algoritma kod koga je cilj rešenje problema kod genetskog programiranja je cilj sinteza programa za rešavanje zadate klase problema [17]. Umesto generisanja populacije potencijalnih rešenja ovde se generiše populacija kompjuterskih programa za rešavanje zadatog problema. Algoritam se vremenom pokazao efikasnim u različitim tehničkim primenama kao što su generisanje električnih, pneumatskih i hidrauličnih kola, sistema automatskog upravljanja i sl. U zavisnosti od oblasti primene za generisanje potencijalnog rešenja koriste se različite komponente. Ako se radi o programiranju to su programske instrukcije, matematičke funkcije, aritmetičke i logičke operacije. Kod generisanja različitih kola i sistema koriste se odgovarajuće sastavne komponente. Adekvatnost rešenja ocenjuje se testiranjem svakog od programskih rešenja na osnovu odgovarajuće funkcije stanja. Realizacija operatora rekombinacije, mutacije i selekcije mora biti kompatibilna sa domenom primene. Za kreiranje inicijalne populacije rešenja takođe moraju biti razvijene posebne strategije [47].

## 5. PREPOZNAVANJE AKUSTIČNOG IZVORA

Akustično prepoznavanje je dobijanje relevantnih informacija o objektu na osnovu zvučnog signala koji on emituje [48]. Pojam prepoznavanje potiče od činjenice da procedura podrazumeva uspostavljanje veze između trenutne akustične impresije i sličnog iskustva obrađivanog u prošlosti. Svaki savremeni čovek je iskusio dijagnostičku proceduru **auskultacije**, jednu od najvažnijih u okviru anamneze i kliničkog pregleda srca, pluća i krvnih sudova. Procedura se obavlja slušanjem zvukova koje proizvode unutrašnji organi pomoću jednostavnog akustičnog instrumenta zvanog **stetoskop**. Čak i događaji koji ne proizvode zvuk direktno mogu biti ispitivani na osnovu akustičnih događaja ukoliko se između ovih uspostavi uzročno posledična veza. Detekcija kašlja, kao akustičnog signala, primenjena je kao biomarker za ispitivanje zagađenja vazduha [49]. Posebna mogućnost je emitovanje zvučnog signala, obično ultrazvuka, u smeru objekta posmatranja. Nakon što se odbije od objekta, eho se analizira i upoređuje sa referentnim obrascima u cilju dobijanja informacija o objektu [50]. Postupak je dobro poznat u medicini, podvodnoj i podzemnoj detekciji, detekciji deformacija u materijalu, geomorfologiji, arheologiji itd. Neke muzičke aplikacije [51,52] omogućavaju prepoznavanje instrumenata, izvođača, automatsku transkripciju zvuka, razdvajanje kanala [53], kategorizaciju i sl. Krajnji, istovremeno i najkompleksniji, problem je prepoznavanje ljudskog govora i njegov prevod u kompjuterski zapis [54,55]. Na ovaj način se obezbeđuje direktna komunikacija između čoveka i računara bez upotrebe različitih interfejsa kao što su tastatura, miš ili komandna palica.

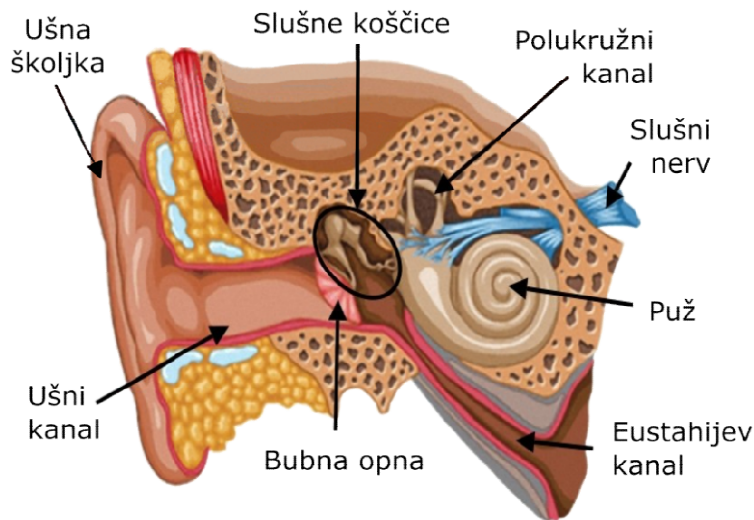
### 5.1. Mehanizam akustičnog prepoznavanja

Prepoznavanje objekata (fenomena) na osnovu akustičnih signala koje proizvodi kod živih bića se naziva slušanje. Pod slušanjem se podrazumeva **receptija** zvučnog signala kao i njegova **percepcija**.

#### 5.1.1. Receptija akustičnog talasa

Mehaničko registrovanje akustičnog signala zajedno sa njegovom transformacijom u neuronski signal naziva se receptija. Receptija se obavlja sa ciljem da se akustični signal učini prepoznatljivim za procesnu jedinicu u kojoj se obavlja faza percepcije.

Najefikasnije sredstvo za recepciju akustičnih talasa su uši, parni organ smešten na glavi kičmenjaka (sa izuzetkom riba). Na slici 5.1 predstavljen je poprečni presek ljudskog uha. **Ušna školjka** ima ulogu sprovođenja akustičnih talasa ka **ušnom kanalu**. Pri tome različite zvukove, u smislu frekvencije i smera prostiranja, pojačava sa različitim faktorom pojačanja. To znači da se već u ušnoj školjci vrši početna selekcija akustičnih parametara. Na kraju ušnog kanala talasi udaraju u **bubnu opnu** izazivajući njeno vibriranje. Uloga **Eustahijevog kanala** je da izjednači pritisak sa obe strane bubne opne. Mehaničke vibracije se od bubne opne prenose putem **slušnih košćica** (**čekića, nakovnja i uzengije**) do spiralnog koščatog organa nazvanog **puž**. Uloga puža je pretvaranje mehaničkih vibracija u nervne impulse koji putem **slušnog nerva** putuju ka centralnom nervnom sistemu. Svaki od ovih organa se u matematičkom smislu može predstaviti kao pojačivač sa izrazito nelinearnom karakteristikom koja je rezultat prilagođavanja na uslove života tokom evolucije. Drugi organizmi, posebno insekti, upražnjavaju naprednu akustičnu komunikaciju sa okolinom pomoću daleko jednostavnijih organa kao što su koža i različite izrasline na njoj (dlačice). Bez obzira na način, organi za recepciju zvuka kod živih bića obezbeđuju elektrohemijski signal na izlazu jer je to način na koji se vrši komunikacija između nervnih ćelija živih bića.

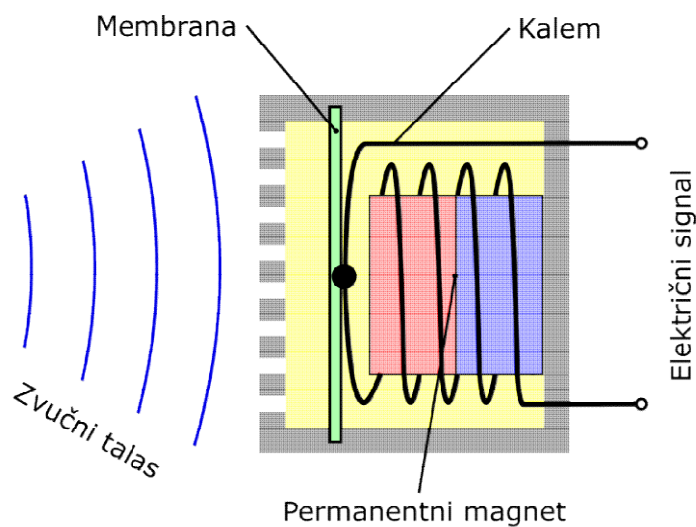


*Slika 5.1 Ljudsko uho [56]*

Različite vrste veštačkih senzora za recepciju zvuka razvijene su za različite potrebe. Tako se senzori za recepciju čujnog opsega, za čoveka između 20 Hz i 20 kHz, nazivaju **mikrofoni**. Za recepciju ultrazvučnih talasa primenjuju se **ultrazvučni senzori**, za akustične talase u

tečnostima **hidrofoni** a za registrovanje pokreta u zemljinoj kori **geofoni**. U zavisnosti od toga koja elektromagnetna pojava je iskorišćena za transformaciju mehaničkog talasa u električni signal u primeni su različite vrste akustičnih senzora kao što su **dinamički**, **kapacitivni** ili **piezoelektrični**. Ove vrste senzora razlikuju se po svojim karakteristikama (osetljivost, distorzija, karakteristika preopterećenja, frekventni odziv, šum i sl.) [57]. Bez obzira na vrstu, svi akustični senzori proizvode električni signal na izlazu zbog toga što se analiza akustičnog signala u veštačkim sistemima vrši primenom električnih kola.

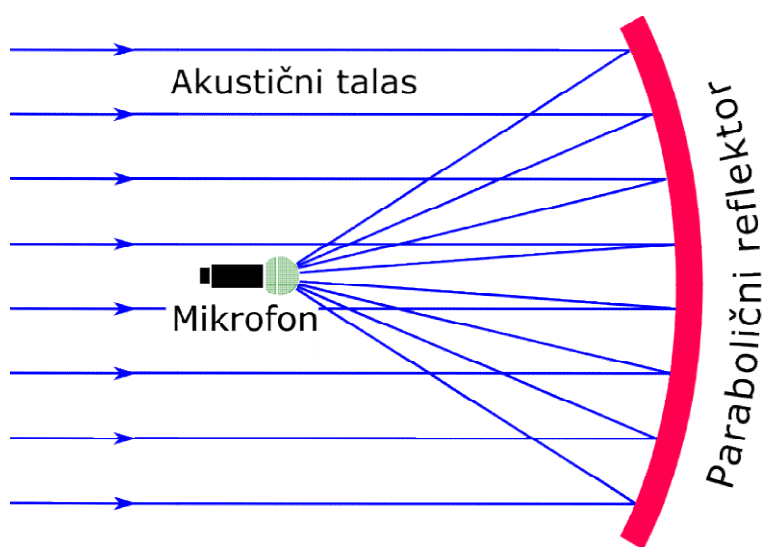
Princip rada akustičnog senzora može se objasniti na primeru dinamičkog mikrofona predstavljenog na slici 5.1. Dinamički mikroskop vrši pretvaranje zvučnog talasa u električni signal koristeći fenomen elektromagnetne indukcije. On sadrži pokretni namotaj koji je vezan za dijafragmu. Namotaj se slobodno kreće u aksijalnom pravcu unutar magnetnog polja koje potiče od permanentnog magneta. Udaranje talasa u dijafragmu uzrokuje njeno oscilovanje koje se preko čvrste veze prenosi na namotaj. Oscilovanje namotaja unutar magnetnog polja uzrokuje indukovanje električne struje čija je jačina proporcionalna intenzitetu zvučnog talasa. Dinamički mikrofoni odlikuju se visokom trajnošću i pouzdanošću. ne zahtevaju posebno električno napajanje i mogu da podnesu visoke akustične pritiske.



Slika 5.2 Dinamički mikroskop

Pre obrade obično se vrši pojačanje signala, primenom različitih električnih pojačivača, kako bi se smanjio gubitak informacija na njegovom putu kroz električne vodove. Električni

pojačivači, međutim, mogu da podignu intenzitet signala ali ne i da povećaju njegov kvalitet. U cilju povećanja kvaliteta recepcije mogu se primeniti mehanički pojačivači kao što je parabolični reflektor prikazan na slici 5.3. Ovaj reflektor funkcioniše ako se usmeri ka izvoru zvuka koji je dovoljno udaljen da se brzine akustičnog talasa u različitim tačkama mogu smatrati paralelnim. Nakon odbijanja od bilo koje tačke na površini reflektora zvuk se vraća ka istoj tački, fokusu parabole, što dovodi do njegove superpozicije. Rezultat je mehanički pojačan talas koji se registruje pomoću mikrofona smeštenog u tački fokusa. Neki od reflektora specijalne namene, od kojih su mnogi rađeni po ugledu na ljudsko uho ili uho životinja, potvrdili su svoju funkcionalnost u laboratorijskim uslovima ali još uvek nisu postali standard u praksi [58].



*Slika 5.3 Mikrofon sa paraboličnim reflektorom*

### 5.1.2. Percepcija akustičnog signala

Kod živih organizama, akustična percepcija predstavlja mentalnu interpretaciju zvučnih nadražaja iz okoline. Percepcija podrazumeva indentifikovanje akustičnih signala, njihovo sakupljanje i čuvanje kao i njihovu interpretaciju u skladu sa raspoloživim znanjem slušaoca. U određivanju osnovnih perceptualnih kvaliteta zvuka polazi se od brojnih psiholoških istraživanja koja su vršena na ovu temu. Lederman, na primer, je ispitivao primenu akustičnih signala za određivanje hrapavosti površine [59], Richards je tragao za akustičnim svojstvom koje jedinstveno karakteriše tip materijala [60], Li, Logan i Pastore su ispitivali mogućnosti prepoznavanja pola osobe na osnovu zvuka koje proizvode za vreme hoda [61], Carello,

Anderson i Kunkler-Peck su ispitivali percepciju dužine objekta na osnovu zvuka [62], Kunkler-Peck i Turvey prepoznavanje oblika na osnovu zvuka koji proizvodi pri udaru [63], Lutfi audio-detekciju šupljina [64]. Istraživanja su dala puno pozitivnih rezultata ali su otvorila još više novih pitanja.

#### 5.1.2.1. Perceptivne karakteristike zvuka

Percepcija zvuka pokazala se kao izuzetno kompleksan proces koji zavisi od puno subjektivnih faktora. Zbog toga je izrada univerzalnog modela akustične percepcije praktičnog nemoguća. Neke karakteristike zvuka su se ipak istakle kao naročito uticajne na iskustvo percepcije.

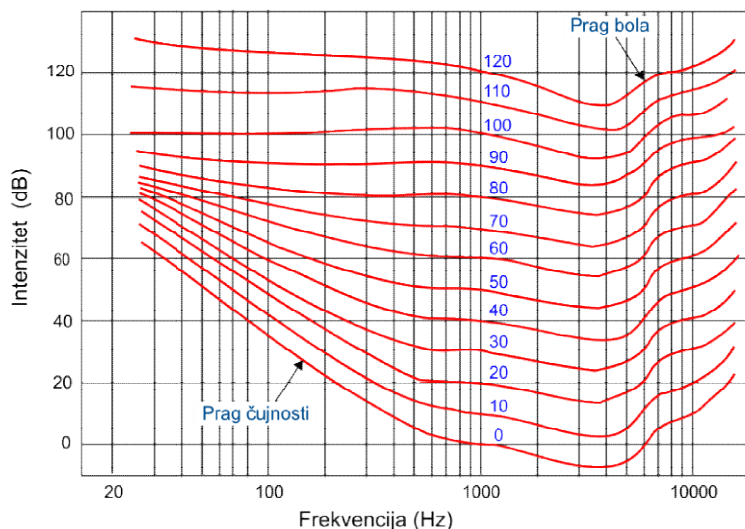
**Visina** (engl. pitch) je perceptivna karakteristika koja omogućava rangiranje zvučnih talasa po frekvenciji. Visina se ne može poistovetiti sa frekvencijom jer je prva subjektivna psihoakustična senzacija a druga fizička karakteristika zvuka pa odnos između ove dve veličine nije linearan. Na primer, najmanja promena visine tona koju ljudsko uho može da registruje zavisi od frekvencije [65]. Za niske frekvencije, ispod 500 Hz, to je 3 Hz a za visoke frekvencije, iznad 1000 Hz, najmanja primetljiva razlika je 6 Hz. Ovi rezultati dobijeni su primenom sinusnih talasa. Broj različitih visina tona koje čovek može da registruje kreće se oko 1400 u opsegu frekvencija od 16 Hz do 16 kHz. Broj nota solmizacije u istom opsegu je 120. Visina tona je karakteristika harmoničnih zvukova što znači da ona može biti registrovana kod zvukova koji imaju čistu i stabilnu frekvenciju u vremenskom periodu dovoljno dugom da može biti registrovana. Kod jednostavnih (sinusoidalnih) talasa visina tona odgovara frekvenciji dok je kod kompleksnih talasa određena frekvencijom osnovnog harmonika.

**Glasnoća** (engl. Loudness) zvuka predstavlja subjektivnu percepciju akustičnog pritiska. Glasnoća se često meša sa fizičkim merama jačine zvuka kao što su akustični pritisak, intenzitet zvuka ili snaga izvora zvuka. Međutim, pored fizičkih osobina samoga talasa na glasnoću utiču fiziološke karakteristike slušnih organa i nervnog sistema kao i psihološke karakteristike slušaoca tako da je preciznu psihofizičku skalu teško odrediti. Za potrebe industrije uspostavljene su empirijske relacije između fizičkih karakteristika zvuka i glasnoće. Analitički primer je zakon Stivena



$$\psi(I) = kI^{0,67} \quad (5.1)$$

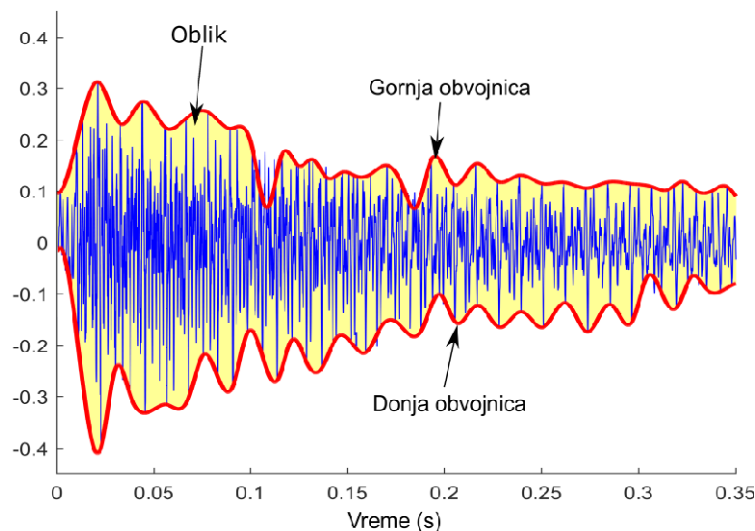
Stivens je ovaj zakon predvideo za subjektivne nadražaje uopšte samo što je za druge nadražaje dobio druge vrednosti eksponenata. Primer grafičke reprezentacije odnosa između glasnoće i intenziteta zvuka su konture jednake (konstantne) glasnoće. Ove konture predstavljaju meru intenziteta zvuka, u decibelima, duž celog frekventnog spektra za koju slušalac doživljava konstantnu glasnoću kada mu je prezentovan čisti stabilan ton. Na slici 5.4 predstavljena je Fletcher-Munson-ova varijanta kontura jednake glasnoće od **praga čujnosti** pa sve do **praga bola**. Prag čujnosti je najmanja glasnoća zvuka koju čovekovo uho može da registruje a prag bola je granica nakon koje zvuk prelazi u bol. Ovu glasnoću uzrokuju pojave kao što su simfonijski orkestar, jaka grmljavina i supersonični avion prilikom probijanja zvučnog zida.



Slika 5.4 Fletcher-Munson-ove konture jednake glasnoće [66]

**Boja** (engl. Timbre) je perceptivna karakteristika kvaliteta zvuka. Ova osobina pravi razliku između tonova proizvedenih na različitim muzičkim instrumentima makar oni bili iste visine i glasnoće. Vrhunski muzičari sposobni su da razlikuju muzičke instrumente istog tipa samo na osnovu boje zvuka. Šta više, oni mogu da menjaju boju zvuka koristeći se različitim tehnikama sviranja ili pevanja. Najbolji dokaz značaja boje na percepciju zvuka je prepoznavanje među ljudima na osnovu boje glasa. U fizičkom smislu boja zvuka određena je frekvencijom i oblikom akustičnog talasa. Pod oblikom se podrazumeva površina ograničena gornjom i donjom obvojnicom signala. Obvojnice su glatke linije koje ograničavaju (dodiruju)

ekstremne vrednosti signala. Na slici 5 prikazan je oblik akustičnog signala proizvedenog na klaviru u trajanju od 0,35 s. Oblik zvuka, a time i njegova boja, u matematičkom smislu, su određeni relativnim odnosom između dominantnih amplituda viših harmonika. Akustični talasi u obliku šuma, koji nemaju stabilne harmonike, ne pružaju percepciju boje zvuka.



*Slika 5.5 Akustični signal proizveden na klaviru*

**Trajanje** je doživljaj fizičkog trajanja dužine zvuka ali sa njim ne može biti poistovećeno. Trajanje počinje u momentu kada nervni sistem dobije informaciju da je signal primećen, a završava se u momentu kada on izgubi ciklični obrazac. Doživljaj trajanja ne mora da bude određen fizičkim trajanjem zvuka. Psihoakustična istraživanja pokazala su da neki tonovi koji se prekidaju, naročito u prisustvu buke i šuma, mogu biti pogrešno registrovani kao kontinualni. Ova pojava je, sa biološkog stanovišta, korisna jer omogućava rekonstruisanje ometenog akustičnog signala na osnovu delova koji su registrovani korektno.

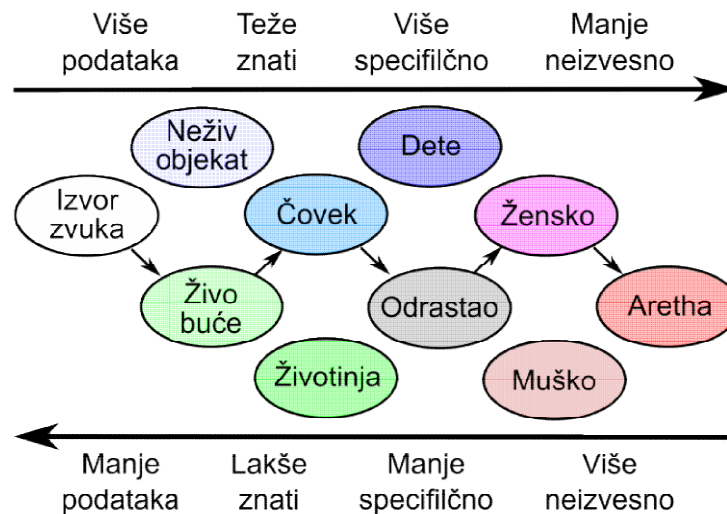
Psihoakustična istraživanja obezbedila su teoretski okvir za razvoj veštačkih sistema za percepciju zvuka. Različite osobine pokazale su se kao krucijalne za različite aplikacije prepoznavanja zvuka a neke metode baziraju se na više karakteristika. Na primer, osobine vremenskog domena kao što su **varijansa** (disperzija) **broj promena znaka** (engl. zero-crossing rate) i **koeficijent tišine**, u kombinaciji sa spektralnim osobinama, kao što su **harmonijski odnos** i **energija pod opsega** (engl. sub-band energy), se koriste za diskriminaciju između govora i muzike [67]. Svaka od navedenih veličina odnosi se na uzorak odnosno prozor (engl.

window) koji je istrgnut iz kompletnog audio signala. Koeficijent tišine predstavlja udeo tišine u odnosu na dužinu uzorka. Muzičke deonice karakterišu se manjim koeficijentom tišine od onih koje sadrže ljudski govor. Da bi se odredio harmonijski odnos uzorak signala se deli na više delova (okvira) i određuje sopstvena frekvencija svakog od njih. Harmonijski odnos je onda količnik broja okvira koji sadrže neku frekvenciju u odnosu na ukupan broj okvira. Govor ima manji harmonijski odnos od muzike jer pored samoglasnika (engl. vowels) koji imaju harmonijsku karakteristiku sadrži i samoglasnike (engl. consonants) koji imaju karakteristiku šuma. Energija podopsega se definiše kao skup energija svih frekvencije u okviru tog opsega. Varijansa energije signala govora je obično manja za niže opsege frekvencija. Za detekciju prirodnih zvukova obično se koriste osobine izvedene iz frekventnog spektra (spektar snage, spektralni centroid, spektralna iregularnost, sadržaj parnih i neparnih harmonika...). Za detekciju signala nalik šumu dominantno se koristi **cepstralna** reprezentacija signala [68]. Cepstralni koeficijenti dobijaju se tako što se koeficijenti frekventnog spektra logaritmuju a onda se nad njima izvrši inverzna Fourier-ova transformacija. U slučaju visoko-dimenzionalnih signala, previše glomaznih za manipulaciju u originalnom obliku, primenjuju se neke od tehnika redukcije dimenzije kao što je analiza glavnih (engl. principal) komponenti. Ova metoda koristi ortogonalnu transformaciju za pretvaranje skupa potencijalno koreliranih promenljivih u skup linearno nekorelisanih promenljivih nazvanih glavnim ili principijelnim komponentama [69]. Principijelne komponente se sortiraju na taj način da prva ima najveću moguću varijansu a svaka naredna mora da ima veću varijansu od preostalih i da bude ortogonalna na sve prethodne komponente. Kada su komponente ovako rangirane, signal može biti redukovan na dimenziju N, zadržavanjem prvih N principijelnih komponenti bez bojazni da je maksimalno moguća količina informacija očuvana. Opšti zaključak je da nema univerzalne osobine zvuka koja će obezbediti najbolju percepciju. Izbor adekvatne perceptualne karakteristike, ili više njih, mora da odgovara konkretnoj primeni pa je uvek predmet promišljanja koje se vrši na osnovu prethodnih iskustava, eksperimenata i intuicije.

### 5.1.3. Kategorizacija akustičnih talasa

Kao kod svih kognitivnih procesa u osnovi procesa prepoznavanja zvuka leži kategorizacija (taksonomija). Kategorije ili klase su grupe objekata koje su napravljene po kriterijumu sličnosti u nekom referentnom okviru. Podela objekata u klase obezbeđuje previla

uopštavanja na osnovu kojih se vrši predviđanje i donošenje odluka. Taksonomija nekog fenomena je, u realnim situacijama, uvek kompleksna i višesmislena. Osim svojstva objekata, ona je uslovljena karakteristikama samog posmatrača i njegovim iskustvom. Slika 5.6 prikazuje jednu od mogućih taksonomija izvora zvuka koja je napravljena na osnovu modela Aaron-a Bobick-a [70,71]. Na vrhu taksonomije je kategorija označena kao izvor zvuka koja obuhvata sve objekte i fenomene koji proizvode zvuk uopšte. Ovu kategoriju je lako predvideti zato što zahteva samo binarnu informaciju, da li ima zvuka ili nema. Na samom dnu taksonomijskog stabla su kategorije koje sadrže isključivo jedan objekat koji proizvodi specifičan, dobro poznati zvuk. Objekti ovog nivoa apstrakcije su usko specifični, zahtevaju više iskustva i zbog toga su veoma teški za predviđanje. Kako nivo apstrakcije raste, potrebno je manje podataka za kategorizaciju ali ona pruža manje informacija o objektu.



Slika 5.6 Prepoznavanje kao apstrakcija u prostoru kategorija

Sa stanovišta veštačke inteligencije kategorizacija audio signala spada u domen prepoznavanja obrazaca. Za ovu namenu predlažu se različite tehnike (Poglavlje 2.3.2.1). Mašina potpornih vektora, kao metod nadgledanog učenja, pokazala se kao upotrebljiva za klasifikaciju kompleksnih (miksovanih) audio signala [67]. Nedostatak ove metode je što ne daje verovatnoću događaja jer je originalno razvijena za rešavanje binarnih problema. Da bi se primenila za rešavanje problema sa više klasa potreban je dodatni napor [72]. Skriveni Markovljev metod posmatra čitav sistem kao Markovljev proces [73]. Osim za kategorizaciju, procedure je često primenjivanja za segmentaciju akustičnih signala [68]. Skriven Markovljev proces je zahtevan, u

pogledu neophodnih memorijskih i računskih kapaciteta. Takođe je kompleksan za primenu jer zahteva inicijalno podešavanje parametara i modela kako bi pružio zadovoljavajuće rezultate. Model Gausovih mešavina spada u nenadgledano učenje i obično se koristi za modeliranje pozadinskih i ostalih neočekivanih zvukova. Model se postepeno podešava za vreme perioda eksploatacije. Jedna od najjednostavnijih metoda kategorizacije je algoritam k-najbližih suseda, tip učenja na osnovu primera ili lenjog učenja [74]. Algoritam je osetljiv na lokalnu strukturu podataka, što znači da se nepoznati uzorak dodeljuje onoj grupi koja u okolini ima najviše predstavnika.

Veštačke neuronske mreže su dokazane u prepoznavanju obrazaca i klasifikaciji [38,75] pa se stoga koriste i za kategorizaciju akustičnih signala. Zahvaljujući hijerarhijskoj strukturi, odlikuje ih masivni paralelizam odnosno istovremena obrada informacija u više procesnih jedinica. Ovakav mehanizam obezbeđuje veliku brzinu u fazi eksploatacije što omogućava primenu u realnim uslovima. Neuronske mreže su jednostavne za implementaciju, adaptivne i kompatibilne sa drugim algoritmima veštačke inteligencije.

## **5.2. Eksperiment akustičnog prepoznavanja**

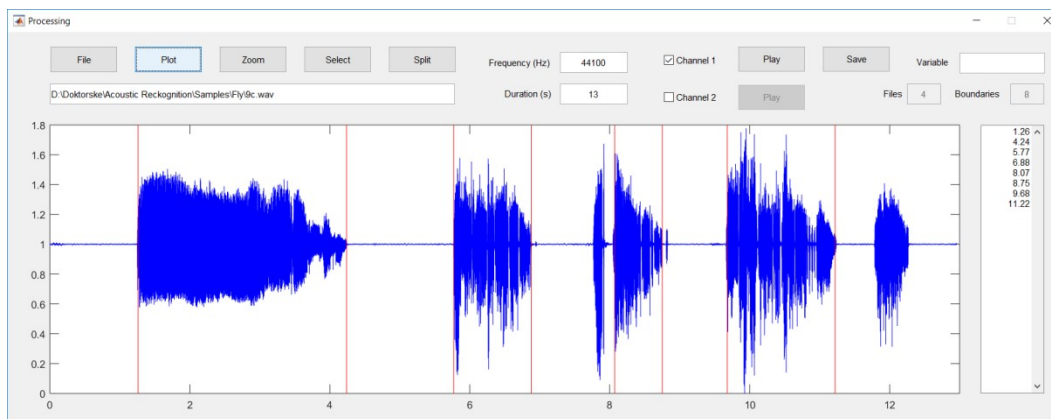
Radi potvrde teoretskih zapažanja izvršeno je eksperimentalno istraživanje sa zadatkom dizajniranja sistema veštačke inteligencije za prepoznavanje akustičnog izvora u realnim uslovima metodom nadgledanog učenja. U skladu sa zadatkom, postavljen je zahtev robustnosti postupka u odnosu na prisustvo slučajnog šuma i konkurentnih signala dok je računaska kompleksnost algoritma ograničena kako ne bi ugrozila primenu u realnom vremenu. Posebno se insistiralo na univerzalnosti rešenja kako bi bilo primenjivo u što većem spektru različitih situacija i od strane najšireg mogućeg kruga korisnika.

### **5.2.1. Prikupljanje akustičnih uzoraka**

Kao uzorci za obučavanje i testiranje sistema primenjeni su zvukovi insekata. Insekti upražnjavaju naprednu akustičnu komunikaciju uprkos relativno nerazvijenim organima za recepciju i obradu signala [76]. Zvukovi insekata se mogu smatrati nepovoljnim sa stanovišta prepoznavanja jer ne poseduju karakteristike harmoničnih signala kao što su visina i boja. Ako

sistem prepoznaje ovakve zvukove onda će sa lakoćom prepoznavati zvukove sa povoljnijom harmonijskom strukturom.

Za potrebe eksperimenta snimljeno je i sakupljeno sa interneta 685 digitalnih uzorka zvuka, na standardnoj frekvenciji od 44,1 kHz, u trajanju od 1 s. U tu svrhu napravljen je jednostavan program koji omogućava preslušavanje akustičnih uzoraka i istovremenu vizuelnu kontrolu signala. Primenom posebnog alata, izvršena je segmentacija signala i eliminisanje onih delova koji ne sadrže zvuk od interesa. Preostale delove signala program je podelio na partije sa minimalno 44100 tačaka diskretizacije (1s). Radi simuliranja različitih nivoa apstrakcije uzorci zvuka su pripadali trima kategorijama. Prvu kategoriju činilo je 32 snimljena uzorka, svaki od njih proizveden od strane istok cvrčka. Uzorci cvrčka predstavljali su najspecifičniju od tri kategorije uzoraka u prostoru apstrakcije kategorija (Slika 5.6). Cvrčak je izabran kao jedan od simbola akustične komunikacije među insektima. Drugu kategoriju činilo je 261 akustičnih uzoraka proizvedenih od strane muva koje su pripadale različitim podgrupama i porodicama muva. Ova kategorija reprezentovala je sredinu prostora apstrakcije. Muve su izabrane kao jedan od najčešćih insekata sa kojima se čovek susreće pa zato postoje potencijali praktične primene njihove detekcije. Poslednju kategoriju činilo je 392 uzorka zvuka različitog porekla, počev od ljudskog glasa, životinja, prirodnih fenomena pa sve do najrazličitijih motora, mašina, vozila, muzičkih instrumenata i ostalih tehničkih sredstava. To su različiti zvukovi koji u realnim uslovima mogu da ometaju akustično prepoznavanje. Ova klasa je jednostavno nazvana “Zvuk” jer predstavlja najmanje specifičnu odnosno najapstraktniju kategoriju u prostoru apstrakcije kategorija.



Slika 5.7 Segmentacija akustičnog signala

Pri izboru uzoraka za treniranje prvo je nasumičnim putem birana jedna od tri kategorije a onda opet nasumično uzorak koji pripada izabranoj kategoriji. Ovim je obezbeđena ravnopravna zastupljenost svih kategorija u skupu podataka za treniranje. Nakon izbora signala, samo njegov deo je bio upotrebljen u procesu treniranja. Početak signala biran je nasumično čime je obezbeđen dodatni nivo uopštavanja, u smislu da se isti signal ne može ponoviti bez obzira na broj uzoraka u skupu podataka za treniranje.

### 5.2.2. Ekstrakcija osobina akustičnog signala

Kao osnova za prepoznavanje zvuka izabran je frekventni spektar, univerzalna karakteristika zvuka koji o njemu pruža najviše informacija. Frekventni spektar leži u osnovi ostalih frekventnih karakteristika kao i perceptualnih karakteristika koje živi organizmi koriste prilikom prepoznavanje zvuka (Poglavlje 5.1.2). Pod uslovom da je sistem predstavljen nizom diskretnih vrednosti

$$x(kT) = x[k] \quad 0 < k < N - 1 \quad (5.2)$$

gde je  $T$  period diskretizacije a  $N$  dužina akustičnog signala, frekventni spektar se može aproksimirati brzom Fourier-ovom transformacijom

$$X_k = \sum_{n=0}^{N-1} x_n e^{-2ikn\pi/N}, \quad 0 < k < N - 1 \quad (5.3)$$

Ova transformacija je, već duže vreme, osnovni alat za transformaciju digitalnih signala iz vremenskog u frekventni domen. Razlog je manji računaska kompleksnost,  $O(N \cdot \log N)$ , u odnosu na osnovni metod,  $O(N^2)$ . Primenom brzih Fourier-ovih transformacija postiže se velika ušteda u računanju, naročito kada se radi o dugačkim signalima. Rezultat brzih Fourier-ovih transformacija (5.3) je niz kompleksnih brojeva sa istim brojem elemenata kao i originalni signal. Zapravo rezultat ima duplo više članova jer je svaki kompleksni broj predstavljen svojim realnim i imaginarnim delom, ili amplitudom i fazom zavisno od reprezentacije. Na sreću, ako su svi članovi signala realni brojevi, što je uvek slučaj u praksi, koeficijenti u drugoj polovine Fourier-ovog niza su konjugovano-kompleksni brojevi koeficijenata prve polovine niza

$$\operatorname{Re}(X_k) = \operatorname{Re}(X_{N-k}), \quad \operatorname{Im}(X_k) = \operatorname{Im}(X_{N-k}), \quad 0 < k < N - 1 \quad (5.4)$$

To znači da se druga polovina koeficijenata može zanemariti pa broj korisnih podataka ostaje isti kao u vremenskom domenu. Za procedure prepoznavanja upotrebljen je samo amplitudni spektar

$$|X_k| = \sqrt{Re^2(X_k) + Im^2(X_k)}, \quad 0 < k < \frac{N}{2} - 1 \quad (5.5)$$

Svaki pokušaj uključivanja faznog spektra

$$\angle X_k = \arctg \frac{Im(X_k)}{Re(X_k)}, \quad 0 < k < \frac{N}{2} - 1 \quad (5.6)$$

u proceduru prepoznavanja pokazao se kao neuspešan. Uprkos dvostrukoj računskoj kompleksnosti tačnost je ostala približno ista. Na kraju, skup podataka za treniranje sačinjen je od parova ulazno-izlaznih veličina

$$S = \left\{ (u_1, z_1), (u_2, z_2), \dots, (u_{N/2}, z_{N/2}) \right\} \quad (5.7)$$

Gde su  $u_i, i = 1, \dots, N/2$ , vektori koji sadrže koeficijente amplitudnog spektra, određene na osnovu jednačine (5.5), a  $z_i, i = 1, \dots, N/2$ , trodimenzionalni vektori koji sadrže očekivane (tačne) vrednosti izlaza. Vektori očekivanih izlaza imaju vrednost 1 na poziciji koja odgovara tačnom odgovoru (kategoriji) i vrednost 0 na ostalim pozicijama.

### 5.2.3. Kategorizacija akustičnih signala

Zbog svoje univerzalnosti, jednostavnosti i brzine, za obradu akustičnih signala u cilju prepoznavanja primenjena je veštačka neuronska mreža sa direktnim prostiranjem. Broj neurona u ulaznom sloju određen je dužinom signala zato što je svaki od  $N/2$  elemenata amplitudnog spektra upotrebljen kao ulaz u neuronsku mrežu. Izlazni sloj je činilo 3 neurona koji odgovaraju trima kategorijama akustičnih uzoraka. Greška neuronske mreže izračunata je kao razlika između očekivanog i ostvarenog rezultata. Ostvarene vrednosti izlaza se, u praksi, uvek kreću između nule i jedinice. Kao rešenje problema prepoznavanja, usvaja se ona kategorija signala kojoj odgovara pozicija sa najvećom izlaznom vrednošću. Kod prepoznavanja jedne klase obrazaca razlikuju se dve principijelno različite greške. Prva greška određuje koliko je elemenata drugih klasa pogrešno svrstano u predmetnu klasu a druga greška meri koliko elemenata predmetne klase je pogrešno svrstano u druge klase. Drugim rečima, svi elementi jedne klase uspešno



klasifikovani ne čine dobru kategorizaciju ako je mnogo elemenata drugih klasa u nju svrstano pogrešno. Ovaj problem se najčešće ilustruje matricom (tabelom) konfuzije (slika 5.8). U ovoj tabeli karakter „C“ simbolizuje kategoriju cvrčaka, „M“ muva a „Z“ kategoriju nazvanu „Zvuk“. Unutar obojenih polja (crvenih i zelenih) prvi karakter označava ciljnu vrednost a druga ostvarenu vrednost prepoznavanja. Očigledno je da dijagonalna polja (obojena zeleno) su tačni slučajevi prepoznavanja gde je cvrčak prepoznat kao cvrčak, muva kao muva a ostali zvuci kao Zvuk. Ostala polja (obojena crveno) predstavljaju slučajeve pogrešne kategorizacije. Na primer, drugo polje prvoga reda govori da je cvrčak prepoznat pogrešno kao muva a treće da je kategorizovan u klasu „Zvuk“. Umesto simbola, prava matrica konfuzije sadri broj uzoraka u procentima. Polja u desnoj koloni, označena sa „×“, pokazuju koliko je elemenata određene klase uspešno prepoznato. Za cvrčka, na primer, ova vrednost se izračunava kao količnik broja cvrčaka koji su korektno prepoznati i ukupnog broja uzoraka koji su prepoznati kao cvrčak

$$\times = \frac{C \rightarrow C}{C \rightarrow C + M \rightarrow C + Z \rightarrow C} \quad (5.8)$$

dok polja u dnu, označena sa „o“, pokazuju koliko procenta uzoraka koji su kategorizovani u neku klasu stvarno pripadaju toj klasi. Za muvu, na primer, ova vrednost izračunava se kao količnik broja muva koje su korektno prepoznate i ukupnog broja muva u skupu ulaznih podataka.

$$\times = \frac{F \rightarrow F}{F \rightarrow C + F \rightarrow F + F \rightarrow Z} \quad (5.9)$$

Obe veličine predstavljaju meru uspešnosti kategorizacije izraženu u procentima. Nedostatak reprezentacije rezultata prepoznavanja u vidu matrice konfuzije je što broj podataka raste sa geometrijskom progresijom u odnosu na broj kategorija.

Pored same kategorizacije, od neuronske mreže se očekuje da vrši i neke druge funkcije. Na primer, podešavanje težinskih koeficijenata ulaznog sloja može se shvatiti kao ekstrakcija osobina zvuka. Ako neki od ulaza ima, nakon procedure obučavanja mreže, koeficijent pojačanja čija je apsolutna vrednost bliska nuli onda se on može zanemariti jer se nije pokazao kao relevantan. Ovo je retka prilika da neuronska mreža pruža povratnu informaciju o procesu koji modelira. Kompletan ulazni sloj neuronske mreže u eksperimentu može se interpretirati kao digitalni filter jer se njime vrši podešavanje koeficijenata amplitudnog spektra.

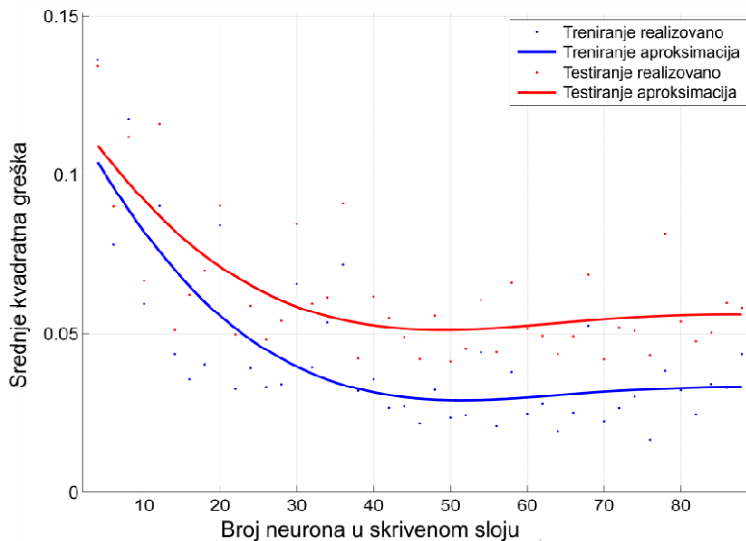
		Očekivano			
		C	M	Z	
Ostvareno	C	C→C	M→C	Z→C	×
	M	C→M	M→M	Z→M	×
	Z	C→Z	M→Z	Z→Z	×
		○	○	○	

Slika 5.8 Matrica konfuzije

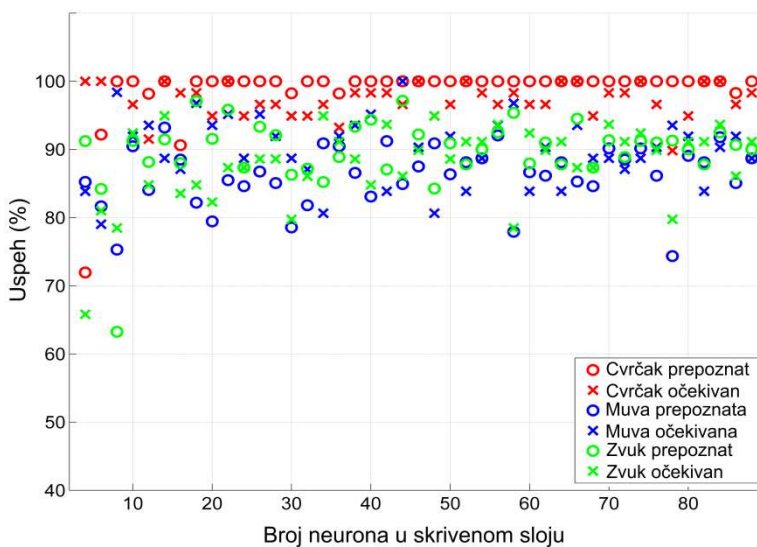
#### 5.2.4. Rezultati

Prvi izazov, u realizaciji opisanog metoda prepoznavanja zvuka bio je definisanje strukture neuronske mreže. Pošto je broj terminala na ulazu uslovljen dužinom signala i njegovom frekvencijom a broj izlaza brojem kategorija (u eksperimentu 3) akustičnih uzoraka, ostalo je da se odredi struktura skrivenih slojeva. U nedostatku univerzalnog pravila za dizajniranje veštačke neuronske mreže, struktura skrivenih slojeva bila je predmet eksperimenta. Testirano je 44 konfiguracije sa počev od 2 pa sve do 88 neurona u jedinstvenom skrivenom sloju. Za obučavanje svake od konfiguracija upotrebljeno je 500 a za testiranje 200 akustičnih uzoraka u trajanju od 0,5 s. Nakon 500 epoha, primenom algoritma sa povratnim prostiranjem greške i momentom, dobijeni su rezultati predstavljeni na slici 5.9. Tačke reprezentuju srednje kvadratne greške treniranja i testiranja a linije aproksimacije ovih veličina primenom kubnog polinoma. Grafik pokazuje da greške opadaju, uz neznatne oscilacije, do konfiguracije sa 40 neurona, u skrivenom sloju, a nakon toga ostaje manje-više konstantne. Interesantno je da se oblik grafika nije mnogo menjao sa promenom dužine signala. Na slici 5.10 predstavljena je uspešnost kategorizacije, u procentima, za svaku od testiranih konfiguracija. Svaka od kategorija predstavljena je dvema merama uspešnosti koje su dobijene na osnovu izraza (5.8) i (5.9). Najmanja devijacija rezultata može se primetiti na sredini grafika, u istoj zoni koja je na prethodnom grafiku označena kao optimalna (slika 5.9). Testiranje konfiguracija sa više skrivenih slojeva nije rezultiralo boljom tačnošću u odnosu na jednoslojne sa istim brojem

neurona. Zbog toga je konfiguracija sa 50 veštačkih neurona u jedinstvenom skrivenom sloju usvojena kao optimalna. Sva naredna testiranja u eksperimentu vršena su primenom neuronske mreže sa ovom konfiguracijom.



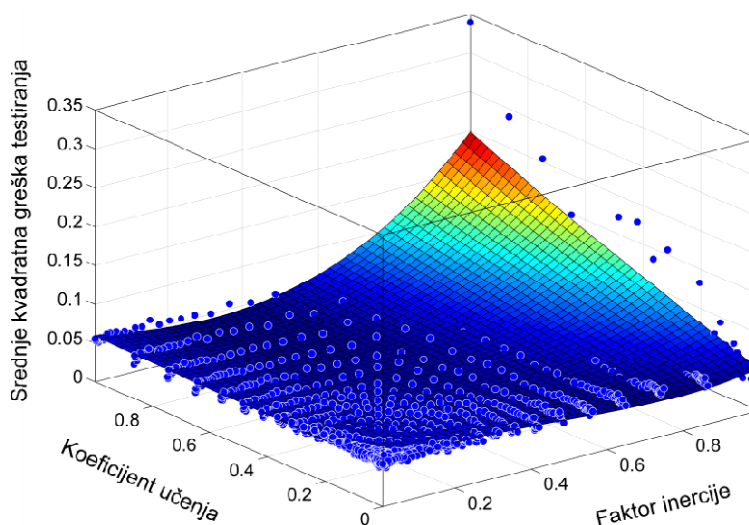
Slika 5.9 Srednje kvadratna greška u zavisnosti od broja neurona u skrivenom sloju



Slika 5.10 Uspeh u zavisnosti od broja neurona u skrivenom sloju

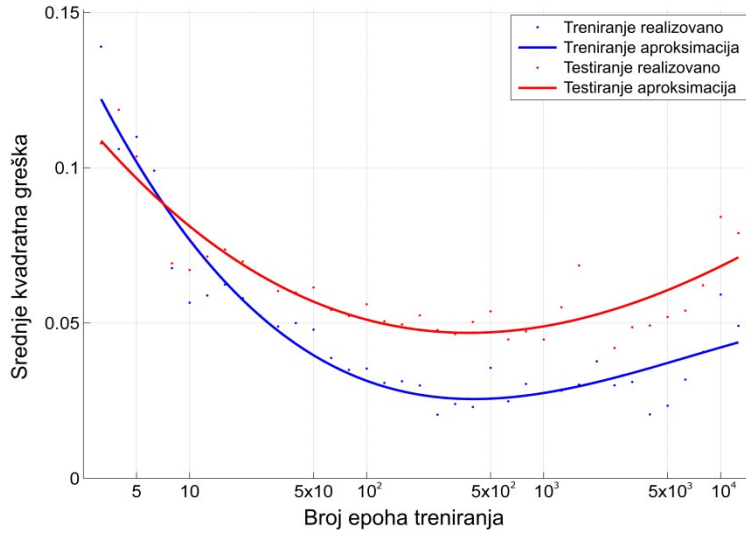
Sledeći zadatak bio je određivanje parametra algoritma za treniranje. Srednje kvadratna greške treniranja i testiranja određene su za 1122 kombinacija parametara sa 33 različite vrednosti koeficijenta učenja, počev od 0,01 pa sve do 0,999, i 34 različite vrednosti faktora

inercije, počev od 0 pa sve do 0,999. Rezultati, koji se odnose na testiranja predstavljeni su, plavim tačkama, na trodimenzionalnom grafiku (slika 5.11). Aproksimaciona površina dobijena je primenom polinomske površine trećeg stepena u oba pravca. Odgovarajući grafik treniranja imao je veoma sličan oblik, naravno sa manjim vrednostima greške. Na osnovu rezultata određene su vrednosti koeficijenta učenja od 0,025 i faktora inercije 0,999 kao optimalne u smislu tačnosti prepoznavanja. Konfuzione matrice koje su određene za svaku od tačaka bile su u saglasnosti sa ovim rezultatima. Ove vrednosti pokazuju da je predložena procedura izuzetno osetljiva na oscilacije u znaku gradijenta [38]. Optimalne vrednosti su primenjene u deljem testiranju neuronske mreže.

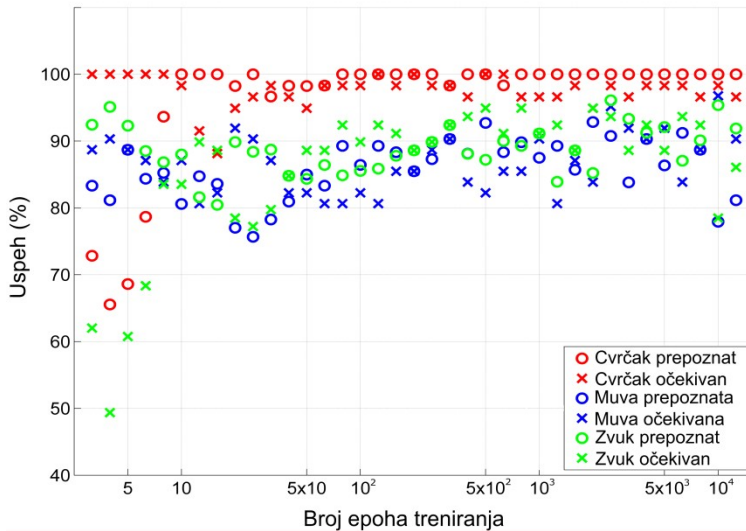


*Slika 5.11 Srednje kvadratna greška testiranja u zavisnosti od parametra treniranja*

U sledećem testu ispitivana je zavisnost srednje kvadratne greške treniranja i srednje kvadratna greške testiranja u odnosu na broj epoha algoritma sa povratnim prostiranjem greške i momentom. Obezbeđeni su rezultati za 37 različitih broja epoha, koje su ravnomerno raspoređene duž logaritamske skale. Najkraće treniranje, među njima, trajalo je 3 epohe a najduže 12589 epoha. Na slici 5.12 prikazani su dobijeni rezultati, u vidu tačaka, kao i aproksimacione krive koje su obezbeđene primenom kubnog polinoma. Slika 5.13 predstavlja odgovarajući grafik uspeha prepoznavanja. Na osnovu oba grafika, najbolji rezultati postignuti su treniranjem neuronske mreže u trajanju između 200 i 500 epoha. Dalji trening, nakon 500 epoha, dovodio je mrežu u stanje pretreniranosti usled čega su se greške povećavale. Na osnovu dobijenih rezultata 400 epoha je usvojeno, kao optimalna dužina treniranja, i primenjeno u daljim testiranjima.



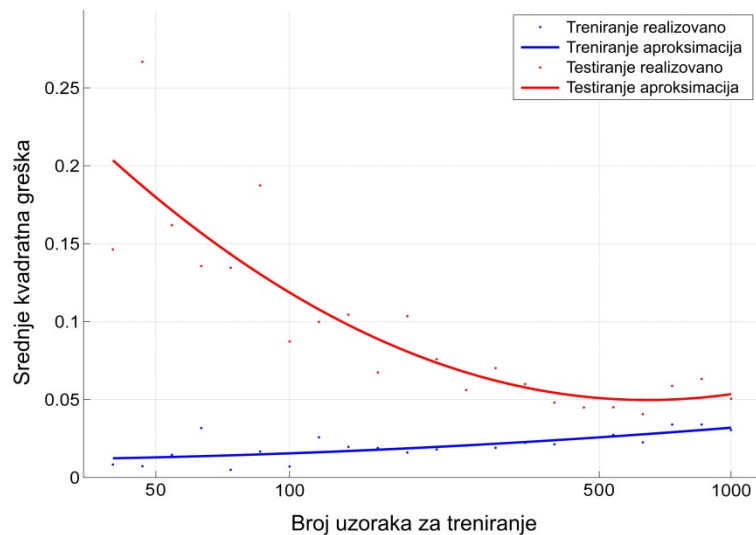
Slika 5.12 Srednje kvadratna greška u zavisnosti od broja epoha treniranja



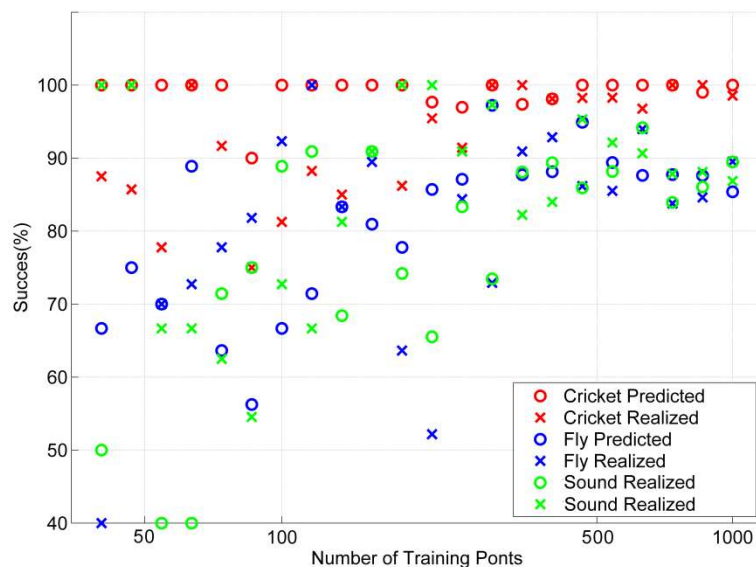
Slika 5.13 Uspeh u zavisnosti od broja epoha treniranja

Poslednji test, vezan za algoritam treniranja, bio je ispitivanje greške u zavisnosti od broja uzoraka za treniranje. Procedura je izvedena za 22 skupa podataka sa različitim brojem elemenata. Broj uzoraka ravnomerno je raspoređen na logaritamskoj skali, počev od najmanjeg skupa sa 40 pa do najvećeg sa 1000 elemenata. Broj podataka za testiranje, u svakom od skupova podataka bio je 40% od ukupnog broja. Dobijeni rezultati su prikazani u vidu tačaka na slici 5.14. a njihove aproksimacije, primenom dvočlane eksponencijalne funkcije, prikazane su linijama. Uspeh prepoznavanja, nad istim skupovima uzoraka, prikazan je na slici 5.15. Rezultati potvrđuju logičnu pretpostavku da više tačaka za treniranje znači veću grešku treniranja i manju

grešku testiranja. U daljim testiranja primenjeno je 500 podataka za treniranje i 200 za testiranje jer dalje povećanje broja uzoraka nije dalo značajno unapređenje tačnosti prepoznavanja.



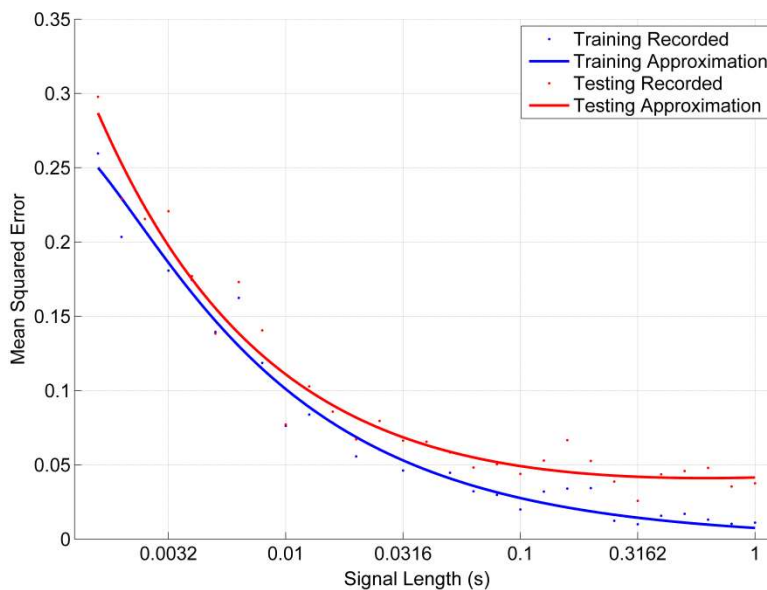
Slika 5.14 Srednje kvadratna greška u zavisnosti od broja uzoraka



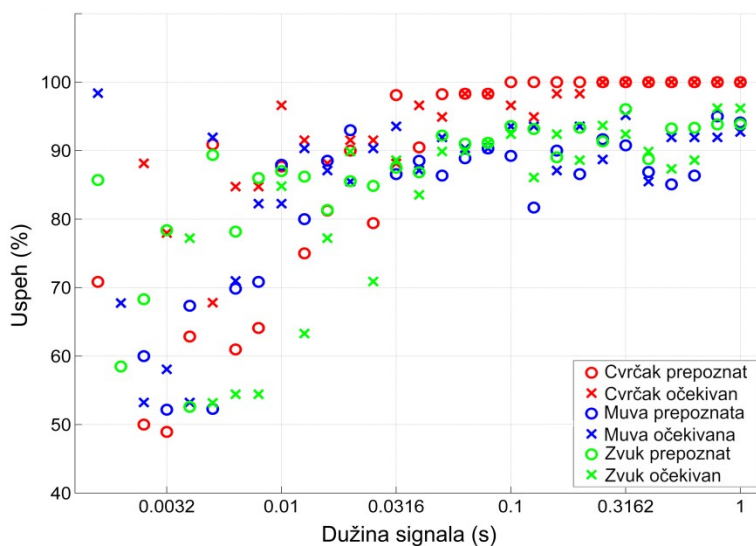
Slika 5.15 Uspeh u zavisnosti od broja tačaka za treniranje

Kako je frekvencija fiksirana na 44.1 kHz, dužina signala određuje koliko će elemenata imati svaki akustični uzorak. Što je signal duži to pruža više informaciju o svom izvoru. Ova zavisnost je snimljena primenom 29 različitih dužina signala, počev od najkraćeg koji je trajao 0,0016 s pa do najdužeg koji je trajao 1 s, ravnomerno raspoređenih na logaritamskoj skali. Ovim dužinama odgovaraju različite konfiguracije ulaznog sloja od kojih najkraćoj odgovara 71 a

najdužoj 44100 ulaznih terminala. Rezultujuće greške (slika 5.16) pokazuju, logično, da duži signal proizvodi manju grešku prepoznavanja. Sa druge strane, duži signal zahteva veći računski napor. Dužina signala od 0,1 s usvojena je kao kompromis jer su greške (naročito testiranja) nakon ove tačke opadale sporije. Aproksimacija rezultata, prikazana linijama, dobijena jednočlanom eksponencijalnom funkcijom. Grafik uspeha, prikazan na slici 5.17, u potpunosti je u saglasnosti sa navedenim zapažanjima.



Slika 5.16 Srednje kvadratna greška u zavisnosti od dužine signala



Slika 5.17 Uspeh u zavisnosti od dužine signala



Ni jedan od prethodnih grafika nije bio dobijen nakon prvog pokušaja. Oni su rezultat nekoliko ciklusa testova sa ciljem nalaženja optimalnih parametara. To znači da svaki od predstavljenih grafika simulira promenu jednog parametra (ili dva) u okoline tačke u prostoru parametara. Optimum srednje kvadratne greške treniranja je iznosio 0,03 a testiranja 0,05 za konfiguraciju mreže sa 50 neurona u jedinstvenom skrivenom sloju, posle 400 epoha treniranja algoritmom sa povratnim prostiranjem greške i inercijom, sa koeficijentom učenja 0,025 i faktorom inercije 0,999. Sporadično je bilo i boljih rezultata ali su zanemareni pa je minimum aproksimacione površine usvojen kao optimalan. Dužina signala od 0,1 s i broj uzoraka za treniranja (500) usvojene su kao kompromisne veličine u skladu sa raspoloživim resursima. Jasno je da svako povećanje ovih vrednosti rezultuje povećanjem tačnosti rezultata. Na slici 5.18 predstavljena je matrica konfuzije u tački optimuma.

		Očekivano			
		C	M	Z	
Ostvareno	C	53	0	1	98.1%
	M	0	55	9	85.9%
	Z	0	6	76	92.7%
		100%	90.2%	88.4%	92.0%

Slika 5.18 Matrica konfuzije u tački optimuma

### 5.2.5. Zaključak

Minimalna vrednost srednje kvadratne greške testiranja od 0,05, koja odgovara tačnosti testiranja između 77 i 78%, može izgledati obeshrabrujuće ali matrica konfuzije (Slika 5.18) iste konfiguracije pokazuje tačnost od 92%. Razlika je posledica digitalizacije izlaza koja se vrši prilikom procesa prepoznavanja. Da bi neki od izlaza bio izabran kao rešenje procedure klasifikacije on ne mora i nikad nema vrednost 1. U opisanom eksperimentu, sa 3 izlaza, teoretski je moguće da obrazac bude izabran sa vrednošću izlaza neznatno većom od 1/3.



Matrica konfuzije i dijagrami tačnosti (Slike 5.10, 5.13, 5.15 i 5.17) pokazuje da je optimalna konfiguracija prepoznavala cvrčka sa tačnošću od gotovo 100% (crveno o). Ponekad, ali veoma retko, u 1 do 2% slučajeva, dešavalo se da ostali uzorci budu pogrešno prepoznati kao cvrčak (crveno x). Tačnost prepoznavanja za ostale dve kategorije varirala je oko 90%. Razlog leži u činjenici da cvrčak zahteva najniži nivo apstrakcije da bi bio prepoznat, jer uzorci potiču od jedne iste jedinice. Na osnovu istog principa, logično bi bilo da muve budu tačnije prepoznate od treće kategorije zvukova. Međutim, jednom kada je iz izbora eliminisan cvrčak, druge dve kategorije uzoraka se međusobno isključuju zbog čega imaju sličnu tačnost prepoznavanja.

Eksperiment je potvrdio uspešno prepoznavanje akustičnih signala srednjeg nivoa apstrakcije. Ako se zvuk sa nižim nivoom apstrakcije učestalo ponavlja ukupna tačnost prepoznavanja može biti unapređena uvođenjem nove klase. Manje apstraktni elementi se prepoznaju sa većom tačnošću što podiže totalni nivo tačnosti akustičnog prepoznavanja.

## 6. LOKALIZACIJA AKUSTIČNOG IZVORA

Akustična lokalizacija predstavlja pozicioniranje izvora zvuka, u odnosu na neki referentni prostorni okvir, kao rezultat analize signala koje proizvodi. Ima najrazličitije vojne, industrijske, naučne i poslovne primene. Neki od primera su inteligentna životna okruženja [77], sistemi za održavanje i strukturalno nadgledanje [78], sistemi za nadzor [79], sonar [80], sistemi za lociranje artiljerijske vatre [81], prepoznavanje govora [82] i sl. Lokalizacija izvora zvuka može se vršiti u čvrstim, tečnim ili gasovitim sredinama na aktivan ili pasivan način. Pasivni sistemi koriste akustični signal koji dolazi od izvora zvuka da bi ih locirali. Sa druge strane, aktivna lokalizacija koristi se za lociranje objekata koji ne moraju biti izvori zvuka. Sam sistem šalje zvučne impulse i hvata njihov eho nakon što se odbiju od okoline. Analizom primljenih signala vrši se lokalizacija traženih ciljeva [80]. Audio lokalizacija može biti primenjena u širokom rasponu dimenzija, što uglavnom zavisi od snage izvora zvuka. Lokalizacija snažnih izvora zvuka, kao što su grmljavina, eksplozije vulkana [83], veštačke eksplozije ili motori aviona [84], mogu se vršiti sa više desetina kilometara dok se kod preciznih mehanizama i ispitivanju strukture materijala radi o mikrometarskoj skali [85,86].

### 6.1. Mehanizam akustične lokalizacije kod živih bića

Mnogi veštački sistemi pokušavaju da oponašaju sposobnost lokalizacije izvora zvuka koje ispoljavaju živa bića [87]. U početku se smatralo da samo frekventne karakteristike zvučnog talasa igraju ulogu u određivanju njegove lokacije, jer unutrašnje uho kodira akustični signal na osnovu spektralnih karakteristika. Savremena istraživanja dokazala su da vremenske karakteristike igraju možda još veću ulogu od frekventnih a takođe i uticaj informacija dobijenih na jednom uhu, nezavisno od drugog [88].

#### 6.1.1. Dupleks teorija

Ljudi i životinje (kičmenjaci pre svega) sposobni su da odrede horizontalni pravac (azimut) iz koga dopire zvuk na osnovu neznatne razlike između signala koji dopiru do ušiju [89]. Ovu pojavu je prvo opisao britanski fizičar Lord Rayleigh (1842-1919) u svojoj teoriji o lokalizaciji zvuka nazvanoj dupleks teorija. Prema njegovom mišljenju lokalizacija izvora zvuka bazirana je na dva fenomena, oba posledica različitog geometrijskog položaja ušiju. Posledica različitog puta

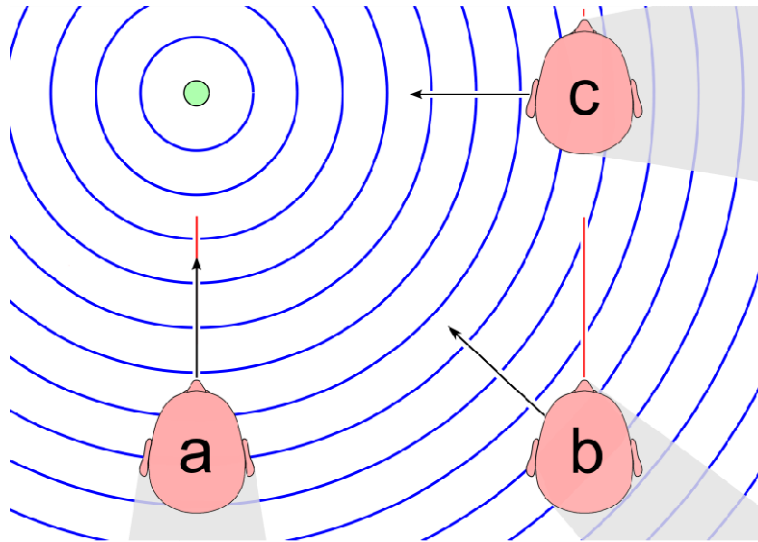
koji signal treba da pređe do levog i desnog uha (Slika 6.1 b i c), osim u specijalnom slučaju kada se izvor nalazi u srednjem planu (vertikalnoj ravni simetrije) glave (Slika 6.1 a), je razlika u momentu registrovanja kao i razlika u intenzitetu akustičnog talasa. Vremenski period između momenata kada signal stiže do levog odnosno desnog uha naziva se interauralna vremenska razlika (engl. Interaural Time Difference). Nervni sistem je sposoban da registruje ovu razliku i na osnovu nje odredi horizontalni pravac izvora zvuka. Za kontinualne signale ova razlika se može definisati kao interauralna razlika faze (engl. Interaural Phase Difference). Sposobnost nervnog sistema da registruje razliku u fazi je limitirana. U opštem slučaju, kod čoveka je ova funkcionalnost potpuna za frekvencije do 1 kHz, onda počinje da opada sve do frekvencija reda veličine 1,5 kHz kada se potpuno gubi. Prema ovome, osetljivost čoveka na razlike u fazi ograničena je na niske frekvencije zvuka. Deo razlike u intenzitetu zvuka (engl. Interaural Intensity Difference) takođe potiče od razlike puta jer intenzitet sfernog akustičnog talasa opada sa kvadratom pređenog puta

$$I = \frac{P}{4\pi r^2} \quad (6.1)$$

gde je  $P$  snaga izvora a  $r$  rastojanje od njega. Drugi uzročnik je čvrsta struktura glave koja **reflektuje** (odbija) i **refraktuje** (menja smer) akustičnog talasa. Ovi efekti su karakteristični za akustične signale čija je talasna dužina reda veličine glave ili manja od nje, odnosno za zvukove visoke frekvencije. Glava pravi akustičnu senku za uho (Slika 6.1 b i c) koje je udaljenije od izvora zvuka, usled čega se javlja dodatna razlika u intenzitetu. Efekat senke na intenzitet zvuka zavisi od položaja njegovog izvora. Razlika u intenzitetu između signala visoke frekvencije registrovanih na različitim ušima može biti i preko 40 dB. Nervni sistem sposoban je da registruje mnogo manju razliku u intenzitetu akustičnog talasa i na osnovu nje generiše informaciju o pravcu njegovog prostiranja. Kod čoveka je ova sposobnost zanemarljiva za zvuke niske frekvencije, ispod 3 kHz, dok za signale iznad ove frekvencije počinje drastično da raste. Navedene sposobnosti i opsezi frekvencija zavise od morfologije glave i ušiju zbog čega individualno variraju.

U današnje vreme postoje brojne opservacije lokalizacije izvora zvuka koje ne mogu biti objašnjene dupleks teorijom kao i brojni eksperimenti koji su u kontradiktornosti sa njom. Najveći dokaz nekompletnosti dupleks teorije je činjenica da ljudi i životinje koje imaju oštećeno

čulo sluha, samo na jednoj strani, takođe ispoljavaju sposobnost lokalizacije izvora zvuka. Osim toga, dupleks teorija ne pruža informaciju o vertikalnom uglu (uglu elevacije) iz koga dopire zvuk. Kao rezultat njene matematičke primene dobija se konusna površina usmerena od uha ka spoljašnosti koja se naziva se **konus konfuzije**. Uprkos nesavršenosti i nekompletnosti istorijska uloga dupleks teorije u objašnjavanju i predviđanju pojava vezanih za lokalizaciju izvora zvuka je neosporna.



*Slika 6.1 Interauralna razlika akustičnog signala*

### 6.1.2. Efekat ušne školjke

Čovekova ušna školjka ili arikula (engl. pinna) ima specifičan, apsolutno nesimetričan, konkavan oblik sa brojnim kanalima i izbočinama (slika 5.1). U poglavlju 5.1.1 je rečeno da se u ušnoj školjci vrši početna selekcija akustičnih parametara jer se različiti signali, u pogledu frekvencije i smera prostiranja, različito pojačavaju. Na osnovu spektra akustičnog talasa, koji je rezultat superpozicije direktnog zvuka i njegove refleksije od ušne školjke, nervni sistem je sposoban da odredi vertikalni pravac njegovo izvora [90]. Sposobnost lokalizacije uspostavlja se procesom učenja kroz mnogobrojne pokrete glave kojima se menja njena relativna pozicija u odnosu na izvor zvuka. Dupleks teorija načelno pominje sposobnost ušiju da filtriraju akustični signal ali o tome ne daje nikakve relacije i modele predviđanja [88]. Prvi kvantitativni modeli funkcije ušne školjke javili su se u drugoj polovini 20 veka ali je tek nedavno počela njena

redovna integracija u proceduru akustične lokalizacije. Efekat glave, zajedno sa ušnim školjkama, na određivanje pravca prostiranja zvuka često se predstavlja uz pomoć prenosne funkcije glave (engl. Head-Related Transfer Function). Ova funkcija modelira način na koji uho transformiše signal iz tačke u slobodnom akustičnom polju do tačke u unutrašnjem uhu. Prenosna funkcija glave zavisi od udaljenosti glave od izvora zvuka, horizontalnog i vertikalnog smera dopiranja zvuka, ugaoone brzine, oblika i dimenzija glave. Levo i desno uho karakterišu različite funkcije prenosa glave a njihova kombinacija pruža binauralnu informaciju o položaju izvora zvuka.

### 6.1.3. Određivanje udaljenosti akustičnog izvora

Sposobnost živih bića da odrede udaljenost izvora zvuka na osnovu signala koji dopiru do ušiju je ograničena ali evidentno postoji, naročito kada se izvor zvuka nalazi neposredno u blizini glave. U tu svrhu upotrebljavaju se različite karakteristike akustičnog signala.

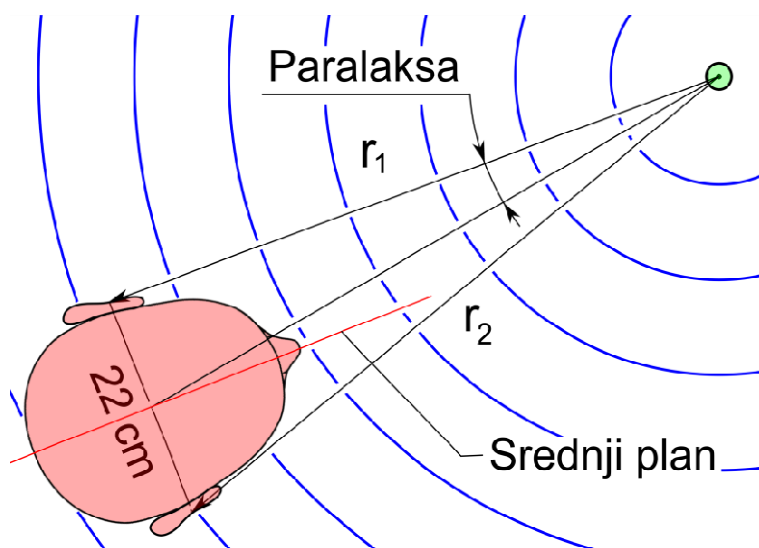
**Jačina** zvuka opada sa kvadratom udaljenosti od njegovog izvora (6.1), što znači da se udaljeni izvori zvuka čuju slabije od bližih. Slušni sistem čoveka i životinja sposoban je da primeti ovu razliku i na osnovu nje stvori percepciju o udaljenosti izvora zvuka. U slobodnom akustičnom polju, na udaljenosti većoj od 1 m od izvora zvuka, razlika u jačini između dve tačke opservacije može se odrediti na osnovu jednačine

$$\Delta L = 20 \log \left( \frac{r_1}{r_2} \right) \text{ dB} \quad (6.2)$$

gde su  $r_1$  i  $r_2$  rastojanja od izvora zvuka do prve i druge tačke opservacije respektivno. Prema jednačini (6.2) svako dupliranje rastojanja od izvora zvuka do tačke opservacije praćeno je redukcijom jačine zvuka od 6 dB. Nedostatak ovog pristupa je što izvor zvuka, u smislu njegove jačine, mora biti dobro poznat.

**Interauralna razlika jačine** zvuka može biti primenjena za percepciju udaljenosti akustičnog izvora čija jačina ne mora biti unapred poznata. Razlika u jačini zvuka između dva uha određuje se na osnovu jednačine (6.2) gde je  $r_1$  rastojanje od izvora zvuka do bližeg uha a  $r_2$  rastojanje od izvora zvuka do daljeg uha (slika 6.2). Razlika se povećava što se izvor zvuka udaljava od srednjeg plana glave ali kada jedno od ušiju padne u akustičnu senku (slika 6.1)

problem se drastično usložnjava. Kada je udaljenost izvora zvuka velika, u odnosu na dimenzije glave, razlika u jačini signala između ušiju je mala. Na sreću ljudsko uho je veoma osetljivo na promenu intenziteta zvuka. Najmanja promena koja može biti registrovana procenjuje se na 0,4 dB. Ovoj osetljivosti, pod pretpostavkom da je rastojanje između ušiju 22 cm odgovara udaljenost izvora zvuka od 4,6 m. Zbog uticaja akustične senke jednačina (6.2) se ne može primeniti za određivanje udaljenosti izvora koji se nalazi suviše bočno od srednjeg plana glave pa je ovaj opseg daleko manji. Na daljinama manjim od 1 m jačina zvuka odstupa od navedenog zakona a percepcija udaljenosti izvora vrši se na osnovu drugih osobina akustičnog talasa.



Slika 6.2 Položaj ušiju u odnosu na izvor zvuka

**Akustična paralaksa** je ugao između smera prostiranja zvuka prema centru glave i smera prostiranja zvuka u prema uhu (slika 6.2). Paralaksa se povećava kako se izvor zvuka približava glavi. Ovaj efekat je primetan naročito za izvore zvuka koji se nalaze u blizini glave. Sa promenom paralakse menja se karakteristika filtriranja akustičnog talasa koju vrši ušna školjka. Što je položaj akustičnog izvora udaljeniji od srednjeg plana glave to se ušna školjka više ponaša kao visokopropusni filter. Nervni sistem čoveka i životinja razvio je sposobnost da registruje promenu karakteristike zvuka u funkciji promene paralakse i na osnovu nje proceni udaljenost izvora, pod uslovom da se on nalazi u neposrednoj blizini glave [91]. Kako se izvor udaljava od glave tako ova indikacija postaje nepreciznija u odnosu na interauralnu razliku jačine zvuka. Kao granična, za primenu jedne ili druge metode, usvojena je distanca od 1m.

**Odnos zvuk/odjek** može biti indikacija za određivanje udeljenosti izvora zvuka u zatvorenim prostorijama ili u prirodnim okruženjima u kojima se javlja odjek. Ovde se pod zvukom podrazumeva samo akustični talas koji putuje direktno od izvora zvuka do tačke opservacije. Ovaj talas sledi inverzni kvadratni zakon (6.1) bez obzira da li se radi o zatvorenom ili otvorenom prostoru. Eho nastaje refleksijom direktnog talasa od zida ili nekog drugog čvrstog tela. Energija eha je konstantna za distance veće od kritične. Pod kritičnom distancom podrazumeva se ona pri kojoj su energije direktnog signala i eha jednake odnosno pri kojoj je odnos zvuk/odjek jednak jedinici. Prema tome, u zatvorenoj prostoriji, odnos direktni signal/odjek za distance veće od kritične obrnuto je proporcionalan rastojanju od izvora.

**Promena spektra** zvučnog signala, kao posledica njegovog prigušenja na putu kroz medijum, koristi se za procenu većih udaljenosti izvora. Što je izvor udaljeniji to je zvuk više prigušen pri čemu reflektujuće površine i prepreke u prirodnom okruženju brže apsorbuju visoke od niskih frekvencija zvuka. Zbog toga ljudi i životinje, na osnovu sopstvenog životnog iskustva, doživljavaju akustične signale niskih frekvencija kao udaljene. Na daljinama većim od 15 m primetno je dodatno prigušenje zvuka usled apsorbcije od strane vazduha pa je ukupan efekat još veći. I ova metoda je beskorisna za nepoznate izvore zvuka jer da bi efekat prigušenja bio procenjen posmatrač mora da dobro poznaje spektralnu karakteristiku originalnog zvuka, pre njegovog prigušenja.

## **6.2. Mehanizam akustične lokalizacije kod veštačkih sistema**

Veštački sistemi za audio lokalizaciju mogu biti realizovani uz pomoć jednog [92], dva [93] pa sve do masivnih nizova akustičnih senzora [94]. Druga podela reprezentativnih metoda je u odnosu na fizičke veličine koja se upotrebljava za obavljanje lokalizacije.

### **6.2.1. Monoauralna lokalizacija**

Lokalizacija izvora zvuku pomoću jednog akustičnog senzora, može biti vršena na osnovu fenomena opadanja energije zvučnog talasa prilikom prostiranja kroz medijum [95]. Nedostatak ovog pristupa je što karakteristike izvora i medijuma kroz koji se prostire zvuk moraju biti unapred poznate. Kao rezultat se dobija samo udaljenosti akustičnog izvora ali njegov geometrijski položaj ostaje nepoznat. Informacija o pravcu dopiranja zvuka, na osnovu jednog

senzora, može biti obezbeđena primenom usmerenih senzora ili različitih akustičnih reflektora (prepreka). Kao rezultat eksperimentalnih merenja definiše se prenosna funkcija akustičnog reflektora  $h_\theta(t)$ , analogno prenosnoj funkciji ušne školjke kod kičmenjaka, koja je zavisna od smera dopiranja zvuka. Signal dobijen na senzoru može se aproksimirati linearnom jednačinom

$$y(t) = h_\theta(t) * s(t) + w(t) \quad (6.3)$$

gde je “\*” operator konvolucije,  $s(t)$  je originalni signal zvuka a  $w(t) \in \mathcal{N}(0, \sigma_m)$  je aditivni Gausov (beli) šum. Pošto je proces vremenski invarijantan češće se predstavlja u frekventnom domenu

$$Y(f) = H_\theta(f)S(f) + W(f) \quad (6.4)$$

### 6.2.2. Binauralna lokalizacija

U cilju percepcije udaljenosti zvuka humanoidni roboti koriste sistem od dva (ili više) senzora sa odgovarajućim akustičnim reflektorima [96] koji su namenjeni simuliranju funkcije ušnih školjki kod čoveka. Kod ovakvih sistema značajan uticaj na karakteristike registrovanog zvuka, osim reflektora, vrši glava kao posebna akustična prepreka. Zbog toga se definiše prenosna funkcija glave odnosno prenosna funkcija koja sublimira uticaj svih akustičnih prepreka na karakteristike registrovanog zvuka. Obučavanjem binauralnih sistema putem procedure sukcesivnog pomeranja i percepcije može se ostvariti sposobnost detekcije pravca pa čak i lokalizacije izvora zvuka u prostoru [97].

### 6.2.3. Prostorno filtriranje

Ako ne postoji mogućnost pomeranja ono se može simulirati primenom tehnike formiranja snopa (engl. Beamforming) odnosno prostornog filtriranja. Ova popularna tehnika sastoji se skeniranju prostora od interesa u čvorovima,  $\vec{\rho}_j, j = 1, \dots, C$ , gde je  $C$  broj čvorova, geometrijske mreže. Svakom signalu pridružuje se odgovarajuće kašnjenje koje je jednako teoretski izračunatom vremenu putovanja akustičnog talasa od izvora zvuka do mikrofona

$$t_{ij} = \frac{|\vec{r}_i - \vec{\rho}_j|}{c_0} \quad (6.5)$$



gde je  $\vec{r}_i$  vektor položaja mikrofona  $i$ ,  $\vec{\rho}_j$  vektor položaja čvora  $j$ , a  $c_0$  brzina prostiranja zvuka kroz medijum. Na ovaj način, signali dobijeni na različitim sensorima poravnavaju se u vremenskom domenu, odnosno koriguju u fazi, nakon čega se izlaz snopa dobija njihovim sumiranjem

$$y(n, \vec{\rho}_j) = \sum_{i=0}^{M-1} q_{ij} s_i(n - t_{ij}) \quad (6.6)$$

gde je  $M$  broj akustičnih senzora,  $s_i(n)$  signal primljen na  $i$ -tom mikrofону,  $t_{ij}$  kašnjenje signala (6.5) i  $q_{ij}$  su faktori pojačanja. Ako je izvor zvuka udaljen od niza senzora, pod pretpostavkom da su karakteristike senzora u nizu uniformne, najčešće se radi uprosečavanje izlaza usvajanjem istih težinskih koeficijenata,  $q_{ij} = 1/M$  za  $i = 1, \dots, M, j = 1, \dots, C$ . Za akustične izvore u okolini niza mikrofona koeficijenti se primenjuju radi kompenzovanja različitih akustičnih pojava. Na primer, signal koji je dobijen na udaljenijem senzoru može se pojačati više od onog koji je primljen na senzoru bližem izvoru zvuka kako bi se eliminisala razlika nastala kao posledica opadanja amplitude sa rastojanjem. Bez obzira na izbor težinskih koeficijenata, ako je u odgovarajućem čvoru prisutan izvor zvuka dolazi do konstruktivne interferencije signala što za posledicu ima pojačavanje amplitude izlaza. U protivnom se signali međusobno poništavaju što govori o odsustvu izvora zvuka. Sumiranje kvadrata izlaza (6.6) na čitavoj dužini digitalnog signala,  $n = 1, \dots, N$ , daje **snagu usmerenog odziva** (engl. steered-power response)

$$P(\vec{\rho}_j) = \sum_{n=0}^{N-1} y(n, \vec{\rho}_j)^2 \quad (6.7)$$

Za najbolju aproksimaciju lokacije izvora usvaja se onaj čvor koji ima najveću energiju usmerenog odziva

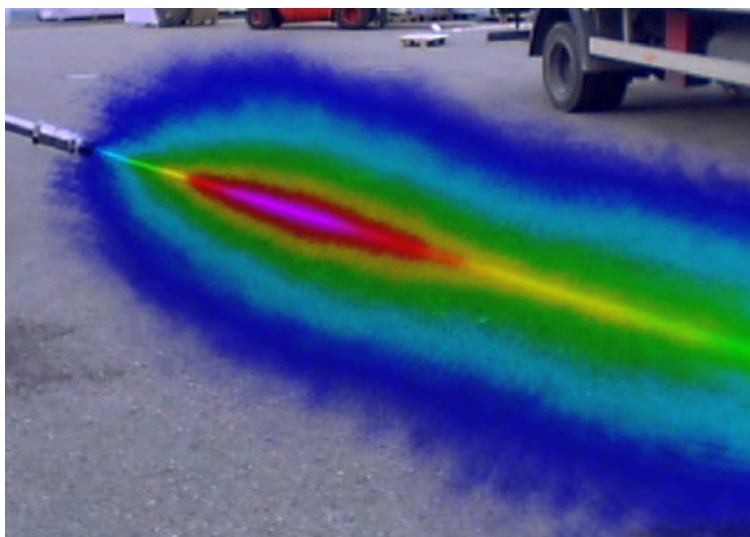
$$\vec{\rho}_l = \arg \max_j (P(\vec{\rho}_j)) \quad (6.8)$$

Kao rezultat testiranja svih čvorova dobija se akustična mapa ispitivanog prostora (Slika 6.3) što ovu metodu čini primenjivom za istovremenu lokalizaciju više izvora zvuka kao i za lokalizaciju kontinualnih izvora. Metode je veoma robusna u odnosu na prisustvo poremećaja ali je rezolucija lokalizacije limitirana frekvencijom uzorkovanja akustičnog signala. Razlike u kašnjenjima koje

su manje od perioda uzorkovanja nije moguće registrovati. Kod primene prostornog filtriranja u realnom vremenu rezolucija je ograničena i raspoloživim računskim kapacitetom. Fourier-ovom transformacijom izraza (6.6) procedura formiranja snopa prevodi se iz vremenskog u frekventni domen.

$$Y(\omega, \vec{\rho}_j) = \sum_{i=0}^{M-1} q_{ij} S_i(\omega) e^{-i\omega t_{ij}} \quad (6.9)$$

Dalja procedura dobijanja snage usmerenog odziva, koja se sastoji u sumiranju izlaza svih mikrofona nad čitavim spektrom kružnih frekvencija, slična je kao kod vremenskog domena [98,99].



*Slika 6.3 Snimak vodenog mlaza, mašine za pranje pod visokim pritiskom, primenom akustične kamere koja funkcioniše po metodi prostornog filtriranja [100]*

#### 6.2.4. Akustična lokalizacija na osnovu vremena prostiranja talasa

Najprecizniji rezultati akustične lokalizacije dobijaju se merenjem **vremena prispeća** (engl. Time of Arrival – TOA) ili **razlike u vremenu prispeća** (engl. Time Difference of Arrival – TDOA) signala prikupljenih na sensorima koji su locirani na različitim geometrijskim pozicijama. Pod pretpostavkom da se akustični signal prostire kroz nedisperzivni homogeni medijum konstantnom brzinom u svim pravcima, akustični izvor može se smatrati jezgrom sferičnog talasa. Ovakav model prostiranja zvuka naziva se sferičnim ili modelom **bliskog polja**.

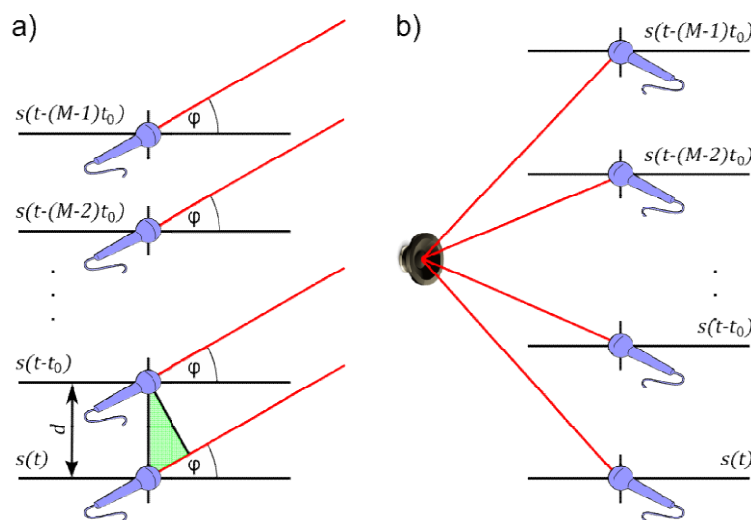
Ako je izvor zvuka dovoljno udaljen, front akustičnog talasa aproksimira ravnim frontom čime se dobija model prostiranja zvuka u **dalekom polju**. U literaturi je usvojen kriterijum rastojanja,  $r_0 > \frac{2D^2}{\lambda_{min}}$ , za modeliranje zvuka u dalekom polju [101], gde je  $D$  dimenzija niza mikrofona a  $\lambda_{min}$  minimalna talasna dužina zvuka. Uprkos uspostavljenoj podeli između akustične lokalizacije u bliskom i dalekom polju postoje algoritmi koji teže da pokriju obe situacije [102].

#### 6.2.4.1.1. Akustična lokalizacija u dalekom polju

Kod lokalizacije u dalekom polju brzine zvuka u svim tačkama fronta su paralelne (Slika 6.4 a). Na osnovu osenčenog trougla jednostavno je odrediti ugao dopiranja zvuka (engl. Angle of Arrival – AOA) [103]

$$\varphi = \arcsin \left( \frac{t_0 c_0}{d} \right) \quad (6.10)$$

Za određivanje ugla dovoljan je par senzora ali se u praksi koristi niz ekvidistantnih senzora. Rezultujući pravac dobija se uprosečavanjem rezultata dobijenih na parovima senzora. Ovako se povećava tačnost rezultata i robustnost u odnosu na prisustvo poremećaja. Nedostatak ove metode je što ne pruža informaciju o udaljenosti izvora. Ovaj problem se može prevazići primenom dva ili više niza senzora postavljenih na različitim pozicijama [104]. U tom slučaju svaki od njih određuje po jedan pravac dopiranja zvuka a izvor se traži u preseku ovih pravaca.



Slika 6.4 Akustična lokalizacija: a) daleko polje, b) blisko polje

#### 6.2.4.2. Akustična lokalizacija u bliskom polju

Lokalizacija u bliskom polju izvora zvuka (Slika 6.4 b) je daleko kompleksnija jer svaki od senzora prima akustični signal pod drugačijim uglom. U zavisnosti od tipa merenja razlikuju se sistemi bazirani na **pređenom rastojanju**, **sumi pređenog rastojanja** i sistemi bazirani na **razlici u pređenom rastojanju** između signala primljenih na sensorima. Svaku od navedenih vrsta karakterišu posebne jednačine odnosno drugačija geometrijska lokalizacija traženog rešenja.

##### 6.2.4.2.1. Sistemi bazirani na apsolutnom vremenu prostiranja zvuka

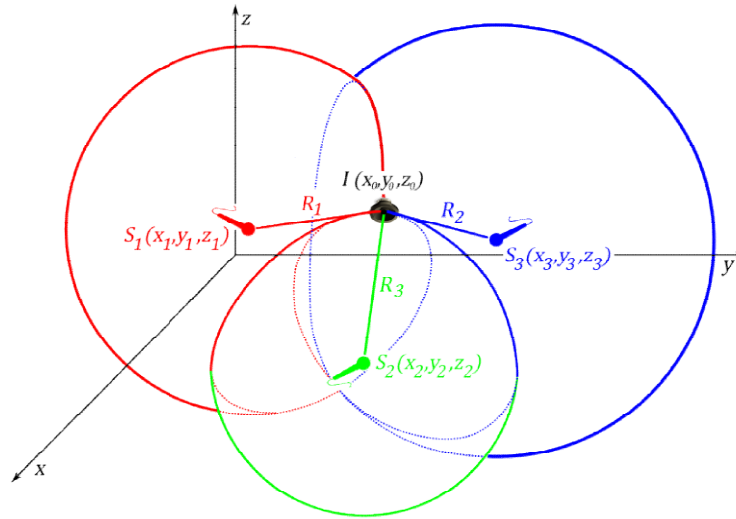
Sistemi bazirani na rastojanju koriste apsolutno vreme,  $t_i$ , potrebno da akustični talas pređe put od izvora,  $I$ , do senzora  $S_i$  za izračunavanje dužine ovog puta

$$R_i = c_0 t_i \quad (6.11)$$

Ovo rastojanje izraženo u funkciji koordinata izvora i koordinata akustičnog senzora iznosi

$$R_i = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2 + (z_i - z_0)^2} \quad (6.12)$$

gde su  $x_i, y_i, z_i$ , koordinate senzora,  $S_i$ , a  $x_0, y_0, z_0$ , koordinate akustičnog izvora,  $I$ . Na osnovu (6.11) i (6.12) očigledno je da informacija o apsolutnom vremenu prostiranja signala do senzora lokalizuje poziciju izvora na površinu sfere prečnika  $R_i$  sa centrom u tom senzoru. Zbog toga se ova metoda pozicioniranja naziva još i **sferična lokalizacija**. Primenom najmanje tri senzora, od kojih svaki daje jednu jednačinu dobija se sistem od tri jednačine sa tri nepoznate koji je dovoljan za izračunavanje prostorne pozicije izvora zvuka. U opštem slučaju, kao rezultat primene minimalne konfiguracije, dobijaju se dve prostorne pozicije izvora što je posledica kvadratnih članova u jednačinama sfera [105]. U odsustvu a priori informacije, koja bi otklonila dvosmislenost rezultata, izbor se može izvršiti uvođenjem četvrtog senzora. Ovaj, i svaki naredni senzor, čine sistem jednačina preodređenim čime nalaženje rešenja postaje nelinearni optimizacioni problem. U takvom slučaju, traži se prihvatljivo aproksimativno rešenje na osnovu nekog od kriterijuma kvaliteta, najčešće greške, sa ciljem njene optimizacije. Najveća preciznost postiže se kada se sfere, koje definišu geometrijsko mesto izvora, seku pod pravim uglom. Kako se ugao preseka udaljava od pravog tako se povećava devijacija dobijenih rezultata.



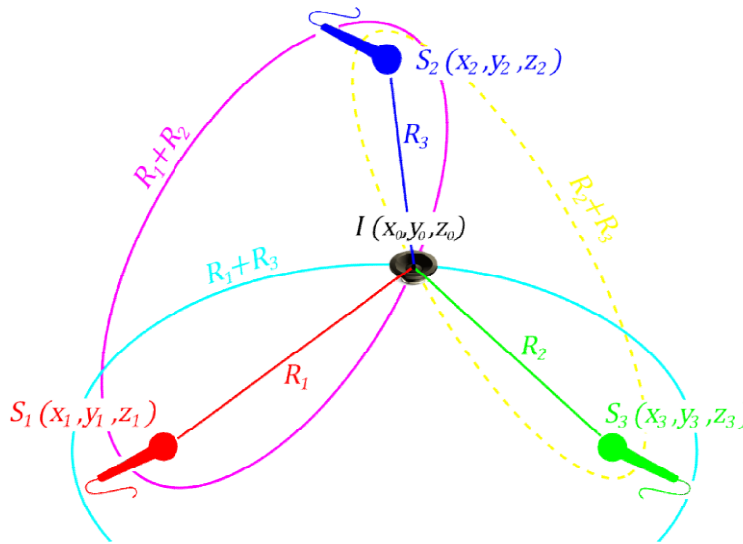
Slika 6.5 Lokalizacija na osnovu vremena prostiranja signala

Sistemi baziranih na sumi pređenih rastojanja, kao informaciju na osnovu koje se vrši akustična lokalizacija koriste zbir apsolutnih vremena prostiranja zvuka za par senzora

$$R_{ij} = R_i + R_j = c_0(t_i + t_j) \quad (6.13)$$

gde je  $R_i$  rastojanje od akustičnog izvora do senzora  $S_i$  a  $R_j$  rastojanje od izvora zvuka do senzora  $S_j$ . Ako se ova rastojanja izraze u funkciji geometrijskih relacija između senzora i mikrofona (6.12), jednačina (6.13) redukuje moguću lokaciju izvora zvuka na elipsoid (tačnije sferoid) sa žižama u akustičnim sensorima  $S_i$  i  $S_j$ . Zbog toga se ova vrsta lokalizacije naziva još i **eliptička lokalizacija**. Na slici 6.6 prikazan je slučaj dvodimenzionalne eliptičke lokalizacije sa minimalnom konfiguracijom koju čine 3 mikrofona. Kod dvodimenzionalne lokalizacije paru mikrofona, na osnovu jednačina (6.12) i (6.13), odgovara elipsa. U preseku dveju elipsa,  $R_1 + R_2$  i  $R_1 + R_3$  dobija se dualno rešenje pa je za jednoznačno određivanje pozicije akustičnog izvora potrebna još jedna informacija. U nedostatku ove informacije može se iskoristiti preostali par mikrofona. Na osnovu iste logike, jednoznačna eliptička lokalizacija u trodimenzionalnom prostoru može se izvršiti sa minimalno 4 akustična senzora koji ne smeju biti u ravni.

Osnovni nedostatak sferične i eliptičke lokalizacije je što što zahtevaju preciznu sinhronizaciju između akustičnog izvora i senzora kako bi vreme prostiranja zvuka između njih bilo moguće registrovati.



Slika 6.6 Eliptička lokalizacija u dve dimenzije

#### 6.2.4.3. Sistemi bazirani na razlici u vremenu pristizanja signala

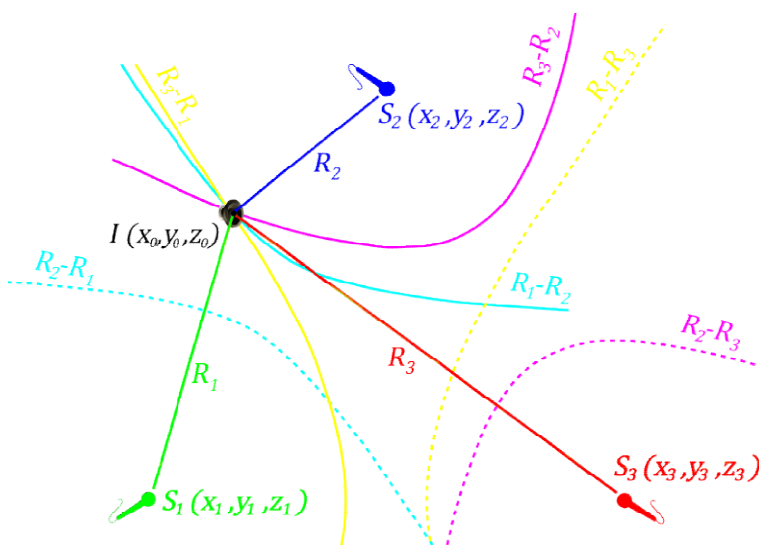
Sistemi bazirani na razlici udaljenosti akustičnih senzora od izvora zvuka koriste vremensku razliku između prispeća signala za određivanje pozicije izvora zvuka

$$R_{ij} = R_i - R_j = c_0(t_i - t_j) = \Delta t_{ij}c_0 \quad (6.14)$$

gde je  $\Delta t_{ij}$  vremensko kašnjenje signala registrovanog na senzoru  $S_i$  u odnosu na signal registrovan na senzoru  $S_j$ . Ako se u jednačini (6.14) rastojanja  $R_i$  i  $R_j$  izraze u funkciji geometrijskih relacija između akustičnih komponenti (6.12) dobija se jednačina rotacionog hiperboloida sa akustičnim sensorima u tačkama fokusa [106]. Zbog toga je metod lokalizacije na osnovu kašnjenja signala poznat još pod nazivom **hiperbolička lokalizacija**. Slika 6.7 prikazuje hiperboličku lokalizaciju u dvodimenzionalnom prostoru sa minimalnom konfiguracijom koju čine tri senzora. U dvodimenzionalnom prostoru jednačina (6.14) definiše hiperbolu sa sensorom u njenom fokusu. U opštem slučaju, presek dveju hiperbola daje dualno rešenje koje se može prevazići na isti način kao kod prethodne metode. Minimalna konfiguracija za prostornu lokalizaciju metodom hiperboličke lokalizacije je 4 prostorno raspoređena senzora.

Hiperbolička lokalizacija se izdvaja po jednostavnosti realizacije jer ne zahteva apsolutno vreme prostiranja talasa od izvora zvuka do senzora. Kada se radi o nepoznatim ili nepristupačnim izvorima, daleko je praktičnije registrovati kašnjenje između signala zbog čega

ovaj pristup dominira u praksi akustične lokalizacije u bliskom polju. Iz istog razloga, hiperbolička lokalizacija je stavljena u fokus ove disertacije.



Slika 6.7 Hiperbolička lokalizacija u dve dimenzije

Rešavanje sistema hiperboličkih jednačina (6.14) nije trivijalan problem čak o ako se radi o minimalnoj konfiguraciji senzora. Za ovu svrhu je, u literaturi, predložen veliki broj algoritama različite namene i različite kompleksnosti. Metod Newton – Kantorovich-a [107] obezbeđuje tačno i robusno rešenje preodređenog sistema ali sa velikim računskim zahtevima. Primenom metoda gradijenta ovaj problem može biti delimično prevaziđen [108] jer nije potrebno izračunavati inverznu matricu. Međutim, ovaj pristup zahteva dobru pretpostavku početnog stanja i ne garantuje konvergenciju u prisustvu velikih poremećaja. Imajući u vidu navedene nedostatke, ušteda u računskoj kompleksnosti primenom metoda gradijenta može se lako ispostaviti kao neracionalna u odnosu na metodu Newton – Kantorovich-a. Algoritam Levenberg–Marquardt-a [109], takođe poznat pod nazivom metod prigušenih najmanjih kvadrata, kombinuje prednosti metode gradijenta i Gauss-Newton-ovog algoritma. Problem je velika mogućnost da zaglavi u zoni lokalnog minimuma. Pod uslovom da je broj akustičnih senzora jednak broju hiperboličkih jednačina (6.14) metod Fang-a [110] obezbeđuje tačno rešenje sa manje računskih operacija u odnosu na iterativne metode. Chan je predložio metod [101] koji funkcioniše u mešovitom, bliskom i dalekom akustičnom polju. Još jedna prednost, u odnosu na metod Fang-a, je što ovaj pristup pruža mogućnost povećanja tačnosti primenom redundantnih senzora dok je osnovna mana osetljivost na uticaj poremećaja.

#### 6.2.4.4. Izračunavanje kašnjenja među signalima

Diskretni akustični signali primljeni na dva prostorno odvojena senzora mogu se modelirati sledećim jednačinama

$$s_i[k] = s_0[k] + n_i[k] \quad (6.15)$$

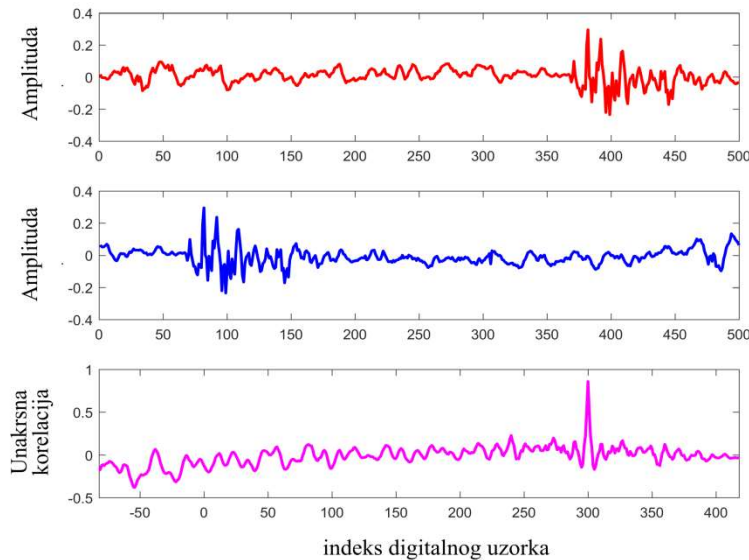
$$s_j[k] = \alpha_{ij}s_0[k - l] + n_j[k] \quad (6.16)$$

gde je  $k$  indeks digitalnog uzorka,  $s_0[k]$  signal emitera (izvora),  $n_i[k]$  i  $n_j[k]$  su nekorelisani slučajni procesi sa Gausovom raspodelom [111],  $\alpha_{ij}$  je koeficijent koji modelira razliku u amplitudi između ova dva signala a  $l$  je kašnjenje između signala izraženo u periodima uzorkovanja,  $T_s$ . Bazična metoda za određivanje kašnjenja između ovih signala je na osnovu funkcije **unakrsne korelacije**

$$R_{ij}(l) = \frac{1}{K} \sum_{k=0}^{K-1} s_i[k] s_j[k - l] \quad (6.17)$$

gde  $K$  predstavlja dužinu digitalnog uzorka, kao argument, u okviru domena mogućih kašnjenja, koji maksimizuje njenu vrednost

$$\Delta t_{ij} = \frac{1}{f_s} \arg \max(R_{ij}[l]), \quad -\frac{T}{2} \leq l \leq \frac{T}{2} \quad (6.18)$$



Slika 6.8 Grafici dva akustična signala i njihova unakrsna korelacija



gde je  $f_s$  frekvencija uzorkovanja digitalnog signala a  $T$  veličina opsega opservacije, izražena u periodima diskretizacije. Dobra aproksimacija korelacije može se dobiti inverznom diskretnom Fourier-ovom transformacijom izraza

$$R_{ij}(l) \approx \frac{1}{K} \sum_{k=0}^{K-1} R_{ij}(f) e^{\frac{j2\pi fl}{K}} \quad (6.19)$$

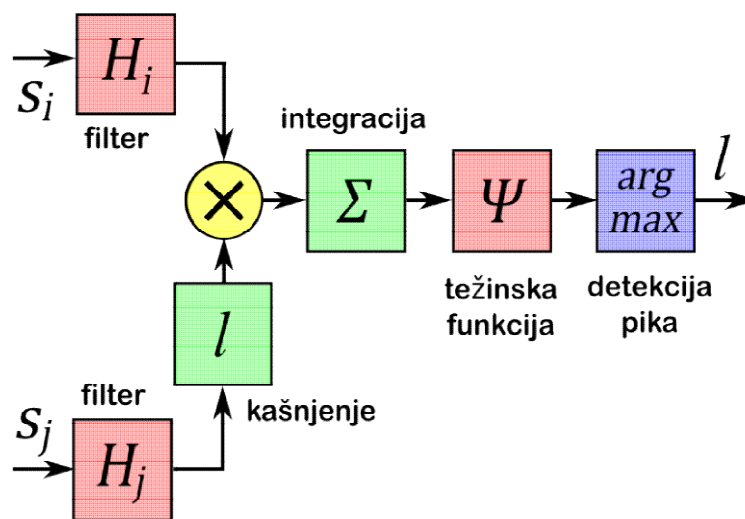
gde je  $R_{ij}(f)$  unakrsna spektralna gustina snage

$$R_{ij}(f) = S_i(f) \bar{S}_j(f) \quad (6.20)$$

U izrazu (6.19),  $S_i(f)$  i  $S_j(f)$  predstavljaju frekventnu reprezentaciju signala  $s_i[k]$  i  $s_j[k]$  a nadvučena crta reprezentuje operator konjugacije kompleksnog broja. Ako se u zbiru (6.19) umetne **težinska funkcija**,  $\Psi(f)$ , poznata drugačije pod nazivom **funkcija prefiltriranja** ili **funkcija prozora**, dobija se uopštenu algoritam unakrsne korelacije

$$R_{ij}(l) \approx \frac{1}{K} \sum_{k=0}^{K-1} \Psi(f) R_{ij}(f) e^{\frac{j2\pi fl}{K}} \quad (6.19)$$

Blok dijagram izvođenja uopštene unakrsne korelacije predstavljen je na slici na Slici 6.8. U odnosu na osnovnu metodu unakrsne korelacije, ovde su dodati filteri, sa prenosnom funkcijom  $H_i$  i  $H_j$ , koji služe za predobradu signala u cilju povećanja robusnosti postupka.



Slika 6.9 Blok dijagram uopštene unakrsne korelacije

U zavisnosti od vrste težinske funkcije razlikuju se različite metode (varijante) uopštene unakrsne korelacije od kojih svaka ima svoje specifičnosti i svoju namenu. Tabela 6-1 prikazuje u praksi najčešće upotrebljavane funkcije prozora. Ako nema težinske funkcije, odnosno njena vrednost je jednaka jedinici, radi se o osnovnoj metodi unakrsne korelacije, bez ikakve predobrade signala. Zbog toga je ova metoda osetljiva na poremećaje i eho signala. Funkcija Wiener-a, nazvana po istoimenoj teoriji optimalne procene (Norbert Wiener, 1894 –1964), pojačava signal proporcionalno relativnom odnosu signala i poremećaja sa ciljem njegovog povećanja [112]. Ovaj odnos ocenjuje se na osnovu funkcije koherencije,  $\gamma_{ij}^2$ , čija je vrednost takođe data u tabeli. Funkcija koju je predložio Roth [113] onda funkcija maksimalne verodostojnosti [114] i transformacija izravnate koherencije (engl. Smoothed Coherence Transform – SCOT) [115] imaju sličan efekat prigušivanja delova spektra sa velikim sadržajem poremećaja u kojima je verovatnije da dođe do greške. Transformacija izravnate koherencije uvedena je iz razloga da umanjí uticaj jakih tonova ali u literaturi postoje dokazi da smanjuje tačnost uopštene unakrsne korelacije čak i u odnosu na osnovni metod [116] Algoritam transformacije faze (engl. Phase Transform – PHAT) uklanja sadržaj energije iz unakrsnog spektra čime obezbeđuje oštar pik unakrsne korelacije [117].

**Tabela 6-1** Težinske funkcije uopštene unakrsne korelacije

Naziv metode	Težinska funkcija
Kvadratni prozor (Cross correlation)	$\Psi_{CC} = 1$
Wiener-ov prozor	$\Psi_{Weiner} = \gamma_{ij}^2(f)$
Roth-ov prozor	$\Psi_{Roth} = 1/S_{ii}(f)$
PHAT prozor (Phase Transform)	$\Psi_{PHAT} = 1/ S_{ij}(f) $
SCOT prozor (Smoothed Coherence Factor)	$\Psi_{SCOT} = 1/\sqrt{S_{ii}(f)S_{jj}(f)}$
ML prozor (Maximum Likelihood)	$\Psi_{ML} = \gamma_{ij}(f)/\left(\left(1 - \gamma_{ij}(f)\right) S_{ij}(f) \right)$
Koherencija	$\gamma_{ij}^2(f) =  S_{ij}(f) ^2/(S_{ii}(f)S_{jj}(f))$

Osim unakrsne korelacije kao kriterijum za određivanje kašnjenja može se primeniti funkcija **srednje kvadratna razlike** (engl. average square difference function – ASDF)

$$D_{ij}(l) \approx \frac{1}{K} \sum_{k=0}^{K-1} (s_i[k] - s_j[k-l])^2 \quad (6.19)$$

Na osnovu koje se nalazi kašnjenje kao argument, u okviru domena mogućih kašnjenja, koji minimizuje njenu vrednost

$$\Delta t_{ij} = \frac{1}{f_s} \arg \min(D_{ij}[l]), \quad -\frac{T}{2} \leq l \leq \frac{T}{2} \quad (6.20)$$

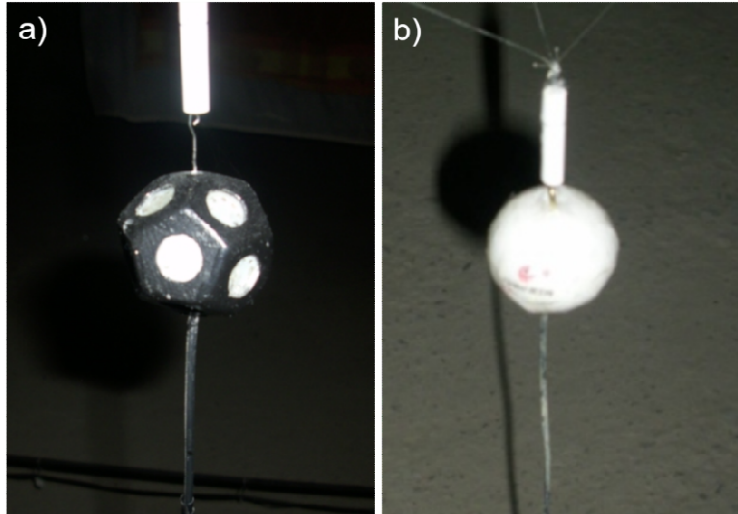
Prednost ovog algoritma je manja računaska kompleksnost u odnosu na unakrsnu korelaciju jer nema množenja između članova. Daje perfektu ocenu kašnjenja ali je izuzetno osetljiv na prisustvo poremećaja [118].

### 6.3. Eksperiment akustične lokalizacije u bliskom polju

Radi potvrde teoretskih zapažanja izvršeno je eksperimentalno istraživanje [119] sa zadatkom dizajniranja sistema veštačke inteligencije za određivanje prostornih koordinata akustičnog izvora u realnim uslovima, metodom nadgledanog učenja. U skladu sa zadatkom, postavljen je zahtev robustnosti postupka u odnosu na prisustvo poremećaja dok je računaska kompleksnost ograničena kako ne bi ugrozila primenu u realnom vremenu. I ovde se insistiralo na univerzalnosti postupka kako bi bio primenjiv u što većem spektru različitih situacija i od strane najšireg mogućeg kruga korisnika.

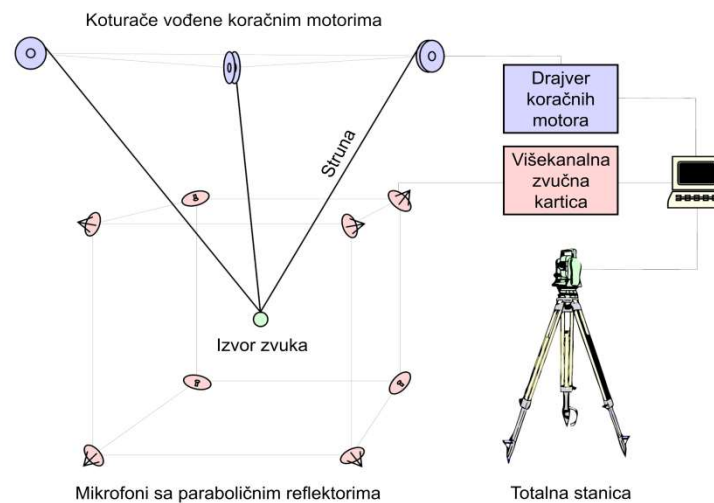
#### 6.3.1. Prikupljanje akustičnih uzoraka

Da bi akustična lokalizacija bila objektivna akustični izvor bi trebao da bude omnidirekciono odnosno da šalje zvuk istog intenziteta u svim pravcima. U eksperimentu je to ostvareno pomoću dva različita modela, približno istih gabarita (oko 40 mm). Prvi model (Slika 6.9 a) sačinjen je od 12. paralelno vezanih. minijturnih zvučnika (prečnika 16 mm), svaki smešten u centru stranice pravilnog dodekahedrona. Drugi izvor (Slika 6.9 b) sastavljen je od ping-pong loptice u kojoj je smešten minijturni vibracioni motor tako da, udarajući u njene zidove, proizvodi akustične vibracije. Inicijalna ispitivanja su pokazala da izbor između ova dva akustična izvora nije imao primetan uticaj na registrovano kašnjenje. Zbog toga je eksperiment nastavljen isključivo primenom akustičnog izvora sa ping-pong lopticom zbog njegove manje mase i jednostavnijeg upravljanja.



Slika 6.10 Dve varijante omnidirekcionog izvora zvuka

Radi prikupljanja akustičnih uzoraka, izvor zvuka je navođen, unutar prostora oblika kocke sa dužinom stranice od 1,6 m, okačen o tri strune. Drugi kraj svake od struna bio je namotan na koturaču vođenu koračnim motorom koja se slobodno okreće oko vertikalne ose. Tri ovakve koturače smeštene su iznad prostora predviđenog za kretanje izvora, u temenima horizontalnog jednakostraničnog trougla sa dužom stranice od 4,35 m, gradeći jednostavan mehanizam za prostorno pozicioniranje (Slika 6.10). Smislenim zamotavanjem i odmotavanjem struna postignuta je bilo koja zadata pozicija unutar predviđenog prostora.



Slika 6.11 Akvizicija akustičnih signala namenjenih lokalizaciji u bliskom polju

Potrebna dužina strune, za postizanje zadate geometrijske pozicije zvuka  $S(x_0, y_0, z_0)$ , određena je na osnovu vertikalnog preseka koji kroz nju prolazi. Ovaj presek u dve različite perspektive i razmere prikazan je na slici 6.12. Dužina strune od tačke iznad izvora zvuka, u kojoj je za njege fiksirana  $S_0$ , do tačke na obodu koturače,  $Q$ , tačno iznad njenog centra, izračunata je kao zbir dveju dužina

$$l_k = \sqrt{R_k^2 - r^2} + \left(\frac{\pi}{2} - \varphi_k\right) r, \quad k = 1, 2, 3 \quad (6.21)$$

gde je  $k$  indeks koturače,  $r$  njen poluprečnik,  $R_k$  rastojanje od izvora zvuka do preseka ose koturače sa pomenutom vertikalnom ravni (Slika 6.12 a)

$$R_k = \sqrt{(x_k - x_0)^2 + (y_k - y_0)^2 + (z_k - z_0)^2} \quad (6.22)$$

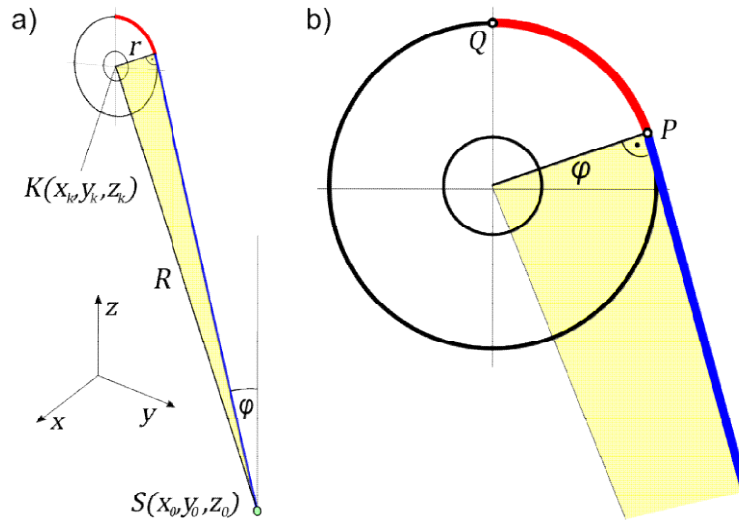
a  $\varphi_k$  ugao koji struna zaklapa sa vertikalnom osom koja prolazi kroz izvor zvuka (Slika 6.12 b)

$$\varphi_k = \arcsin\left(\frac{z_k - z_0}{R_k}\right) - \arctan\left(\frac{r}{R_k}\right) \quad (6.23)$$

Prvi član zbira (6.21), na slici označen plavom linijom, jednak je dužini odmotanog dela strune, od izvora zvuka,  $S_0$ , do tačke  $P$  na obodu koturače koja deli odmotani (pravolinijski) od zamotanog (kružnog) dela strune. Drugi član, na slici označen crvenom linijom, jednak je dužini zamotanog dela strune od tačke  $P$  pa do tačke  $Q$ . Jednačina (6.21) omogućava izračunavanje broja koraka step motora potrebnih da se izvor zvuka iz trenutne, poznate, pozicije dovede u narednu poziciju

$$N_k[n] = k_m(l_k[n] - l_k[n - 1]), \quad k = 1, 2, 3 \quad (6.24)$$

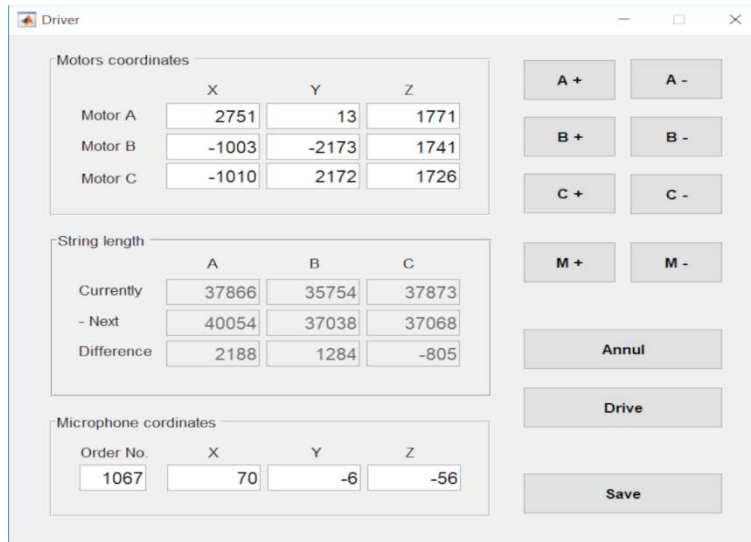
gde je  $n$  indeks zadatog položaja akustičnog izvora a  $k_m$  konstanta pojačanja sklopa sačinjenog od koračnog motora i koturače. Ova konstanta jednaka je broju potrebnih koraka motora za odmotavanje jedinične dužine strune. Konstanta  $k_m$  u praksi varira kao posledica neravnomernog namotavanja strune u radijalnom i aksijalnom pravcu. Međutim, na osnovu inicijalnih ispitivanja, utvrđeno je da su ove varijacije zanemarljive, na predviđenom opsegu rastojanja, pa je u narednim fazama eksperimenta primenjena konstantna vrednost,  $k_m = 12,743 \text{ mm}^{-1}$ .



Slika 6.12 Vertikalni presek eksperimentalne postavke kroz strunu

Da bi jednačina (6.24) mogla da bude primenjena izvor zvuka je morao da bude precizno pozicioniran u početni položaj. Za merenje geometrijskih relacija između komponenti eksperimentalne opreme primenjena je Totalna Stanica Sokkia SET630R. Instrument omogućava lasersko merenje prostornih koordinata, sa tačnošću od  $\pm 3\text{mm}$  na primenjenom opsegu rastojanja, memorisanje i prenos u računar preko RS-232 porta. Nakon upoređivanja realizovanih sa zadatim pozicijama, na oko 3000 uzoraka, prosečna apsolutna greška koju je napravio mehanizam za pozicioniranje akustičnog izvora bila je nešto manja od 10 mm. Ova greška ocenjena je kao prihvatljiva s obzirom da je višestruko manja (oko 4 puta) od dimenzije akustičnog izvora.

Radi upravljanja procesom akvizicije signala programiran je drajver (slika 6.13) sa zadatkom upravljanja procesom prikupljanja akustičnih signala. Drajver je imao funkciju da sukcesivno čita koordinate zadatih tačaka i na osnovu njih izračunava potrebnu korekciju položaja koračnih motora (6.24). Informaciju o korekciji drajver je prosleđivao mikrokontroleru ATmega328P, preko USB porta, koji je dalje navodio koračne motore (NEMA 23LM-C005-18) putem modula za upravljanje koračnim motorima (A4988). Nakon postizanja svake od zadatih pozicija, drajver je aktivirao akustični izvor i beležio signale registrovane na mikrofona, po unapred utvrđenom rasporedu, pre nego krene na naredni položaj. Primenom ovog drajvera proces prikupljanja signala za potrebe akustične lokalizacije u bliskom polju u potpunosti je automatizovan.

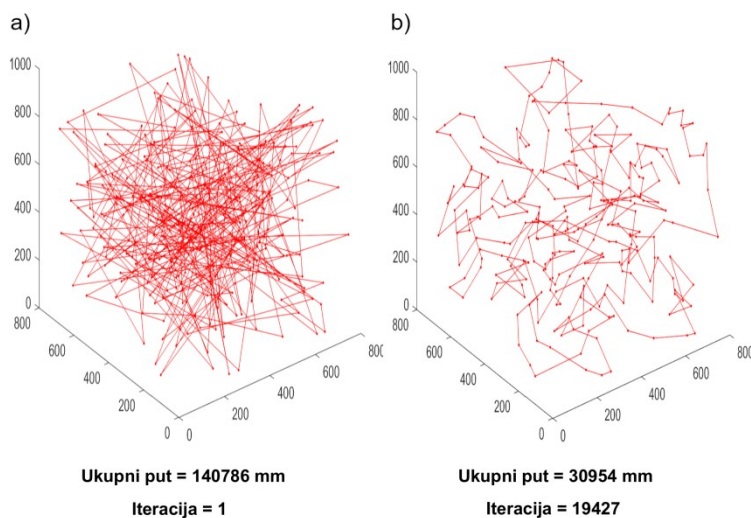


*Slika 6.13 Grafički interfejs drajvera za upravljanje akvizicijom signala za potrebe akustične lokalizacije u bliskom polju*

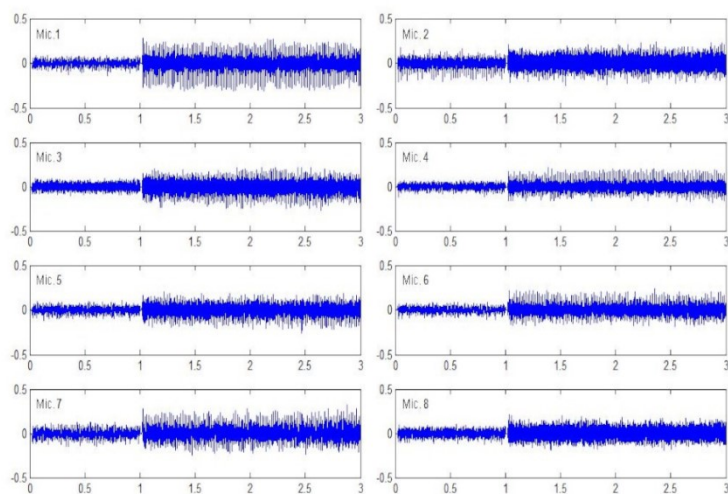
Unutar pomenutog opsega kretanja predviđeno je 2000 slučajno izabranih pozicija akustičnog izvora namenjenih treniranju algoritma lokalizacije. Radi povećanja efikasnosti treniranja izvršena je optimizacija putanje izvora sa ciljem njene minimizacije. U skladu sa temom disertacije, u ovu svrhu primenjen je genetski algoritam kao jedan od najopštijih algoritama veštačke inteligencije. Pošto se minimizacija putanje sa 2000 tačaka, zbog velike računске kompleksnosti, pokazala neefikasnom tačke su grupisane po kvadrantima pa je optimizacija obavljena u svakom od 8 kvadranta posebno. Slika 6.13 prikazuje putanju izvora zvuka, duž slučajno izabranih tačaka koje pripadaju prvom kvadrantu, pre (a) i posle (b) postupka optimizacije. Nakon izvođenja postupka u svih 8 kvadranta, prosečna ušteda u rastojanju iznosila je 4-5 puta. Skraćenje puta uticao je i na smanjenje greške pozicioniranja jer duža putanja izaziva veće varijacije koeficijenta  $k_m$ .

Za registrovanje akustičnih signala predviđeno je 8 jednostavnih, lako dostupnih, minijaturnih mikrofona, koji pripadaju standardnoj opremi za nadzor. Mikrofonu su prostorno raspoređeni u temenima zamišljene kocke, sa dužinom stranice od 2 m, oko prostora pomeranja izvora zvuka, sa namerno uvećanom tolerancijom prostorne pozicije od 20 mm. Iza svakog mikrofona postavljeni su parabolični reflektori, usmereni ka centru prostora pomeranja mikrofona, u funkciji mehaničkih pojačivača akustičnih talasa. Povezivanje mikrofona na računar ostvareno je pomoću osmokalnalne zvučne kartice TerraTec EWS88 MT. U svakoj od

2000 prostornih pozicija zabeležen je signal, na standardnoj frekvenciji uzorkovanja od 44,1 kHz, i u trajanju od 3s. U prvoj sekundi registrovan je samo šum, pre aktiviranja izvora zvuka, dok druge dve sadrže akustični signal registrovan nakon aktiviranja izvora. Slika 6.14 daje primer registrovanih signala koji odgovaraju jednoj od 2000 slučajno izabranih prostornih pozicija akustičnog izvora.



*Slika 6.14 Putanja izvora zvuka u jednom kvadrantu: a) pre optimizovanja, b) nakon optimizovanja*



*Slika 6.15 Primer akustičnih signala istovremeno zabeleženog na 8 kanala*



### 6.3.2. Određivanje kašnjenja između signala

Sa osam prostorno raspoređenih mikrofona svakoj prostornoj poziciji odgovaraju 28 parova mikrofona odnosno 28 različitih kašnjenja (6.18). Za 2126 realizovanih pozicija izvora zuka u eksperimentu lokalizacije to je 59528 različitih kašnjenja. Literatura pruža veliki izbor postupaka za određivanje kašnjenja među signalima ali ne postoji univerzalno rešenje koje bi bilo primenjivo u svim situacijama. U zavisnosti od vrste signala i uslova primene različite metode ispoljavaju svoje prednosti i mane (Poglavlje 6.2.4.4). Sa druge strane, konfiguracija algoritma je pitanje kompromisa između tačnosti i računске kompleksnosti. U cilju očuvanja univerzalnosti postupka, u konkretnom slučaju, pitanje izbora metode ostavljeno je otvoreno. Kao polazna osnova usvojena je unakrsna korelacija sa ograničenim kašnjenjem, bez obzira da li se radi o vremenskoj ili frekventnoj reprezentaciji. Konačan izbor metode i konfiguracija parametra predviđena je u okviru postupka optimizacije sa ciljem minimizacije srednje kvadratne greške rezultujućih kašnjenja. Greška je ocenjivana u odnosu na referentna (teoretska) kašnjenja među signalima koja su izračunata primenom jednačina (6.11), (6.12) i (6.14). Pri izračunavanju referentnih kašnjenja primenjene su realizovane koordinate izvora zvuka utvrđene totalnom stanicom.

Izvođenje unakrsne korelacije (Tabela 6-1) neodvojivo je od predobrade signala koja se preduzimaju sa ciljem povećanja robustnosti sistema u odnosu na prisustvo poremećaja. Zbog toga je postupak određivanja kašnjenja detaljno opisan u narednom poglavlju koje je u potpunosti posvećeno predobradi signala.

### 6.3.3. Rešavanje sistema hiperboličkih jednačina

Na osnovu jednačina (6.12) i (6,14), svako kašnjenje između para signala daje po jednu hiperboličku jednačinu. Prema tome, rezultat procedure određivanje kašnjenja nad 8-kanalnim akustičnim signalom je sistem od 28 hiperboličkih jednačina sa 3 nepoznate. Zbog svoje univerzalnosti, jednostavnosti i brzine, za rešavanje ovog sistema predviđena je neuronska mreža sa direktnim prostiranjem. Mreža je izgrađena od veštačkih neurona sa sigmoidnom aktivacionom funkcijom a trenirana primenom algoritma sa povratnim prostiranjem greške i momentom. Kao ulaz su primenjena raspoloživa kašnjenja između signala što znači da je broj neurona u ulaznom sloju bio određen brojem upotrebljenih kašnjenja (maksimalno 28). Princip

rada neuronske mreže se ne menja znatno sa promenom broja ulaznih terminala što je omogućilo jednostavno eksperimentisanje sa dodavanjem i oduzimanjem redundantnih merenja. Broj izlaznih neurona bio je uvek tri, što odgovara broju prostornih koordinata izvora zvuka. U eksperimentu su korišćene pravouglo koordinata ali neuronska mreža pruža mogućnost jednostavnog formatiranja izlaza. Na primer, preformatiranjem izlaza, u raspoloživim podacima za treniranje, mreža se može obučiti da daje izlaz u polarnim koordinatama. Još jedna prednost predloženog postupka je što ne zahteva koordinate mikrofona. Kao rezultat procedure treniranja, uticaj rasporeda mikrofona na vrednost kašnjenja sadržan je u vrednostima faktora pojačanja neuronske mreže. Svako pomeranje mikrofona zahtevalo bi kompletno novu proceduru obučavanja, počev od akvizicije signala, odnosno nove faktore pojačanja. Konfiguracioni parametri eksperimentalnog postupka, neuronske mreže i algoritma za treniranje, koji nisu bili uslovljeni postavkom, testirani su, nad raspoloživim skupom podataka. Kao rezultat testiranja određeni su optimalni parametri koji obezbeđuju minimalnu grešku lokalizacije.

#### 6.3.4. Rezultati

Performanse akustične lokalizacije prvo su ispitane u pogledu rasporeda mikrofona. Uticaj rasporeda bio je evidentan, naročito za slučaj lokalizacije sa 4 mikrofona. Pošto su mikrofoni bili fiksirani u temenima zamišljenog kubnog prostora različit raspored je postignut uključivanjem i isključivanjem pojedinih mikrofona. Rezultati su dobijeni treniranjem jednoslojne neuronske mreže sa 10 neurona u skrivenom sloju, na bazi 1866 uzoraka za treniranje, nakon 1600 epoha primenom algoritma sa povratnim prostiranjem greške i momentom. U algoritmu je primenjen koeficijent učenja 0,7 i faktor inercije 0,9. Greška treniranja definisana je kao srednja apsolutna greška između realizovanih (stvarnih) koordinata izvora zvuka, izmerenih totalnom stanicom, i pozicija tačaka za treniranje dobijenih kao rezultat postupka akustične lokalizacije. Nakon ispitivanja različitih prostornih rasporeda, sa 4 mikrofona, na osnovu vrednosti greške lokalizacije su se izdvojile 3 grupe (Tabela 6-2). Najveća greška dobijena je primenom mikrofona smeštenih u ravni koja seče eksperimentalni kubni prostor. Nešto manja greška dobijena je primenom mikrofona smeštenih u horizontalnoj ili vertikalnoj ravni koja dodiruje kubni prostor. Značajno manja greška od prethodnih dobijena je primenom konfiguracije mikrofona prostorno raspoređenih izvan ravni. Logično je pretpostaviti da bi standardni, linearni raspored mikrofona pružio ekstremno loše rezultate lokalizacije. Na osnovu prethodnog je

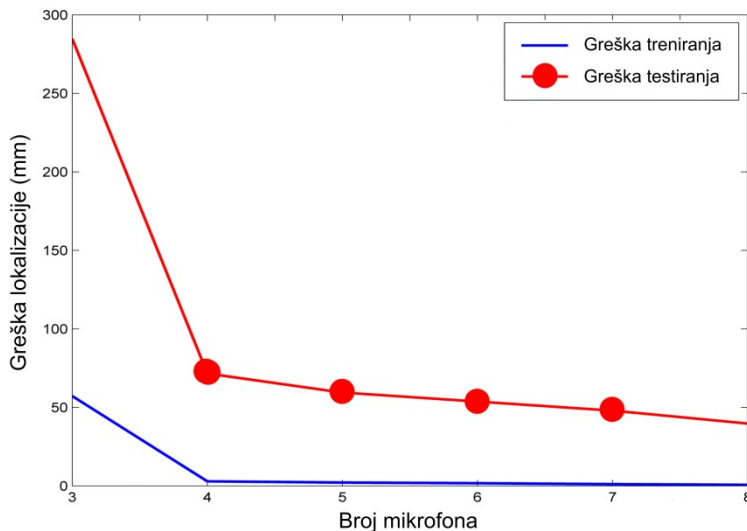
zaključeno da mikrofoni moraju da budu ravnomerno raspoređeni u prostoru kako bi pružili maksimalnu tačnost prostorne lokalizacije u bliskom polju.

**Tabela 6-2** Srednje kvadratna greška u zavisnosti od prostornog rasporeda mikrofona

Raspored mikrofona	Greška treniranja	Greška testiranja
Horizontala/Vertikalna ravan	53 mm	305 mm
Kosa ravan	64 mm	329 mm
Izvan ravni	3 mm	72 mm

Zavisnost greške akustične lokalizacije od broja iskorišćenih mikrofona za registrovanje signala prikazana je na slici 6.16. Rezultati su dobijeni primenom istih konfiguracionih parametara, kao u prethodnom slučaju, i najpovoljnijeg prostornog rasporeda za svaki broj mikrofona. Oni pokazuju prirodnu tendenciju opadanja grešaka testiranja i treniranja sa povećanjem broja senzora. Sa druge strane, broj raspoloživih kašnjenja, samim tim i računaska kompleksnost postupka, raste sa povećanjem broja senzora,  $n_s$ , po kvadratnom zakonu

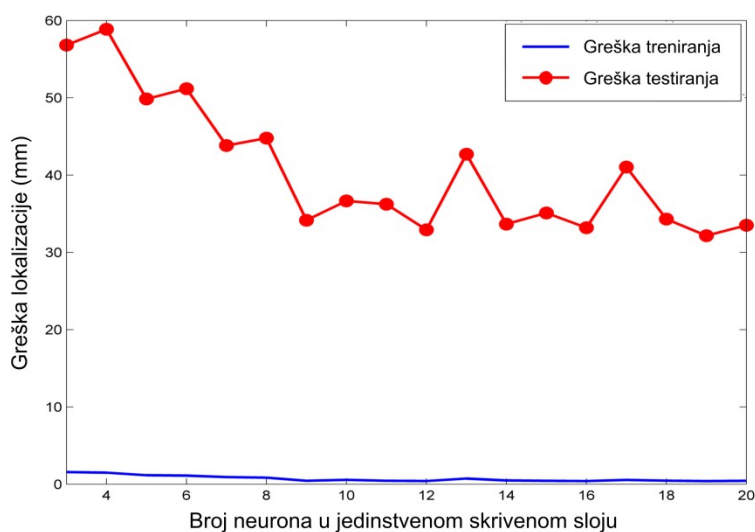
$$N_t = \frac{N_s(N_s - 1)}{2} \quad (6.25)$$



*Slika 6.16 Srednje apsolutna greška lokalizacije u zavisnosti od broja mikrofona*

Stoga, optimalni broj senzora za izvođenje akustične lokalizacije mora biti kompromis između željene tačnosti i dozvoljene računске kompleksnosti postupka.

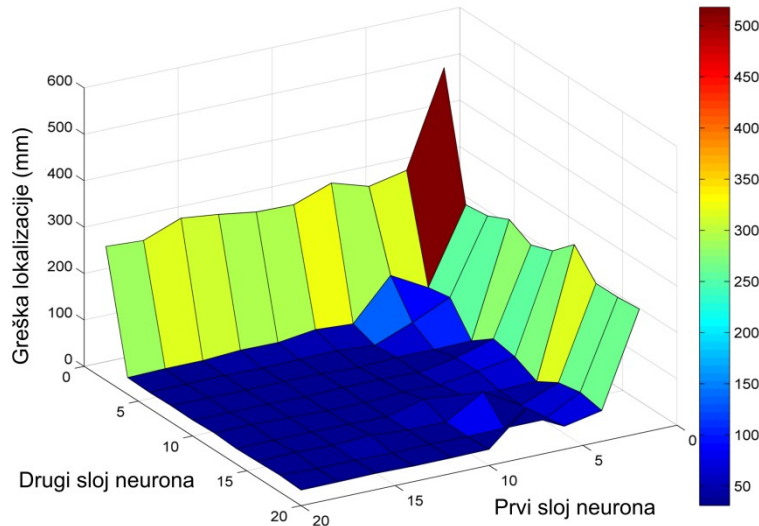
Nakon konfiguracije senzora, predmet testiranja bila je konfiguracije neuronske mreže. Kako je broj neurona u ulaznom sloju bio uslovljen brojem mikrofona (6.25) a broj neurona u izlaznom sloju brojem prostornih koordinata, ostalo je da se odredi konfiguracije skrivenih slojeva. Na osnovu opšteg pravila optimalni broj slojeva i broja veštačkih neurona u njima postignut je kada se kvalitet modela ne povećava sa dodavanjem novih elemenata (slojeva i neurona) [120]. U skladu sa ovim, testiranje je započeto primenom jednoslojne neuronske mreže. Kao kriterijum za ocenu kvaliteta konfiguracije usvojena je srednje apsolutna greška testiranja. Grafici na slici 6.17 predstavljaju srednje apsolutnu grešku treniranja i srednje apsolutnu grešku testiranja u zavisnosti od broja neurona u jedinstvenom skrivenom sloju. Rezultati su dobijeni sa 8 mikrofona, nakon 5000 epoha treniranja i ostalim parametrima nepromenjenim u odnosu na prethodno testiranje. Na osnovu grafika, usvojena je konfiguracija od 10 neurona jer dalje povećanje njihovog broja nije dovođilo do primetnog unapređenja tačnosti lokalizacije.



*Slika 6.17 Srednje apsolutna greška lokalizacije u zavisnosti od broja neurona u jedinstvenom skrivenom sloju*

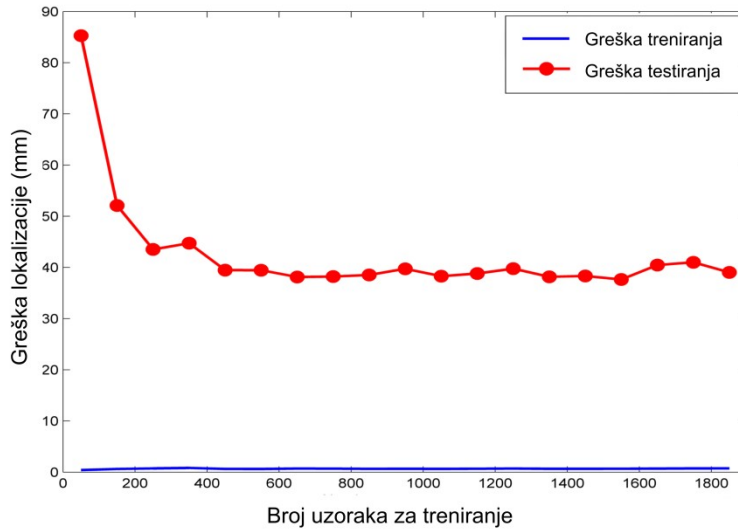
Sledeći pokušaj je bio uvođenje novog skrivenog sloja. Zavisnost srednje apsolutne greške testiranja u zavisnosti od broja neurona u dva skrivena sloja prikazana je na slici 6.18 a odgovarajući grafik srednje apsolutne greške treniranja imao je vrlo sličan oblik. Dvoslojna

konfiguracija skrivenog dela mreže nije donela unapređenje tačnosti u poređenju sa jednoslojnom neuronskom mrežom sa istim brojem neurona. Zbog toga je jednoslojna konfiguracija sa 10 veštačkih neurona usvojena kao optimalna.



*Slika 6.18 Srednje apsolutna greška lokalizacije u zavisnosti od broja neurona u dva skrivena sloja*

Grafik na slici 6.19, dobijen sa nepromenjenim konfiguracionim parametrima i usvojenom konfiguracijom mreže, prikazuje srednja apsolutnu grešku treniranja i srednje apsolutnu grešku treniranja u zavisnosti od broja uzoraka koji su upotrebljeni u proceduru treniranja. Optimalni broj uzoraka generalno zavisi od broja neurona i kompleksnosti mreže. Veća gustina uzoraka obezbeđuje bolju tačnost aproksimacije ali povećava računsku kompleksnost procedure treniranja. U literaturi postoji preporuka da broj podataka za treniranje mora da bude bar dva puta veći od broja faktora pojačanja koji figurišu u mreži [120]. Prema ovom kriterijumu za treniranje usvojene konfiguracije neuronske mreže, sa 10 mikrofona u skrivenom sloju i 8 mikrofona, koja ima 310 faktora pojačanja, potrebno nešto preko 600 uzoraka da bi bila kvalitetno obučena. Greška testiranja, predstavljena na slici 6.19, potvrđuje da 600 podataka za treniranje obezbeđuje dostizanje optimalne tačnosti postupka lokalizacije. Dalje povećanje gustine uzoraka nije davalo приметno smanjenje greške testiranja. Na osnovu grafika, 500 je usvojen kao optimalni broj uzoraka za treniranje neuronske mreže u navedenim uslovima. Isti grafik pokazuje da greška treniranja blago raste sa povećanjem broja podataka za treniranje jer mreža nije u stanju da, sa istim preciznošću, aproksimira veliki broj merenja.

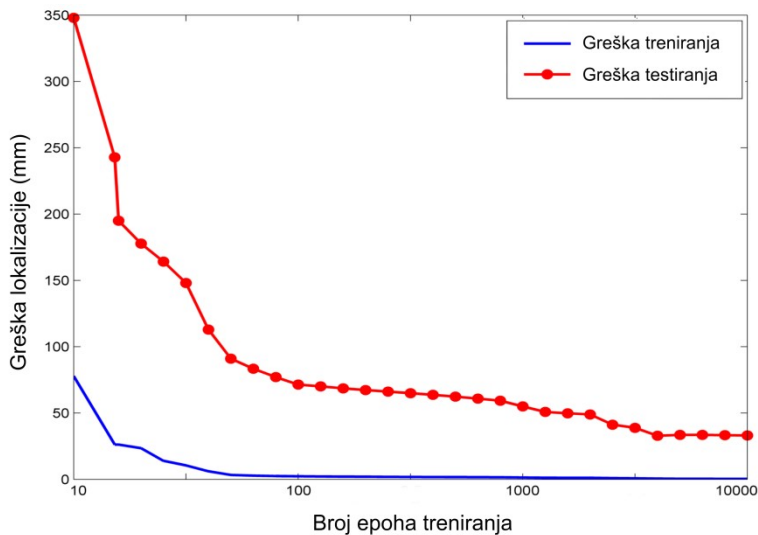


Slika 6.19 Srednje apsolutna greška lokalizacije u zavisnosti od broja uzoraka za treniranje

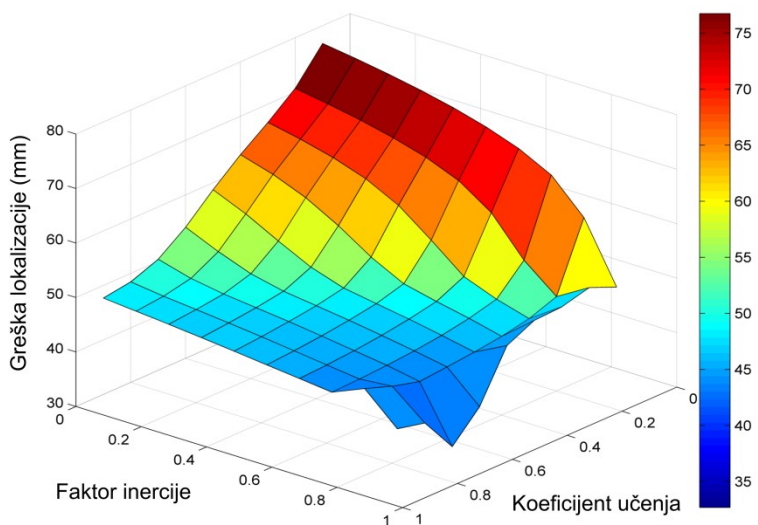
Na kraju su ispitane performanse neuronske mreže u odnosu na parametre algoritma sa povratni prostiranjem greške i momentom. Kada zadata greška nije a priori poznata jedini kriterijum za prekid algoritma je maksimalni dozvoljeni broj epoha treniranja. Druga mogućnost je vremensko ograničavanje treniranja ali nema racionalnog razloga jer ono traje znatno kraće od ostalih poslova koji prate postupak akustične lokalizacije (sklapanje opreme, programiranje, akvizicija podataka...). U odsustvu drugih kriterijuma, dozvoljeni broj epoha za treniranje može biti određen po istom kriterijumu kao i broj neurona, kao onaj broj epoha nakon koga dodavanje novih ne dovodi do značajnog smanjenja greške. U praksi, nakon određenog broja epoha, greška treniranja počinje da osciluje. Kako bi bilo izbegnuto ovakvo ponašanje, maksimalni dozvoljeni broj epoha treba da bude neznatno manji. Zavisnost greške lokalizacije od broja izvedenih epoha u proceduri treniranja, u eksponencijalnoj skali, data je na slici 6.20. Podaci su dobijeni sa 500 uzoraka za treniranje i nepromenjenim ostalim parametrima u odnosu na prethodni test. Grafik pokazuje da obe greške konstantno opadaju sa povećanjem broja epoha sve do približno 4000. Dalji nastavak treniranja dovodio je neminovno, pre ili kasnije, do oscilovanja srednje apsolutne greške.

Izborom adekvatnih parametra algoritma moguće je postići kompromis između trajanja postupka treniranja i njegove tačnosti. Slika 6.21 predstavlja rezultat traganja za optimalnim vrednostima ovih parametara u dvodimenzionalnom prostoru na osnovu srednje apsolutne greške treniranja kao kriterijuma kvaliteta. Rezultati su dobijeni nakon 4000 epoha treniranja

algoritmom sa povratnim prostiranjem greške i inercijom i nepromenjenim ostalim parametrima u odnosu na prethodni slučaj. Na osnovu rezultata testiranja usvojene su vrednosti 0,7 za koeficijent učenja i 0,9 za faktor inercije kao optimalne u uslovima obavljenog eksperimenta.



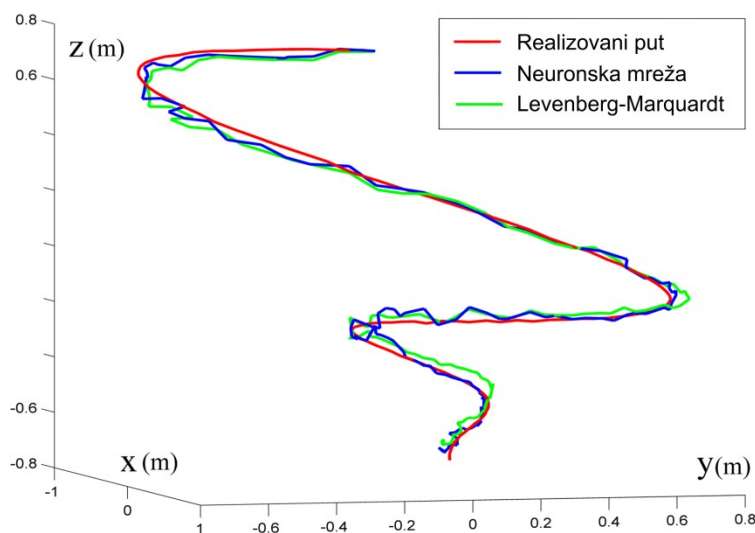
Slika 6.20 Srednje apsolutna greška lokalizacije u zavisnosti od broja epoha treniranja



Slika 6.21 Srednje apsolutna greška lokalizacije u zavisnosti od parametara algoritma sa povratnim prostiranjem greške, koeficijenta učenja i faktora inercije

Konačni rezultat akustične lokalizacije predstavljen je na slici 6.22. Glatka (crvena) kriva predstavlja stvarnu (realizovanu) putanju akustičnog izvora sastavljenu od 126 ravnomerno raspoređenih prostornih pozicija. Plava linija predstavlja rezultat akustične lokalizacije

obezbeđen primenom 8 mikrofona i neuronske mreže sa eksperimentalno izabranim vrednostima konfiguracionih parametara. Rezultujuća srednje apsolutna greška testiranja akustičnog izvora bila je 35.7 mm što je ocenjeno kao zadovoljavajuće u poređenju sa njegovim gabaritima. Rezultat lokalizacije primenom Levenberg-Marquardt-ovog algoritma sa istim skupom podataka prikazan je zelenom linijom. Srednje kvadratna greška bila je 36,3 mm što je približno isti rezultat kao u slučaju neuronske mreže. Međutim, računaska kompleksnost je znatno veća što je uticalo da registrovano vreme trajanja procedure bude približno  $10^4$  puta veće nego u slučaju neuronske mreže.



*Slika 6.22 Stvarna i putanje akustičnog izvora dobijene akustičnom lokalizacijom*

### 6.3.5. Zaključak

Eksperiment je dokazao efikasnost veštačke neuronske mreže sa direktnim prostiranjem signala u rešavanju problema hiperboličke lokalizacije u prostoru. Akustični signali prikupljeni su u prisustvu korelisanog šuma i odjeka primenom mikrofona najnižeg kvaliteta sa tolerancijom u prostornom položaju i položaju akustičnih reflektora. U ovakvim uslovima lokalizacija ne bi bila moguća bez predobrade signala koje je objašnjeno u sledećem odeljku. Osim predobrade na povećanje tačnosti lokalizacije uticao je izbor adekvatne dužine signala i broja redundantnih merenja. Postupak je analiziran na velikom broju akustičnih uzoraka, za različite dužine signala, različite prostorne konfiguracije mikrofona, konfiguracije neuronske mreže i parametre treniranja. Srednje apsolutna greška lokalizacije od 36,3 mm manja je od 42,8 mm što odgovara



srednje apsolutnoj grešci izračunatih kašnjenja ( $0,128 \text{ ms}$ ), pri brzini prostiranja zvuka od  $334 \frac{\text{m}}{\text{s}}$ .  
To znači da je neuronska mreža uspešno kompenzovala brojne neizvesnosti u eksperimentalnim podacima uprkos relativno malom broju redundantnih merenja.

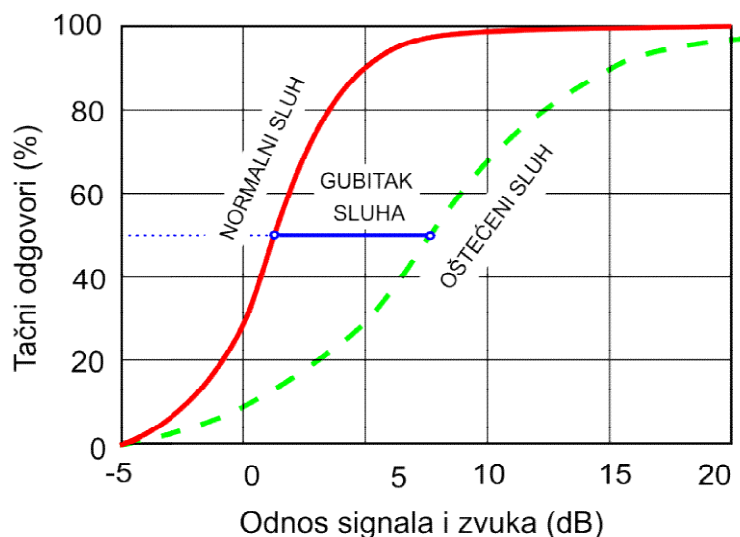
## 7. PREDOBRAĐA I POSTOBRAĐA AKUSTIČNIH SIGNALA

Akustična opservacija funkcioniše besprekorno u laboratorijskim uslovima ali problemi nastaju kada se primeni u praksi. Realne okolnosti podrazumevaju prisustvo poremećaja, različitog porekla, koje ima negativan uticaj na tačnost opservacije. Prisustvo poremećaja u signalu izražava se preko relativnog **odnosa signala i poremećaja** (engl. signal to noise ratio – SNR)

$$SNR (dB) = 20 \log_{10} \left( \frac{A_{sig}}{A_{por}} \right) \quad (7.1)$$

gde su  $A_{sig}$  amplituda signala a  $A_{por}$  amplituda poremećaja (šuma). Slika 7.1 daje krive tačnosti percepcije reči u rečenici, kod osoba normalnog i oštećenog sluha, u zavisnosti od odnosa signala i poremećaja. Razlika između grafika, za 50% tačnih odgovora, predstavlja meru gubitka sluha.

Najčešće se pod šumom podrazumeva nekorelisani (slučajni) poremećaj koja potiče od nesavršenosti medijuma kroz koji se prostire zvuk i instrumenata za njegovu detekciju. Još veći problem može predstavljati korelisani poremećaj koji je, takođe, neizbežan u praksi. On potiče od različitih pojava koje prate svakodnevicu kao što su tehnička sredstva (kompjuterski ventilatori, hard diskovi, motori, prevozna sredstva...), živa bića i ostali prirodni fenomeni.



Slika 7.1 Tačnost akustične percepcije u zavisnosti od odnosa signala i zvuka za osobe normalnog i oštećenog sluha [121]

## 7.1. Metode predobrade

Radi umanjenja uticaja poremećaja, odnosno povećanja sadržaja korisnih informacija, primenjuju se različite metode predobrade akustičnog signala. U pojedinim slučajevima, predobrada podataka može predstavljati najznačajniju i najkompleksniju fazu u čitavom procesu dobijanja konačnog rešenja. Zahvaljujući sposobnosti nelinearnog mapiranja ulazno-izlaznih veličina, neuronske mreže su sposobne da obavljaju određene procese predobrade. Zbog toga je logično pretpostaviti da predobrada ulaznih podataka nema značaja kod primene neuronskih mreža za njihovu obradu. Uprkos tome, obrada signala, pre nego se oni dovedu na ulazne terminale neuronske mreže, po pravilu, daje pozitivan efekat na tačnost izlaznih veličina. Najjednostavnije metode predobrade podrazumevaju linearnu transformaciju ulaznih veličina dok složenije mogu uključivati redukciju dimenzija. Bez obzira na metod, rezultat predobrade su novi signali a procedura njihovog dobijanja naziva se ekstrakcija karakteristika (engl. feature extraction) akustičnih signala.

### 7.1.1. Normalizacija

Normalizacija je jedna od najčešćih metoda predobrade koja se sastoji u linearnom pojačanju ulaznih signala. Ovaj postupak se primenjuje kako bi se nivo ulaza akustičnog signala u doveo u sklad sa značajem koji on treba da ima na vrednost izlaza sistema. Najjednostavnija metoda normalizacije akustičnog signala,  $s_i[k]$ , je dovođenje njegove maksimalne amplitude na jediničnu (ili zadatu) vrednost

$$s_i^n[k] = \frac{s_i[k]}{\max_k(|s_i[k]|)}, \quad k = 1, \dots, K \quad (7.2)$$

Postupak je poznat pod nazivom **normalizacija pika**. Nedostatak je što postojanje samo jednog neželjenog pika može da diskredituje ostale podatke. Dobra alternativa može biti normalizacija po sledećoj formuli

$$s_i^n[k] = \frac{s_i[k] - \bar{s}_i}{\sigma_i}, \quad k = 1, \dots, K \quad (7.3)$$

gde je  $\bar{s}_i$  aritmetička sredina signala

$$\bar{s}_i = \frac{1}{K} \sum_{k=1}^K s_i[k] \quad (7.4)$$

a  $\sigma_i$  je standardna devijacija

$$\sigma_i = \frac{1}{K-1} \sum_{k=1}^K (s_i[k] - \bar{s}_i)^2 \quad (7.5)$$

signala,  $s_i[k]$ , na uzorku dužine  $K$ . Ovaj algoritam normalizacije kompenzuje odstupanje srednje vrednosti amplitude od nule, u literaturi poznato kao DC pomeraj ili DC pojačanje (engl. DC offset). Da bi se akustični signal učinio invarijantnim u odnosu na intenzitet, primenjuje se normalizacija u odnosu na kvadratnu sredinu (engl. Root mean squared – RMS) [122]. Osim pomenutih metoda, koje imaju opštu primenu, razvijene su brojne metode normalizacije akustičnih signala sa specijalnom namenom, na primer za obradu signala namenjenih prepoznavanju govora [123].

### 7.1.2. Filtriranje

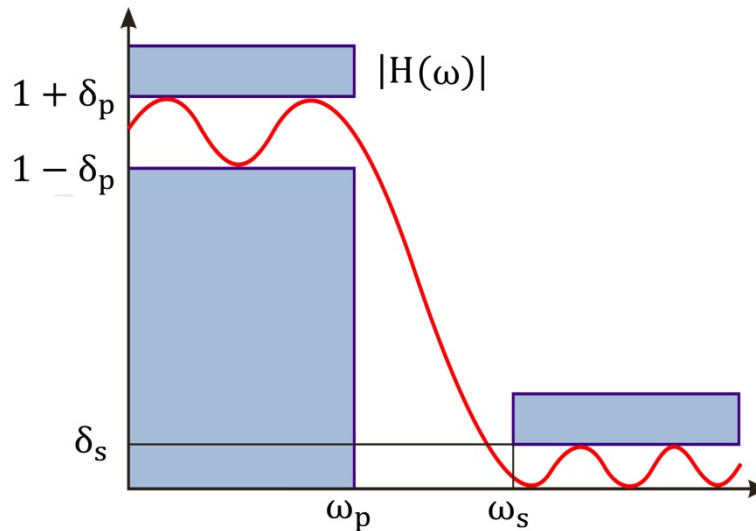
Opšta procedura za poboljšanje kvaliteta akustičnog signala, odnosno odnosa signala i šuma, je filtriranje. U ovu svrhu primenjuje se digitalni filter, sredstvo koje teži umanjeњу nivoa šuma uz očuvanje originalnog signala nepromenjenim. Digitalni filter je linearni sistem čije buduće stanje,  $y[n]$ , zavisi od konačnog broja,  $S$ , prethodnih ulaza i konačnog broja,  $Q$ , prethodnih izlaza

$$y[k] = \sum_{i=0}^P b_i x[k-i] + \sum_{i=1}^Q a_i y[k-i] \quad (7.6)$$

Opšti oblik digitalnog filtera (7.6) naziva se rekurzivni ili filter sa beskonačnim impulsnim odzivom (engl. Infinite Impulse Response – IIR). Ako izlaz filtera ne zavisi od veličine prethodnih izlaza,  $a_i = 0, i = 1, \dots, Q$ , filter je sa konačnim impulsnim odzivom (engl. Finite Impulse Response – FIR). Prenosna funkcija IIR filtera izračunava se z-transformacijom izraza (7.6)

$$H(z) = \frac{\sum_{i=0}^P b_i z^{-i}}{1 - \sum_{i=1}^Q a_i z^{-i}} \quad (7.7)$$

Na osnovu prenosne funkcije dobija se odziv filtera u frekventnom domenu,  $H(\omega)$ . Uobičajena grafička reprezentaciju digitalnog filtera je u vidu amplitudno-frekventnog odziva (Slika 7.2).



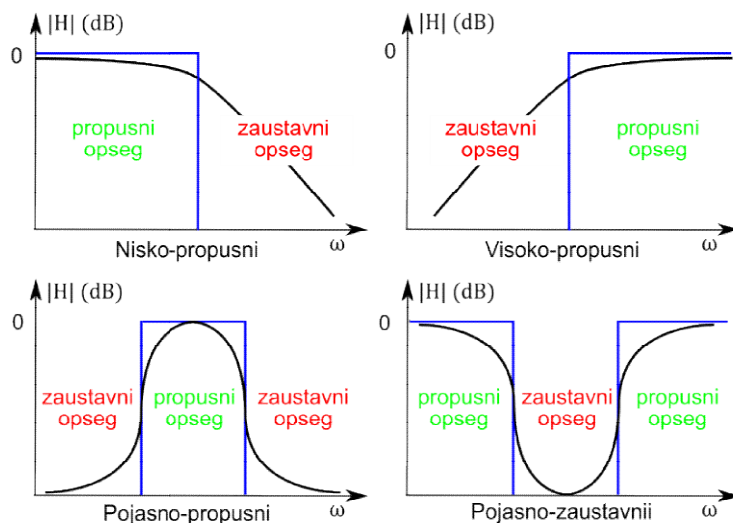
Slika 7.2 Frekventni odziv digitalnog filtera

Projektovanje digitalnog filtera je traganje za adekvatnim koeficijentima prenosne funkcije (7.7) koji obezbeđuju željene karakteristike. Filter je najčešće zadat u vidu amplitudno-frekventnog odziva ili u formi specifikacije karakteristika

- $[0, \omega_p]$  – propusni opseg
- $[\omega_z, \pi]$  – zaustavni opseg
- $[\omega_p, \omega_s]$  – prelazni opseg
- $20 \log_{10}(1 + \delta_p)$  – talasanje propusnog opsega (dB)
- $20 \log_{10}(\delta_z)$  – talasanje zaustavnog opsega (dB)

U nekim slučajevima, osim amplitudnog značajno je poznavanje i faznog odziva,  $\angle H(\omega)$ , digitalnog filtera. Nepovoljna fazna karakteristika može da, na primer, dovede do gubitka informacija o kašnjenju među signalima. Amplitudno-frekventni odzivi tipičnih vrsta digitalnih filtera, u pogledu propusnog opsega, predstavljeni su na slici 7.3. Na svakom od 4 amplitudno-frekventnih odziva, plavom linijom označen je idealna karakteristika, bez prelaznih opsega, a

crnim krivama realna. Što je red filtera veći to se realna karakteristika približava idealnoj, međutim, idealna karakteristika je u praksi neostvariva. Veština projektovanja efikasnog digitalnog filtera ogleda se u nalaženju kompromisa između kompleksnosti filtera, koja se ogleda u njegovom redu, i kvaliteta aproksimacije. Karakteristike idealnog odziva biraju se na osnovu uvida u frekventni spektar signala i poremećaja sa ciljem prigušenja ovih drugih.



Slika 7.3 Amplitudno-frekventne karakteristike tipičnih vrsta filtera

Napredna varijanta klasičnog digitalnog filtera (7.6) je adaptivni filter koji ima sposobnost podešavanja sopstvene prenosne funkcije, čak i za vreme eksploatacije, prema utvrđenom algoritmu optimizacije. Za slučaj uskopojasnog zvuka i širokopojasne buke adaptivni filter ne zahteva bilo kakvu a priori informaciju o ovim signalima. U protivnom, je potrebno poznavanje željenog izlaza. Još jedna prednosti ovih filtera, kao posledica adaptivnosti, je sposobnost reagovanja na promene u nestacionarnim uslovima. Adaptivni filteri se obično realizuju u formi filtera sa konačnim impulsnim odzivom. Ovi filteri su uvek stabilni ali računski zahtevniji i daju lošije performanse od filtera istoga reda sa beskonačnim impulsnim odzivom. Algoritam za podešavanje koeficijenata adaptivnog filtera odvija se u odnosu na neki kriterijum optimizacije. Najčešće je to srednje kvadratna razlika između izlaznog i željenog signala ali pored ovog primenjuju se brojni drugi kriterijumi [124]. Digitalni filteri sa konačnim odzivom imaju unimodalnu površinu srednje kvadratne greške što opravdava primenu gradijentnih metoda za optimizaciju njihovih koeficijenata. Filteri sa neograničenim odzivom imaju više od jednog minimuma zbog čega zahtevaju složenije strategije optimizacije [125].

### 7.1.3. Ekstrakcija karakteristika

Osnovni cilj ekstrakcije karakteristika digitalnog signala je redukcija njegovih dimenzija. Najjednostavniji način da se ona obavi je zanemarivanje pojedinih delova originalnog signala. Drugi način za to mogu biti linearne ili nelinearne kombinacije ulaznih veličina. Redukcija dimenzija smanjuje kompleksnost podataka ali, neminovno, dovodi do gubitka sadržaja korisnih informacija u njemu. Dobro strategija predobrade mora da obezbedi da relevantne informacije u signalu budu sačuvane u što većoj meri. Količina izgubljenih informacija digitalnog signala, u odnosu na original, naziva se **greška kvantifikacije**. Minimalna vrednost ove greške limitirana je rezolucijom digitalnog signala.

Ekstrakcija akustičnih osobina u suštini predstavlja proceduru pretraživanja (Poglavlje 2.2.1) u prostoru različitih karakteristika zvuka. Da bi pretraživanje bilo smisleno potreban je kriterijum za ocenu kvaliteta postignutih rešenja. Pretraživanje se može obavljati na osnovu neke od brojnih strategija sistematskog pretraživanja ili iscrpnim pretraživanjem (engl. exhaustive search) nad skupom mogućih rešenja. Drugi pristup je računski daleko zahtevniji, jer podrazumeva obradu kompletne populacije mogućih rešenja, ali jedini koji garantuje dostizanje globalnog minimuma.

U literaturi su opisane brojne metode, posebne namene, za ekstrakciju osobina akustičnog signala. Procedure izračunavanja perceptualnih karakteristika zvuka, pomenutih u poglavlju 5.1.2.1, takođe spadaju u ekstrakciju karakteristika akustičnog signala. Eksperimentalno prepoznavanje zvuka, opisano u poglavlju 5.2, obavljeno je na osnovu amplitudno-frekventnog spektra. U funkciji ekstrakcije osobina signala, odnosno za izračunavanje spektra, izvršena je diskretna Fourier-ova transformacija ulaznih signala. Pošto je za prepoznavanje upotrebljena samo polovina amplitudnog spektra (5.7), znači da je ekstrakcijom osobina izvršena redukcija dimenzije ulaznog signala na polovinu. Uprkos redukciji dimenzije, perceptivni kvalitet signala je unapređen jer je prepoznavanje zvuka na osnovu frekventne reprezentacije neuporedivo efikasnije u poređenju sa vremenskom. Lokalizacija izvora zvuka, opisana u poglavlju 6.3, obavljena je na osnovu kašnjenja među signalima. Izbor adekvatne metode za ekstrakciju karakteristika, u smislu izračunavanja kašnjenja među signalima (poglavlje 6.2.4.4), izvršen je u okviru postupka nelinearne optimizacije opisanog u poglavlju 7.6.

## 7.2. Postobrada rezultata akustične opservacije

Rezultati akustične opservacije su numerički a u određenim slučajevima mogu biti i kodirani odnosno nerazumljivi za krajnjeg korisnika. Zbog toga je često potrebno izvršiti postobradu izlaznih signala radi tumačenja i prezentovanja dobijenih rezultata. Zbog brojnosti i raznolikosti postupaka postobrade njihova detaljna analiza nije predmet ove disertacije. Dat je samo jedan primer kao ilustracija mogućnosti primene ovih metoda u postupcima akustične opservacije.

### 7.2.1. Kalmanov filter

Reprezentativni primer algoritma za ocenu stanja merenja na osnovu zašumljenih ili nekompletnih merenje je Kalmanov filter. U matematičkom smislu ovaj filter predstavlja rekurzivni postupak kojim se dobija aproksimacija tražene veličine metodom najmanjih kvadrata. Algoritmi koji su se u ovu svrhu upotrebljavali pre Kalmanovog za estimaciju stanja sistema u narednom koraku koristili su informacije iz svih prethodnih koraka. Kalmanov metod filtriranja revolucionaran je u tom smislu što upotrebljava merenje samo iz prethodnog koraka za dobijanje a-priori i a-posteriori predviđanja. Ova predviđanja su međusobno su povezana rezultatima novog merenja.

Uobičajen način modeliranja fizičkog sistema je njegovo predstavljanje sistemom diferencijalnih jednačina. Da bi matematički model tačno opisivao stanje fizičkog sistema neophodno je ustanoviti koje od veličina su stvarno relevantne a koje nemaju značaj. Često se, međutim, dešava da relevantna promenljiva nije raspoloživa za direktno merenje. U tom slučaju se do nje dolazi posredno, izračunavanjem na osnovu merljivih veličina. Dodatni problem predstavlja činjenica da ponašanje realnog sistema nikada nije potpuno determinističko jer na njega deluju različiti poremećaji (šumovi). Merenje fizičkih veličina, u cilju dobijanja veličina stanja, takođe je zašumljeno i izvesnoj meri nepouzđano. Zbog toga ove veličine zapravo predstavljaju slučajne promenljive što navodi na zaključak da je osim determinističkih za predstavljanje realnih sistema neophodno uvesti i stohastičke diferencijalne jednačine. Vremenski diskretan linearan proces, koji nije potpuno deterministički, može biti opisan diferencnom jednačinom stanja



$$\mathbf{X}_k = \mathbf{A}\mathbf{X}_{k-1} + \mathbf{B}\mathbf{U}_{k-1} + \mathbf{W}_{k-1} \quad (7.8)$$

i diferencnom jednačinom mernog sistema

$$\mathbf{Y}_k = \mathbf{H}\mathbf{X}_k + \mathbf{V}_k \quad (7.9)$$

gde je  $\mathbf{X}_k$   $n$ -dimenzionalni vektor stanja koji sadrži sve relevantne veličine u koraku  $k$ . Vektor  $\mathbf{X}_{k-1}$  pomnožen  $n \times n$ -dimenzionalnom matricom sistema  $\mathbf{A}$  čini deterministički deo jednačine stanja.  $\mathbf{U}_k$  je  $m$ -dimenzionalni vektor ulaza u koraku  $k$  koji pomnožen  $n \times m$ -dimenzionalnom matricom ulaza čini nehomogeni deo jednačine stanja.  $\mathbf{W}_k$  je  $n$ -dimenzionalni slučajni vektor poremećaja (šum objekta) u koraku  $k$  koji čini stohastički deo jednačine stanja. Diferencna jednačina mernog sistema pokazuje da je  $l$ -dimenzionalni vektor izlaza (merenih veličina)  $\mathbf{Y}_k$  takođe jednak zbiru determinističkog i stohastičkog dela. Deterministički deo mernog sistema čini proizvod  $l \times n$ -dimenzionalne matrice  $\mathbf{H}$  i vektora stanja a stohastički deo  $l$ -dimenzionalni vektor  $\mathbf{V}_k$  koji predstavlja vektor poremećaja mernog sistema u koraku  $k$ . Zbog jednostavnosti izvođenja pretpostavljeno je da su matrice  $\mathbf{A}$ ,  $\mathbf{B}$  i  $\mathbf{H}$  vremenski invarijantne mada one u praksi mogu biti promenljive.

Cilj Kalmanovog filtriranja je nalaženje vektora stanja  $\mathbf{X}_k$  koji nije direktno merljiv na osnovu merljivog vektora izlaza  $\mathbf{Y}_k$ . Pri izvođenju filtera smatra se da šumovi objekta  $\mathbf{W}_k$  i mernog sistema  $\mathbf{V}_k$  predstavljaju međusobno nezavisne slučajne procese sa Gausovom raspedelom koja ima nulto matematičko očekivanje. Ovi slučajni vektori mogu biti simbolički predstavljeni sledećim izrazima

$$\mathbf{W} \sim N(0, Q) \text{ i } \mathbf{V} \sim N(0, R) \quad (7.10)$$

gde su matrice kovarijanse ova dva vektora definisane na sledeći način

$$\mathbf{Q} = E[\mathbf{W}\mathbf{W}^T] \text{ i } \mathbf{R} = E[\mathbf{V}\mathbf{V}^T] \quad (7.11)$$

Kalmanov filter koristi se za ocenu stanja objekta u određenom vremenskom trenutku. On je napravljen u vidu mehanizma upravljanja sa povratnom spregom i izvršava se u dva koraka. Predikcioni korak algoritma predstavlja vremensko podešavanje ili projekciju ocene stanja iz prethodnog koraka ( $k - 1$ ) na naredni korak ( $k$ ). Rezultat predikcione faze predstavlja a priori ocenu stanja. U fazi korekcije sistem prima informacije o procesu na koga se odnosi a priori

pretpostavka i na osnovu ovih informacija vrši njenu korekciju. Rezultat korekcionog koraka naziva se a posteriori ocena. A-posteriori ocena nakon vremenske projekcije postaje a-priori ocena u sledećem koraku.

Jednačine vremenskog podešavanja

$$\hat{\mathbf{X}}_k^- = \mathbf{A}\hat{\mathbf{X}}_{k-1} + \mathbf{B}\mathbf{U}_{k-1} \quad (7.12)$$

$$\boldsymbol{\Sigma}_k^- = \mathbf{A}\boldsymbol{\Sigma}_{k-1}\mathbf{A}^T + \mathbf{Q} \quad (7.13)$$

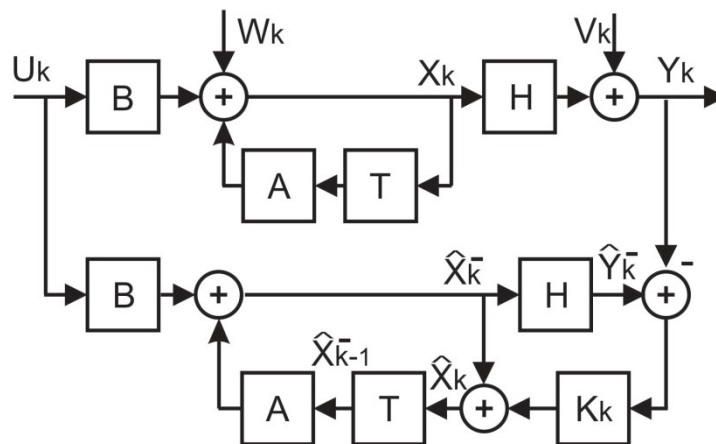
Jednačine podešavanja na osnovu rezultata merenja

$$\mathbf{K}_k = \boldsymbol{\Sigma}_k^- \mathbf{H}^T (\mathbf{H}\boldsymbol{\Sigma}_k^- \mathbf{H}^T + \mathbf{R})^{-1} \quad (7.14)$$

$$\hat{\mathbf{X}}_k = \hat{\mathbf{X}}_k^- + \mathbf{K}_k (\mathbf{Y}_k - \mathbf{H}\hat{\mathbf{X}}_k^-) \quad (7.15)$$

$$\boldsymbol{\Sigma}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{H}) \boldsymbol{\Sigma}_k^- \quad (7.16)$$

gde su  $\hat{\mathbf{X}}_k^-$  a-priori ocena stanja u koraku  $k$  (što podrazumeva poznavanje stanja u prethodnom koraku),  $\hat{\mathbf{X}}_k$  je a-posteriori ocena stanja u koraku  $k$  (što podrazumeva poznavanje rezultata merenja  $\mathbf{Y}_k$ ),  $\mathbf{e}_k^-$  je greška a-priori procene,  $\mathbf{e}_k$  je greška a-posteriori procene,  $\boldsymbol{\Sigma}_k^-$  je matrica kovarijanse a-priori greške i  $\boldsymbol{\Sigma}_k$  matrica kovarijanse a-posteriori greške. Jednačine (7.14-7.16) izvedene su uz pomoć tehnika minimizacije koje se oslanjaju na matricnu algebru. Blok dijagram izvođenja Kalmanovog filtera dat je na slici 7.4.



Slika 7.4 Blok dijagram Kalmanovog filtera

### 7.3. Eksperiment akustičnog prepoznavanja

Drugi deo eksperimenta akustičnog prepoznavanja, opisanog u poglavlju 5.2, bio je posvećen ispitivanju robustnosti postignutog rešenja u odnosu na prisustvo poremećaja i merama namenjenih prevenciji njihovog uticaja.

#### 7.3.1. Metodologija

Premda su originalni uzorci, sakupljeni iz različitih izvora, već sadržali izvesnu količinu šuma, pošlo se od pretpostavke da su oni čisti. U praksi nije ni moguće zabeležiti akustični signal koji je apsolutno lišen šuma. Uticaj slučajnog (belog) šuma i zbunjujućih (konkurentnih) zvukova iz okruženja, na grešku akustičnog prepoznavanja testiran je za različite udele ovih komponenti u mešavini (engl. mix) sa čistim signalom. Udeo svake od komponenti poremećaja ravnomerno je povećavan, u logaritamskoj skali, počev od 0, što odgovara odsustvu komponente poremećaja, pa sve do 1, kada su amplitude signala i odgovarajućeg poremećaja iste. Na ovaj način dobijena je pravouga mreža sa  $22 \times 22$  tačke od kojih svaka odgovara različitoj kombinaciji udela tri akustične komponente u testiranom uzorku. Slučajni šum generisan je primenom funkcije za generisanje slučajnih brojeva dok je signal iz okruženja biran pomoću iste funkcije iz baze akustičnih signala iz okruženja koji su primenjeni u eksperimentu akustičnog prepoznavanja. U svakoj od tačaka izračunata je srednje kvadratna greška testiranja nad skupom podataka koji su opredeljeni za ovu namenu na način opisan u eksperimentu akustičnog prepoznavanja (Poglavlje 5.2). Za testiranje je upotrebljena optimalna konfiguracija veštačke neuronske mreže dobijena kao rezultat ovog eksperimenta.

S obzirom da je prepoznavanje vršeno na osnovu amplitudno-frekventnog spektra akustičnog signala logično je očekivati od ulaznog sloja neuronske mreže da vrši funkciju njegovog filtriranja. Zbog toga je, pod navedenim uslovima, bio upitan efekat digitalnog filtera na unapređenje tačnosti akustičnog prepoznavanja. U eksperimentu je ovaj efekat testiran primenom digitalnog filtera drugog reda sa neograničenim impulsnim odzivom. Na osnovu (7.7) prenosna funkcija ovog filtera ima oblik

$$H(z) = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2}}{1 - a_1 z^{-1} - a_2 z^{-2}} \quad (7.17)$$

Projektovanje filtera (7.17) svodi se na proceduru nelinearne optimizacije. sa ciljem minimizacije greške. Literatura nudi brojne adaptivne metode (Poglavlje 2.3) za izračunavanje prenosne funkcije (7.17) na osnovu raspoloživog skupa ulazno izlaznih podataka (5.7). Zbog svoje jednostavnosti, stohastičke prirode i univerzalnosti, u eksperimentu je za ovu namenu izabran evolucioni račun. Tačnije, primenjena je evolucionna strategija (poglavlje 4.2.2) jer su genotip jedinke činili koeficijenti digitalnog filtera koji imaju realne vrednosti.

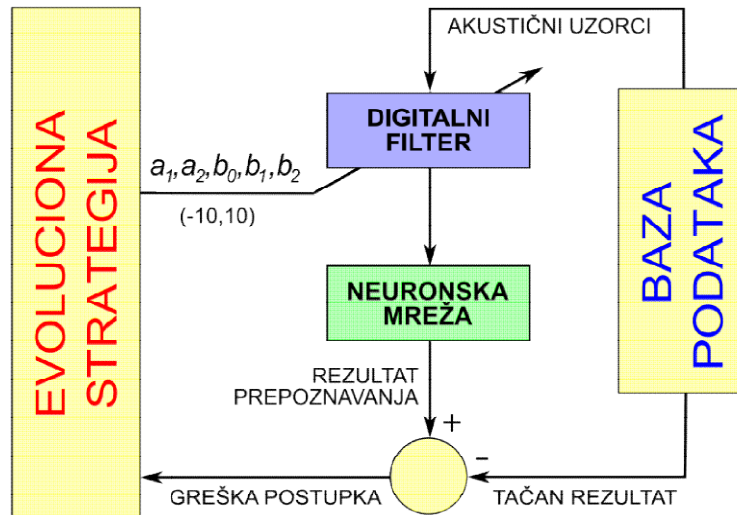
$$X = [a_1 \quad a_2 \quad b_0 \quad b_1 \quad b_2]^T \quad (7.18)$$

Algoritam je startovan sa inicijalnom populacijom od 50 filtera čiji su koeficijenti izabrani na slučajan način u okviru predviđenog opsega,  $(-10,10)$ . Razlika primenjene procedure u odnosu na bilo koji drugi algoritam evolucione strategije sastojala se u načinu evaluacije jedinki. Radi evaluacije efikasnosti, filter je testiran pomoću velikog broja uzoraka smeštenih u bazi podataka (Slika 7.5). Način uzorkovanja signala (Poglavlje 5.2.1), sa stohastičkom komponentom, omogućio je iscrpna testiranja bez ponavljanja sa limitiranim skupom raspoloživih signala. Signali su prethodno zagađivani poremećajima čije je prisustvo, u pogledu sadržaja i intenziteta, takođe dozirano nasumično. Nakon filtriranja signali su dovođeni na ulaz neuronske mreže radi upoređivanja rezultata prepoznavanja sa ciljanim vrednostima smeštenim u bazi podataka. Kao kriterijum kvaliteta usvojena je srednje kvadratna greška prepoznavanja nad čitavim skupom testiranih uzoraka. Nakon evaluacije populacija filtera nastavljala je svoj standardni put u okviru ciklusa evolucione strategije. Kao kriterijum za prekid algoritma usvojen je maksimalan broj uzastopnih evolucionih ciklusa koji ne dovode do smanjenja greške prepoznavanja.

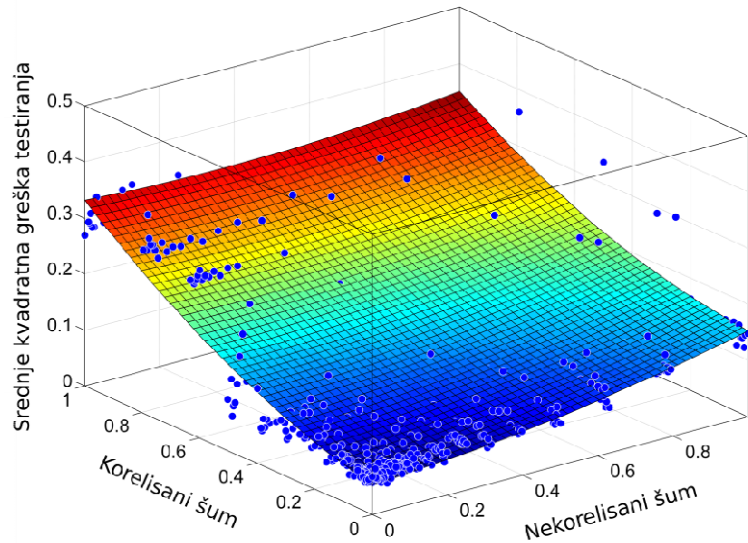
### 7.3.2. Rezultat

Slika 7.6 svedoči o uticaju poremećaja na srednje kvadratnu grešku akustičnog prepoznavanja. Rezultujuće greška, u svakoj od tačaka pravouglo mreže, predstavljena je plavim kružićem a njihova aproksimacija polinomskom površinom drugog stepena u oba pravca. Tački sa koordinatama  $(0,0)$  odgovara srednje kvadratna greška testiranja dobijena prepoznavanjem originalnih signala, bez dodatnih poremećaja, iz eksperimenta akustičnog prepoznavanja. (Poglavlje 5.2). U odsustvu drugih poremećaja, dodavanje belog šuma sa amplitudom jednakom amplitudi signala, na grafiku tačka sa koordinatama  $(1,0)$ , dovelo je do povećanja srednje

kvadratne greške testiranja za približno 3 puta. Prisustvo zvukova iz okruženja sa amplitudom jednakom amplitudi signala povećalo je ovu grešku za 6-7 puta.



Slika 7.5 Blok dijagram algoritma za projektovanje digitalnog filtera

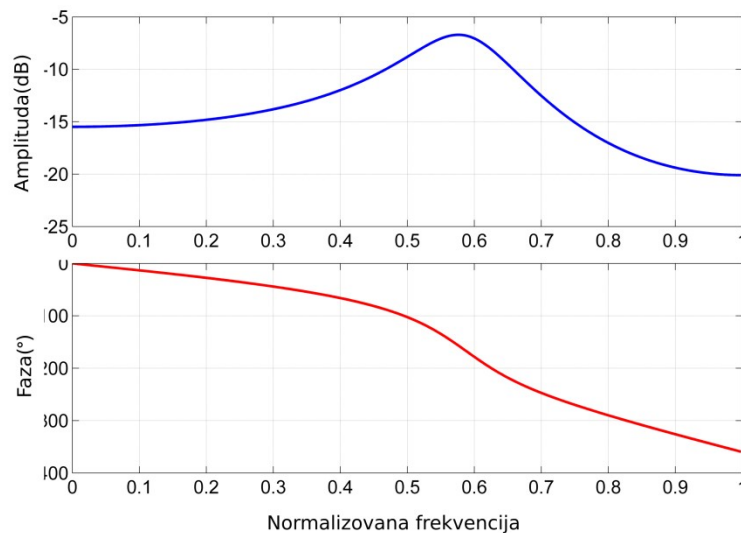


Slika 7.6 Srednje kvadratna greška akustičnog prepoznavanja u zavisnosti od sadržaja poremećaja u signalu

Nakon završetka prethodnog opisanog algoritma, najbolje rezultate, odnosno najveći efekat na povećanje tačnosti akustičnog prepoznavanja, pružio je filter sa sledećom prenosnom funkcijom

$$H(z) = \frac{0,0246 + 0,1105 z^{-1} + 0,2028 z^{-2}}{1 + 0,4136 z^{-1} + 0,5957 z^{-2}} \quad (7.19)$$

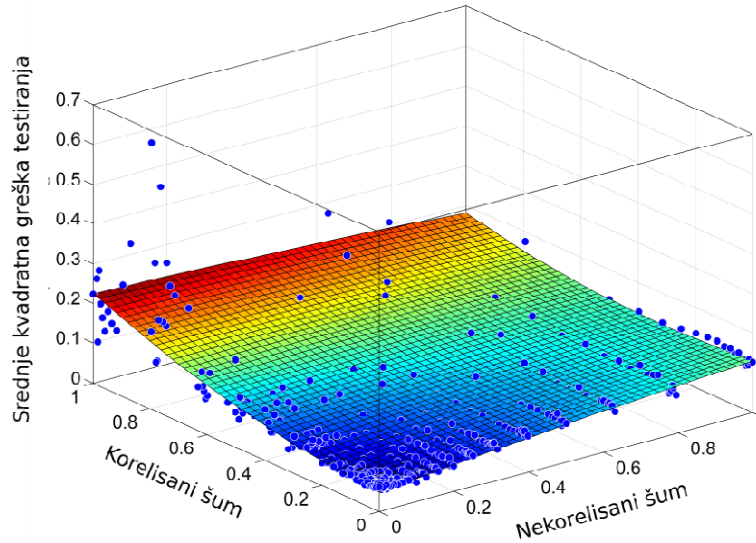
Frekventna karakteristika ovog filtera predstavljena je na slici 7.7 a njegov efekat na srednje kvadratnu grešku testiranja na slici 7.7. Rezultati na poslednjoj slici dobijeni su na potpuno isti način kao i u prethodnom slučaju (slika 7.6) osim što su signali, pre nego su dovedeni na ulaz neuronske mreže, bili podvrgnuti dobijenom digitalnom filteru drugog reda sa neograničenim impulsnim odzivom (7.19). I ovde je aproksimacija rezultata izvršena polinomskom površinom drugog stepena u oba pravca. Na grafiku je primetno opšte smanjenje greške prepoznavanja kao rezultat primene digitalnog filtera.



*Slika 7.7 Frekventna karakteristika digitalnog filtera u eksperimentu akustičnog prepoznavanja*

### 7.3.3. Zaključak

Rezultati eksperimenta sa slike 7.4 svedoče o većem uticaju korelisnog šuma na smanjenje tačnosti akustičnog prepoznavanja u odnosu na nekorelisani šum. Upotrebom digitalnog filtera drugog reda sa neograničenim impulsnim odzivom (7.19) primetno je smanjena greška prepoznavanja kako u zonama uticaja korelisnog tako i u zonama uticaja slučajnog šuma. Pošto je filter projektovan metodom nelinearne optimizacije nad skupom raspoloživih signala logično je pretpostaviti da on neće dati očekivani efekat u nepoznatom okruženju. To znači da je efikasnost predložene metode strogo uslovljena adekvatnim izborom reprezentativnih uzoraka signala i konkurentnih zvukova na osnovu kojih se vrši projektovanje filtera.



*Slika 7.8 Srednje kvadratna greška prepoznavanja prethodnog filtriranog signala u zavisnosti od sadržaja poremećaja*

## 7.4. Eksperiment lokalizacije

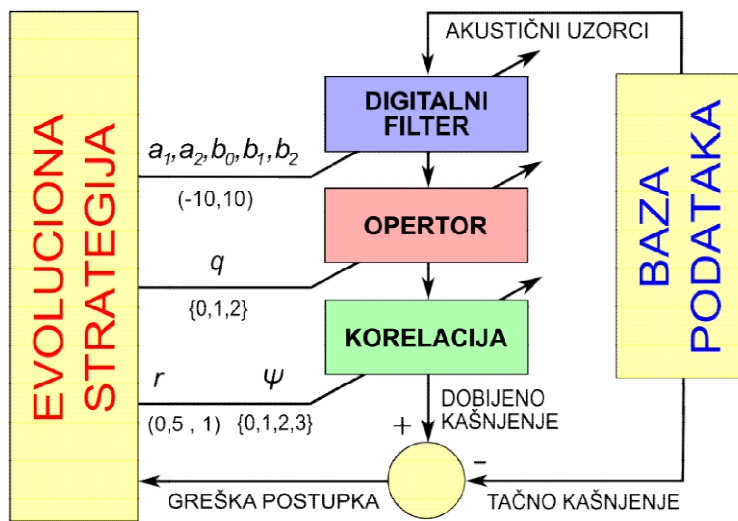
Eksperiment akustične lokalizacije u bliskom polju, predstavljen u poglavlju 6.3, obavljen je u izrazito nepovoljnim uslovima koji se gledaju u snažnom prisustvu šuma i odjeka. Presentovani rezultati dobijeni su tek nakon predobrade akustičnih signala prikupljenih sa niza prostorno raspoređenih mikrofona. Izračunavanja kašnjenja među signalima, kao osnova za predviđeni postupak lokalizacije, neodvojivo je od drugih postupaka predobrade akustičnih signala kao što su normalizacija i filtriranje. Zbog toga je predviđen poseban eksperiment sa ciljem konfigurisanja kompletnog postupka predobrade signala namenjenih akustičnoj lokalizaciji u bliskom polju.

### 7.4.1. Metodologija

Način akvizicije podataka namenjenih akustičnoj lokalizaciji u bliskom polju predstavljen je u poglavlju 6.3.1. Na osnovu realizovanih pozicija izvora zvuka i koordinata mikrofona izračunate su teoretske vrednosti kašnjenja, primenom jednačine (6.14). Poznavanje referentnih kašnjenja omogućilo je eksperimentisanje sa različitim postupcima predobrade akustičnog signala. Nekoliko uobičajenih metoda je eksperimentalno testirano u pogledu njihovog uticaja na tačnost izračunatih kašnjenja. Izračunavanje je pri tome vršeno metodom unakrsne korelacije u

vremenskom domenu (6.18) a kriterijum za evaluaciju bio je prosečno odstupanje od referentnih vrednosti kašnjenja na čitavom skupu raspoloživih uzoraka.

Na kraju je osmišljen i testiran algoritam za konfigurisanje kompletnog postupka predobrade signala (Slika 7.9) koji podrazumeva više postupaka predobrade. Iz istih razloga i na isti način, kao kod eksperimenta akustičnog prepoznavanja, za ovu namenu je primenjena evolucionarna strategija. Osim podešavanja konfiguracionih parametara, u ovom slučaju je upotrebljena i za opredeljivanje između različitih postupka predobrade i različitih postupaka za izračunavanje kašnjenja među signalima. Ovo je zahtevalo značajno proširenje genotipa jedinke u kvantitativnom i kvalitativnom smislu. Pored 5 realnih koeficijenata digitalnog filtera, dodata je realna promenljiva,  $r$ , i dve celobrojne promenljive,  $q$  i  $\Psi$ .



Slika 7.9 Blok dijagram algoritma za projektovanje predobrade akustičnog signala namenjenog lokalizaciji u bliskom polju

Nakon intuitivnog testiranja, primena nekih jednostavnih operacija nad akustičnim signalima dala je pozitivne rezultate na tačnost rezultujućih kašnjenja. Promenljiva  $q$  sa tri celobrojne vrednosti u genotip predobrade uvedena je da opredeli između primene apsolutnog operatora i operatora kvadriranja. Ove opcije jedna drugu isključuju pošto kvadriranje istovremeno znači uklanjanje informacije o znaku signala. Treća vrednost promenljive odgovara opciji kada ni jedan od ova dva operatora nije primenjen.



Pošto različite metode za određivanje kašnjenja među digitalnim signalima odgovaraju različitim uslovima primene (Poglavlje 6.3.2) izbor između njih nije izvršen a priori. U cilju povećanja univerzalnosti postupka, izbor adekvatne metode je odložen da bi bilo napravljen u toku izvođenja algoritma, na osnovu skupa raspoloživih podataka. Za tu svrhu iskorišćena je promenljiva  $\Psi$  sa 4 celobrojne vrednosti koje simbolizuju 4 međusobno isključive opcije. To su unakrsna korelacija u vremenskom i frekventnom domenu, metoda transformacije faze i korelacija sa izravnatim koeficijentom koherencije.

Teoretske vrednosti kašnjenja bile su, po apsolutnoj vrednosti, daleko manje od dužine akustičnih uzoraka. Radi toga je opseg testiranih vrednosti funkcije korelacije (6.17) u potrazi za maksimumom (6.18) ograničen samo na deo od interesa

$$\Delta t_{ij} = \frac{1}{f_s} \arg \max(R_{ij}[l]), \quad -\frac{l_{max}}{2} \leq l \leq \frac{l_{max}}{2} \quad (7.20)$$

Pošto su svi raspoloživi akustični uzorci sakupljeni na pozicijama izvora unutar kubnog prostora sa dužinom ivice od 1,6 m, maksimalno očekivano kašnjenje može se izračunati na osnovu sledeće jednačine

$$\bar{l}_{max} = 1.6\sqrt{3} \cdot \frac{F_s}{c_0} = 365.11 \quad (7.21)$$

gde je  $F_s = 44,1 \text{ kHz}$  frekvencija akustičnog signala a  $c_0 = 334,3 \frac{m}{s}$  njegova brzina prostiranja, u uslovima eksperimenta. Prirodno je da postoji neka rezerva, odnosno da maksimalno kašnjenje primenjeno u eksperimentu,  $l_{max}$ , bude nešto veće od teoretski izračunatog (7.20). U tu svrhu je definisana realna promenljiva  $r$ , kao odnos između ove dve veličine

$$r = \frac{l_{max}}{\bar{l}_{max}} \quad (7.22)$$

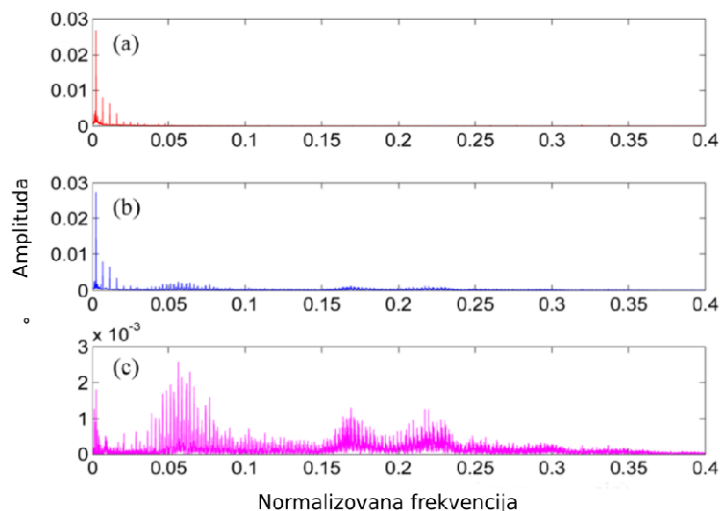
Nakon proširenja, genotip jedinke, koja simbolizuje kompletan postupak predobrade akustičnog signala, bio je sačinjen od 6 realnih i 2 celobrojne promenljive

$$X = [a_1 \ a_2 \ b_0 \ b_1 \ b_2 \ r \ \Psi \ q]^T \quad (7.23)$$

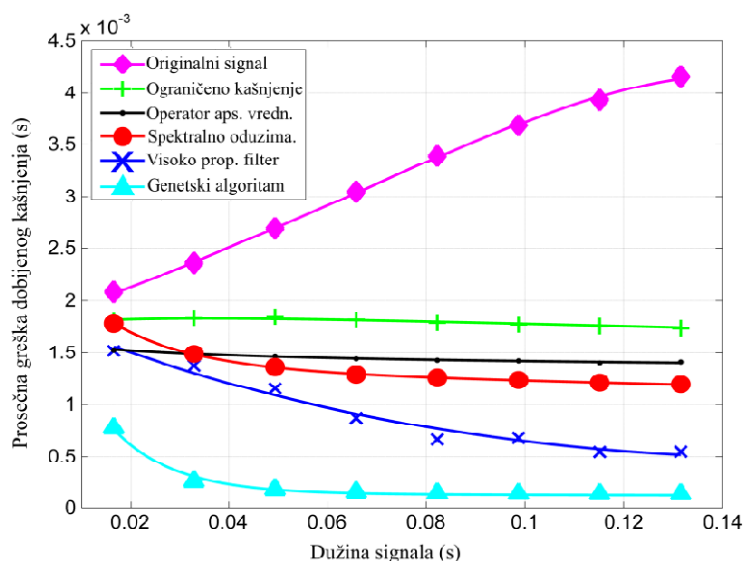
Algoritam je startovan sa inicijalnom populacijom od 50 jedinki, čiji je genotip izabran nasumičnim putem. Vrednosti realnih promenljivih su ograničene, za koeficijente filtera na opseg  $(-10,10)$  a za relativnu vrednost maksimalno očekivanog kašnjenja,  $r$ , na  $(\frac{1}{2}, \frac{3}{2})$ . Kao mera kvaliteta predobrade usvojena je srednja kvadratna greška rezultujućih kašnjenja nad čitavim skupom raspoloživih podataka. Evaluacija jedinki, u okviru algoritma evolucione strategije, vršena je u skladu sa blok dijagramom predstavljenim na slici 7.9. Kao kriterijum za prekid algoritma upotrebljen je maksimalan broj uzastopnih iteracija koje ne dovode do unapređenja tačnosti izračunatih kašnjenja.

#### 7.4.2. Rezultati

Odnos signala i zvuka u eksperimentalnim uzorcima određen je na osnovu frekventnih odziva prikazanih na slici 7.10. Prvi grafik (a) predstavlja amplitudno frekventni odziv poremećaja određen brzom Fourier-ovom transformacijom prve sekunde akustičnog uzorka (Slika 6.15), zabeležene pre nego što je izvor zvuka aktiviran. Drugi grafik (b) reprezentuje frekventni odziv signala. On je rezultat iste procedure nad poslednje dve sekunde uzorka, snimljenih nakon što je izvor aktiviran. Oba grafika dobijena su uprosečavanjem odziva čitave kolekcije raspoloživih uzoraka namenjenih eksperimentu akustične lokalizacije u bliskom polju. Poslednji grafik (c), dobijen oduzimanjem prethodna dva, za potrebe određivanja odnosa signala i buke je upotrebljen kao frekventni odziv čistog signala, u potpunosti lišenog buke. Na osnovu prvog i poslednjeg grafika, odnos maksimalne amplitude čistog signala i maksimalne amplitude buke je bio približno 0,1 što odgovara logaritamskoj vrednosti od -20 dB (7.1). Literatura predviđa donju granicu odnosa signala i buke, koja omogućava precizno izračunavanje kašnjenja primenom korelacionih metoda, u intervalu između -13 dB i -13,5 dB [126]. Prema ovome, sakupljeni akustični signali sadržali su previše poremećaja da bi bili upotrebljivi bez prethodne obrade. To potvrđuje i zavisnost srednje apsolutne greške rezultujućih kašnjenja od dužine signala, na slici 7.11 obeležena purpurnom linijom sa deltoidima, dobijena na osnovu neobrađenih signala. Marker, na svim krivama ovog grafika, označavaju eksperimentalno utvrđenu srednje kvadratnu grešku izračunatih kašnjenja nad čitavom kolekcijom podataka, a linije njihovu aproksimaciju dobijenu primenom eksponencijalne funkcije sa dva člana. Grafik pokazuje povećanje greške izračunatih kašnjenja sa povećanjem dužine signala što je matematički besmisleno.



Slika 7.10 Amplitudno frekventni odziv: a) poremećaj, b) signal, c) razlika

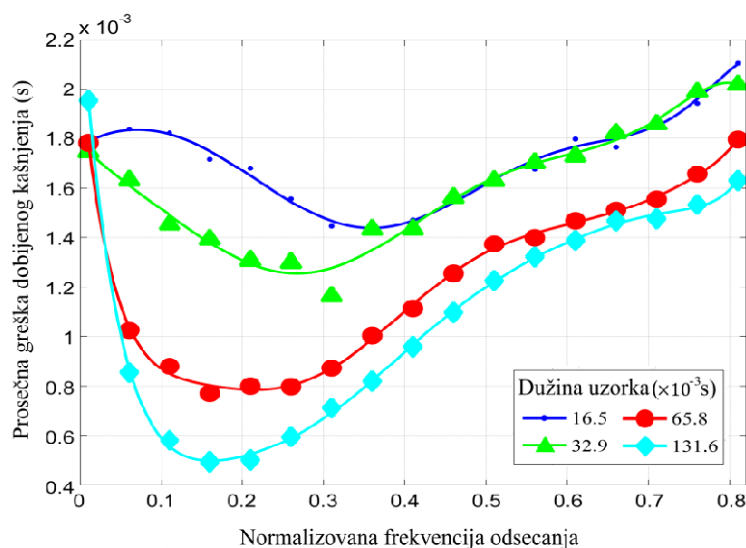


Slika 7.11 Srednja greška izračunatih kašnjenja u zavisnosti od dužine signala, za različite metode predobrade akustičnog signala

Pošto je najveći deo energije poremećaja u uzorcima ležao u niskom opsegu frekvencija (Slika 7.10 a), ispod  $0.05 \cdot \pi$ , razumno je bilo pretpostaviti će dobar efekat na povećanje tačnosti dobijenih kašnjenja dati visokopropusni filter. Zbog svoje jednostavnosti, za ovu namenu izabran je Butterworth-ov filter (Stephen Butterworth) drugog reda koga karakteriše pojačanje

$$G(\omega) = |H(j\omega)| = \frac{1}{\sqrt{1 + \left(\frac{\omega}{\omega_c}\right)^2}} \quad (7.24)$$

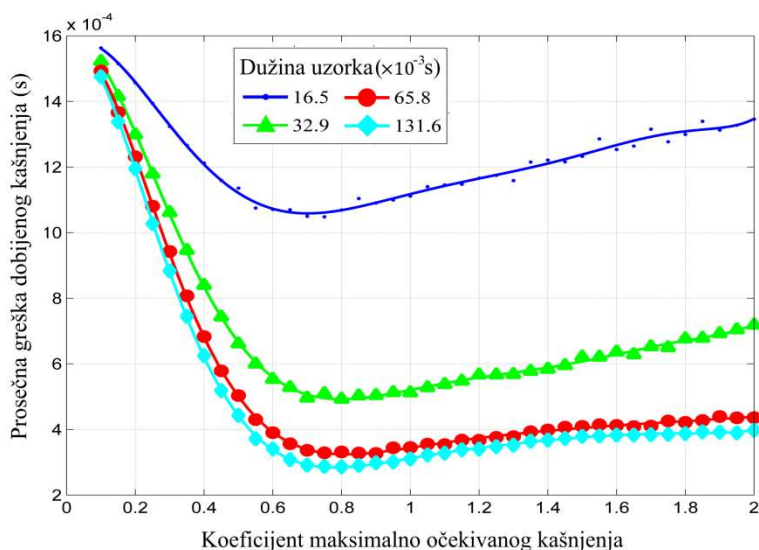
gde je  $\omega_c$  frekvencija odsecanja. Slika 7.12 predstavlja zavisnost srednje apsolutne greške dobijenih kašnjenja od frekvencije odsecanja filtera Butterworth-a drugog reda. Svaka od 4 krive odgovara različitoj dužini testiranih uzoraka,  $T$ . Markeri odgovaraju eksperimentalnim vrednostima a linije su dobijene polinomskom aproksimacijom osmog stepena. Na svim graficima je očigledno postojanje frekvencije odsecanja koja minimizuje grešku rezultata. Ova frekvencije je različita za različite dužine uzoraka. Generalno, što je uzorak duži to je manja frekvencija odsecanja koji obezbeđuje najbolje prigušenje poremećaja. Plava kriva sa x markerima na slici 7.11 ilustruje efekat primene visoko-propusnog filtera na tačnost dobijenih rezultata. U zavisnosti od dužine uzorka korišćen je filter sa adekvatnom frekvencijom odsecanja koja minimizuje vrednost greške. Kriva svedoči da je filtriranje smanjilo nivo buke u signalu i time ga učinilo upotrebljivim za akustičnu lokalizaciju.



Slika 7.12 Srednja greška izračunatih kašnjenja u zavisnosti od frekvencije presecanja visoko-propusnog filtera

Uticaj maksimalno dopuštenog kašnjenja, u formi relativnog odnosa sa teoretskim opsegom kašnjenja (7.22), na veličinu greške je takođe detaljno ispitan a rezultati, za 4 različite dužine signala, su prezentovani na slici 7.13. Uzorci su prethodnog podvrgnuti adekvatnom filteru Butterworth-a drugog reda koji je obezbeđen na način opisan u prethodnom paragrafu. Markerima su označene eksperimentalne vrednosti a linijama njihova aproksimacija polinomom osmog stepena. Grafik pokazuje ekstremno veliku grešku rezultata za signale sa malim

relativnim odnosom maksimalnih kašnjenja,  $r$ . Ovi uzorci su previše kratkotrajni tako da je većina kašnjenja ispadala iz opsega ispitivanja (7.20). Nakon dostizanja minimuma, u vrednosti između 0,75 i 0,8 u zavisnosti od dužine uzorka, greška je ponovo rasla. Razlog leži u činjenici da kod previše dugih signala, koji se jednoliko ponavljaju, algoritam unakrsne korelacije može da preskoči susedni dominantni pik i da registruje neki od narednih. Na ovaj način izračunato kašnjenje se lažno uvećava za višestruku vrednost dominantne talasne dužine signala. Minimumi greške, na prikazanim krivama, predstavljaju tačke balansa između prethodno opisanih negativnih efekta. Kriva na slici 7.11 obeležena zelenom linijom sa krstićima prikazuje efekat ograničavanja opsega pretraživanja u algoritmu unakrsne korelacije na neobrađenim (sirovim) signalima. Očigledno je da ova mera umanjuje grešku ali ipak ne može da obezbedi smislene rezultate u odsustvu drugih metoda predobrade, pre svega filtriranja.



*Slika 7.13 Srednja greška izračunatih kašnjenja u zavisnosti od relativne vrednosti maksimalno dopuštenog kašnjenja*

Sledeće dve krive, na slici 7.11, predstavljaju efekat apsolutnog operatora, označen crnom linijom sa tačkama, i metode spektralnog oduzimanja, označene crvenom linijom sa kružićima, na grešku rezultujućih kašnjenja. Apsolutni operator ima statističkog efekta na smanjenje greške ali ne može da obezbedi smisljeno izračunavanje kašnjenja u odsustvu drugih postupaka predobrade. Interesantno je da je kvadriranje signala, odnosno izračunavanje kašnjenja na osnovu njegove snage, dalo najgore rezultate. Greška nije predstavljena jer je višestruko prevazilazila opseg grafika. Metod spektralnog oduzimanja je regularni metod predobrade

diskretnih signala koji podrazumeva oduzimanje frekventnog spektra šuma od frekventnog spektra zagađenog signala. Vremenski ekvivalent rezultujućeg spektra koristi se kao aproksimacija signala lišenog šuma. Pošto je svaki od raspoloživih uzoraka sadržao deo sa šumom i deo sa zagađenim signalom ovaj postupak je primenjen direktno na osnovu njih uz pomoć brze Fourier-ove transformacije (5.3). Dobijena kriva pokazuje da je spektralnim oduzimanjem omogućeno smisljeno izračunavanje kašnjenja ali je efekat bio manji nego kod visoko-propusnog filtera.

Na kraju je testiran postupak u kome se izbor i konfiguracije kompletne predobrade signala vrši primenom evolucione strategije na osnovu raspoloživih uzoraka (Slika 7.9). Rezultat je prikazan, na slici 7.11, tirkizno plavom linijom sa trouglovima. Za svaku od 8 ispitanih trajanja uzoraka, koji su označeni markerima, algoritam je izvođen ponovo radi dobijanja adekvatne konfiguracije predobrade. Primera radi, primena algoritma na osnovu uzoraka u trajanju od 131 ms dala je sledeći genotip predobrade signala

$$X = [0,1516 \quad -0,2411 \quad 4,7328 \quad 0,7597 \quad -5,4728 \quad 1,1313 \quad 1 \quad 1]^T \quad (7.25)$$

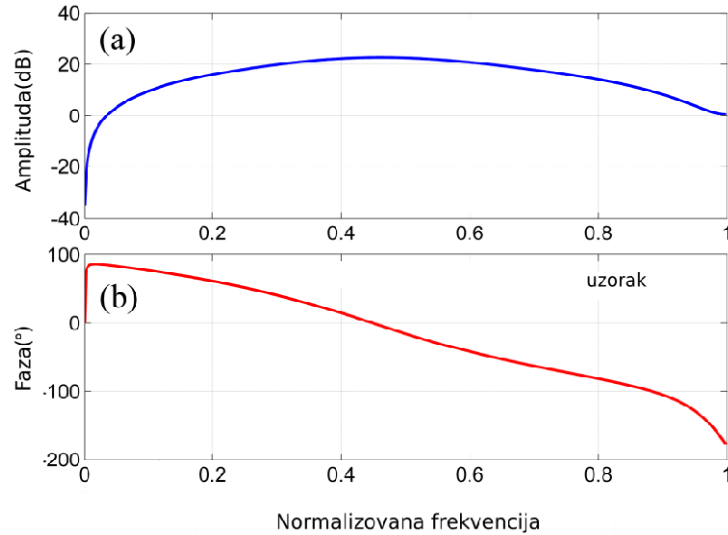
Prvih pet realnih vrednosti određuju koeficijente prenosne funkcije filtera koji je imao optimalni efekat na smanjenje greške dobijenih kašnjenja

$$H(z) = \frac{4,7328 + 0,759 z^{-1} - 5,4728 z^{-2}}{1 - 0,1516 z^{-1} + 0,2411 z^{-2}} \quad (7.26)$$

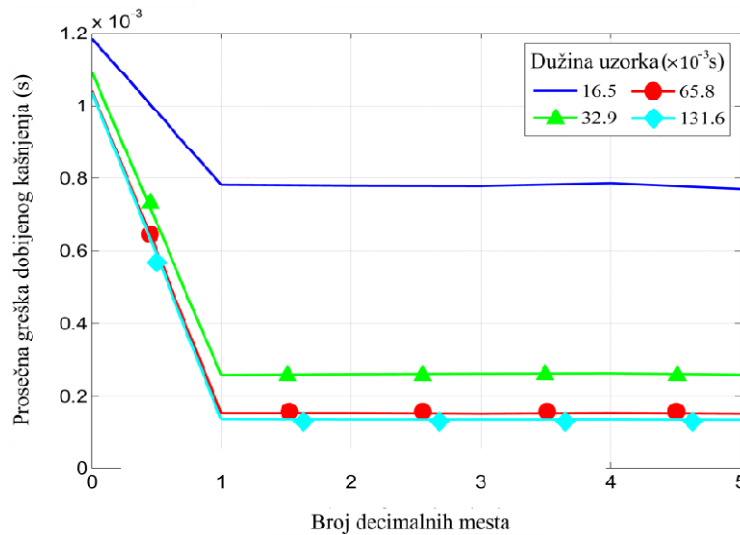
Frekventni odziv ovog filtera predstavljen je na slici 7.14. Interesantno je da je stohastički postupak dao približno linearnu karakteristiku faze koja je neophodna radi očuvanja informacije o kašnjenju. Vrednost realne promenljive  $r = 1,1313$  u kombinaciji sa jednačinama (7.21) i (7.22) dao je optimalnu širinu opsega testiranih kašnjenja

$$l_{\max} = r \cdot \bar{l}_{\max} \approx 412 \quad (7.27)$$

Poslednja dva člana u genu, sa vrednostima jednakim jedinici, govore da je najpreciznije izračunavanje kašnjenja obavljeno metodom unakrsne korelacije u vremenskom domenu ali tek nakon što je iz signala uklonjena informacija o znaku.



Slika 7.14 Frekventni odziv sa optimalnim efektom na smanjenje greške izračunatih kašnjenja kod uzoraka koji traju 131 ms



Slika 7.15 Srednja greška izračunatih kašnjenja u zavisnosti od broja decimalnih mesta u parametrima predobrade

Na kraju je izvršena analiza konačne dužine reči (engl. finite word length analysis) radi definisanja potrebne tačnosti realnih promenljivih u genu a samim tim i uslova za prekid algoritma [127]. Ova analiza, čiji su rezultati predstavljeni na slici 7.15, pokazuje kako zaokruživanje realnih koeficijenata u genu utiče na veličinu greške dobijenih kašnjenja. Samo prvi rezultat, sa celobrojnim vrednostima koeficijenata, dao je značajnu grešku. Posle dodavanja jedne decimale greška se ustalila i nije padala sa daljim povećavanjem preciznosti koeficijenata.

### 7.4.3. Zaključak

Rezultati eksperimenta je jednostavna heuristička metoda za izbor optimalne između različitih ponuđenih strategija za predobradu signala namenjenih lokalizaciji izvora zvuka u bliskom polju. Efektivnost metode ispitana je na velikom broju primeraka sakupljenim, uz pomoć akustične opreme niskog kvaliteta, u izrazito nepovoljnim uslovima koji se ogledaju u prisustvu korelisanog, nekorelisanog šuma i odjeka. Slika 7.11 svedoči o tome da je pravilan izbor konfiguracionih parametara obezbedio drastično smanjenje greške izračunatih kašnjenja među signalima i time omogućio lokalizaciju izvora u bliskom polju. Osim projektovanja koeficijenata filtera, algoritma je omogućio jednostavno rešavanje nekih od većitih dilema vezanih za lokalizaciju izvora zvuka kao što su izbor metode za izračunavanje kašnjenja. Primena evolucionog računa obezbedila je jednostavnost, univerzalnost i adaptivnost postupka tako da se on može vrlo lako prilagoditi u pogledu broja parametara, opcija i cilja. Nedostatak je što nema garancije uspeha, jer je efekat uslovljen adekvatnim izborom reprezentativnih uzoraka.



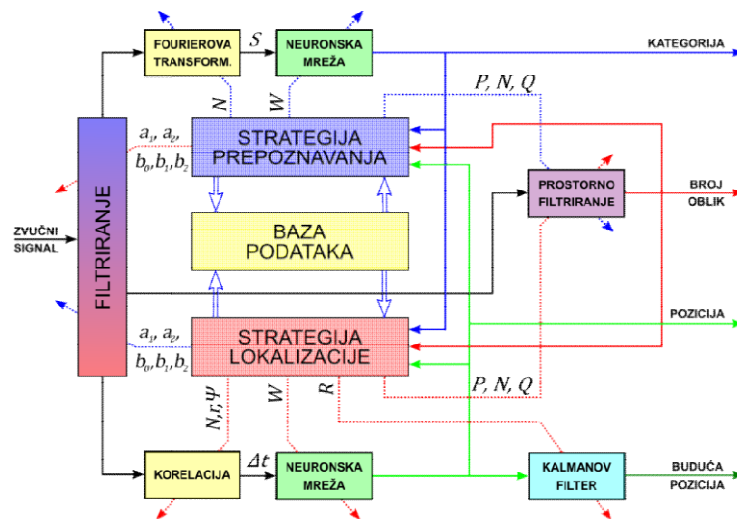
## 8. HIBRIDNI SISTEM ZA PERCEPCIJU ZVUKA

U skladu sa ciljem i rezultatima preduzetih, teoretskih i eksperimentalnih, istraživanja predložen je hibridni sistem namenjen akustičnoj opservaciji. Sistem, koji funkcioniše prema blok dijagramu predstavljenom na slici 8.1, ostvaren je kombinovanjem različitih metoda veštačke inteligencije.

### 8.1. Realizacija

Da bi se stvorio odnos prema nekom objektu pre svega ga je neophodno identifikovati. Zbog toga percepcija počinje prepoznavanjem objekata od interesa odnosno njegovim svrstavanjem u određenu kategoriju objekata. Prepoznavanje je predviđeno u okviru iterativnog postupka kod koga se svaka iteracija obavlja po algoritmu objašnjenom u poglavlju 5.2 ali svaka sa drugim konfiguracionim parametrima. Prema tome, svaka iteracija uključuje faze filtriranja signala, izračunavanja amplitudno-frekventnog spektra i, na osnovu njega, određivanja kategorije objekta primenom neuronske mreže. Odgovarajući konfiguracioni parametri postupka su koeficijenti digitalnog filtera  $a_1, a_2, \dots, b_0, b_1, b_2, \dots$  dužina akustičnog uzorka,  $N$ , na osnovu koga se vrši izračunavanje frekventnog spektra i skup pojačanja neuronske mreže,  $W$ , koja je namenjena prepoznavanju. Parametri prepoznavanja određuju se na način opisan u poglavlju 5.2 a koeficijenti digitalnog filtera postupkom objašnjenim u poglavlju 7.2. Skup pojačanja određuje i konfiguraciju neuronske mreže, u uslovima kada je izbor ograničen na mreže sa direktnim prostiranjem signala. Takođe određuje i dužinu akustičnog uzorka jer ona mora biti dvostruko veća od broja ulaznih terminala mreže. Konfiguracija postupka prepoznavanja, u svakom koraku, usklađuju se sa strategijom prepoznavanja, koja je ostvarena primenom logičkog programiranja. Prepoznavanje počinje sa konfiguracijom namenjenom registrovanju najšireg spektra očekivanih zvukova, u najopštijem slučaju zvuka uopšte. Usmeravanje toka pretraživanja prostora akustičnih objekata (slika 5.6), u svakoj iteraciji, vrši se na osnovu rezultata dobijenih u prethodnoj iteraciji. Pri tome je logično da uspešno prepoznavanje vodi ka nižim nivoima apstrakcije a neuspešno ka višim. Međutim i neuspešno prepoznavanje može doprineti sužavanju prostora pretraživanja, tako što iz njega privremeno eliminiše određenu kategoriju objekata. Kada postupak dostigne najniži nivo apstrakcije on nastavlja iterativni ciklus sa poslednjom konfiguracijom sve dok registruje prisustvo akustičnog objekta.

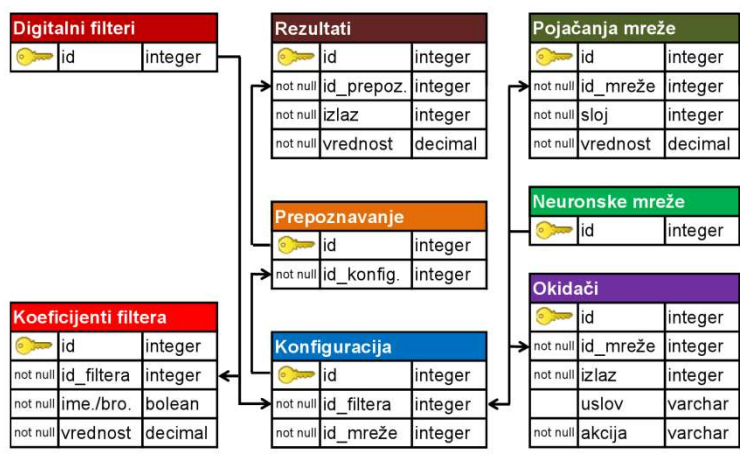
Očigledno je da preciznost predviđenog modela prepoznavanja zavisi od broja, reprezentativnosti i uzročno posledične povezanosti raspoloživih konfiguracija za prepoznavanje akustičnih objekata. Radi smeštanja podataka o različitim konfiguracijama neizbežna je baza podataka koja može biti smeštena u lokalnoj memoriji ili na internetu. Prisustvo baze podataka upotrebljeno je, takođe, za smeštanje pravila rezonovanja, prilikom prepoznavanja akustičnog izvora. Standardne baze podataka omogućavaju jednostavnu realizaciju pravila u formi okidača i poseduju ugrađen interpretator za njihovo sprovođenje. Adekvatno implementaciji, pravila pretraživanja organizuju se u skladu sa jedom od dve moguće strategije pretraživanja (poglavlje 2.2.3). S obzirom da se radi o percepciji, u konkretnom slučaju predloženo je pretraživanje vođeno činjenicama.



Slika 8.1 Blok dijagram hibridnog sistema za percepciju akustičnih signala

Slika 8. predstavlja primer minimalne realizacije baze podataka koja omogućava izvođenje prethodnog opisanog postupka. Tabela “Prepoznavanje” upotrebljena je za registrovanje preduzetih ciklusa prepoznavanja i upotrebljenih konfiguracija postupka prepoznavanja. Tabela “Konfiguracija” sadrži informacije o upotrebljenom filteru i neuronskoj mreži. Podaci o digitalnim filterima, koeficijentima digitalnih filtera, neuronskim mrežama i pojačanjima neuronskih mreža smešteni su u istoimenim tabelama u skladu sa pravilima organizovanja informacija u relacionoj bazi podataka. U tabeli “Rezultati” može se smeštati samo redni broj dominantnog izlaza, kao rezultata ciklusa prepoznavanja, ili alternativno vrednost svih izlaza neuronske mreže. Veća količina informacija o rezultatima omogućava ostvarenje kompleksnijih

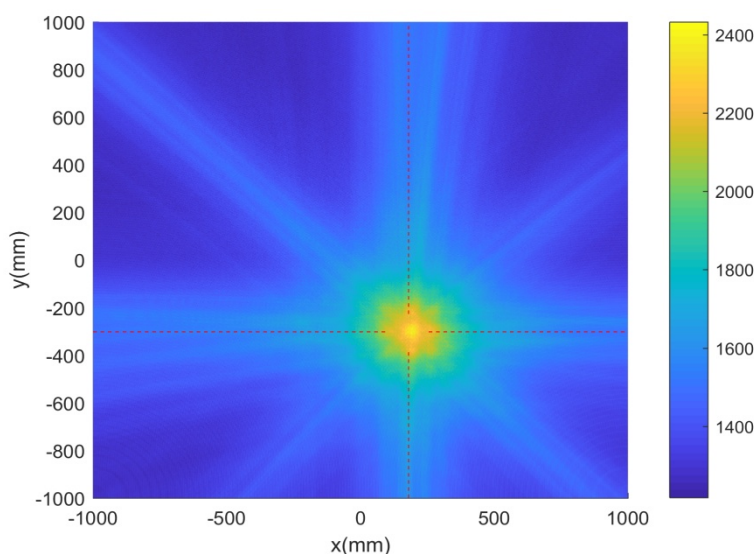
uslova za iniciranje okidača, samim tim kompleksnije strategije prepoznavanja. Tabela pod nazivom “Okidači”, sadrži informacije o radnjama koje treba preduzeti u slučaju ostvarenja određenih rezultata prepoznavanja kao i preduslove za iniciranje ovih radnji. Već je napomenuto da baze podataka sadrže ugrađene okidače odnosno potprograme koji se iniciraju kao posledica manipulacije podacima, odnosno njihovog dodavanja (engl. insert), podešavanja (engl. update) i brisanja (engl. delete). U konkretnom slučaju, tok prepoznavanja ostvaruje se, neposredno, primenom okidača koji se inicira nakon dodavanja (engl. After Insert) podataka u tabelu rezultati. Naredna akcija je najčešće ponavljanje procedure prepoznavanja primenom iste konfiguracije radi potvrde rezultata prepoznavanja odnosno potvrde prisustva registrovanog objekta. Alternativno, rezultat prepoznavanja se može proveriti primenom druge konfiguracije. Osim rezultatima prepoznavanja, izbor naredne akcije može biti uslovljen rezultatima drugih metoda akustične percepcije. Na primer, rezultat akustičnog prepoznavanja može biti diskvalifikovan ako njegov oblik odstupa od rezultata prostornog filtriranja ili ako njegova brzina i obrazac kretanja odstupaju od rezultata dobijenih akustičnom lokalizacijom.



Slika 8.2 Baza podataka za realizaciju strategije prepoznavanja

Prostorno filtriranje se može izvoditi paralelno sa procesom prepoznavanja ili, ako za tim nema potrebe, njegovo iniciranje može biti definisano prema potrebi pomoću okidača. Način izvođenja ovog postupka objašnjen je u poglavlju 6.2.3. Prostorno filtriranje se vrši radi potvrde rezultata akustične percepcije pa se u skladu sa tim biraju i njegovi parametri. Na primer, radi uštede u računskoj kompleksnosti, prostorno filtriranje može biti vršeno samo u okolini akustičnog objekta. Pri tome je pozicija ove okoline dobijena metodom akustične lokalizacije a

njena veličina, odnosno opseg filtriranja, u skladu sa dimenzijama akustičnog objekta koje se očekuju na osnovu rezultata akustičnog prepoznavanja. Slika 8.3. predstavlja rezultat prostornog filtriranja signala iz eksperimenta akustične lokalizacije (Poglavlje 6.3). Rezultati se odnose na horizontalnu ravan koja prolazi kroz maksimum a isprekidane linije predstavljaju stvarnu poziciju izvora zvuka.



*Slika 8.3 Slika akustičnog izvora dobijena metodom prostornog filtriranja*

Konfiguracija neuronske mreže namenjene akustičnoj lokalizaciji, uključujući i njene težinske koeficijente, ne zavisi od prirode akustičnog objekta. Zbog toga ona ostaje ista dok god su geometrijske relacije među mikrofona konstantne. Međutim, strategija lokalizacije treba da obezbedi izbor digitalnog filtera u skladu sa rezultatima prepoznavanja akustičnog izvora. Na ovaj način obezbeđuje se uspešno izolovanje signala koji potiče od ciljanog objekta iz izvorne mešavine signala. Izbor metode za određivanje kašnjenja među signalima kao i dužina akustičnih uzorka za ovu namenu, takođe, mogu biti usklađivani sa rezultatima drugih postupaka akustične opservacije

Na kraju je predviđena i postobrada rezultata akustične lokalizacije primenom Kalmanovog filtera. Cilj navedenog postupka, koji je objašnjen u poglavlju 7.2.1, je prognoza buduće pozicije akustičnog objekta. Buduća pozicija je od značaja kada je potrebno izvršiti neku akciju na osnovu pozicije akustičnog objekta u pokretu. To može biti, na primer, usmeravanje paraboličnih reflektora ka njemu sa ciljem mehaničkog pojačanja registrovanog signala.

## **8.2. Zaključak**

Kombinovanjem različitih metoda veštačke inteligencije ostvaren je hibridni sistem za akustičnu opservaciju koji pruža podatke o kategoriji, lokaciji, obliku i broju akustičnih objekata i daje prognozu o njegovoj budućoj poziciji. Primenom reprezentacije znanja uz pomoć pravila omogućeno je efikasno kontrolisanje postupka akustične percepcije koja uključuje interakciju između različitih metoda. Smeštanjem različitih strategija akustične percepcije i odgovarajućih konfiguracionih parametara u bazu podataka ostvarena je univerzalnost postupka. Ovakva baza omogućava jednostavno unapređenje postupka, na osnovu novih saznanja, i efikasnu razmenu iskustava putem digitalnih medija.

## 9. ZAKLJUČAK I PRAVCI DALJIH ISTRAŽIVANJA

U okviru opisanog naučnog istraživanja ostvareni su sledeći rezultati:

- Sistem za akviziciju i obradu signala namenjenih akustičnom prepoznavanju
- Sistem za akviziciju i obradu akustičnih uzoraka namenjenih akustičnoj lokalizaciji u bliskom polju
- Algoritam za prepoznavanje (kategorizaciju) akustičnih signala na osnovu amplitudno frekventnog spektra
- Sistema za lokalizaciju izvora zvuka na osnovu kašnjenja među signalima zabeleženim na različitim prostornim pozicijama
- Algoritam za konfigurisanje postupka predobrade akustičnih signala sa ciljem unapređenja tačnosti akustične opservacije.
- Hibridni sistem za percepciju akustičnih objekata

Primenom algoritama računске inteligencije, nastalih po ugledu na biološke sisteme, postignuta je univerzalnost i adaptivnost navedenih metoda. Kod obrade signala akcenat je stavljen na temporalnoj rezoluciji (odnosu preciznosti rezultata i trajanja obrade signala). Iz tog razloga je za ovu namenu, u oba postupka akustične opservacije (prepoznavanje i lokalizacija), predviđena veštačka neuronska mreža sa direktnim prostiranjem signala. Metode za treniranje neuronskih mreža zahtevaju daleko veći kapacitet izračunavanja ali se treniranje obavlja jednokratno tako da nema uticaja na fazu eksploatacije. Konfiguracija postupaka predobrade, u cilju postizanja maksimalne univerzalnosti, obavljena je primenom evolucionog računa. I ovaj postupak se obavlja jednokratno, u fazi dizajniranja sistema, tako da ne utiče na produženje vremena obrade signala u eksploataciji.

Rezultati naučnog istraživanja predstavljaju proširenje postojećih saznanja, od teoretskog i praktičnog značaja, za dalji razvoj akustične opservacije. U uslovima sve veće potrebe za interakcijom veštačkih sistema sa okolinom navedeni postupci predstavljaju alternativu ostalim tehnikama opservacije jer ne zahtevaju skupu opremu i veliko znanje o obradi digitalnih signala a pružaju veliku efikasnost u smislu tačnosti i brzine. Rezultati predloženih naučnih istraživanja treba da omoguće implementaciju novog hibridnog sistema za akustičnu opservaciju, koji će biti sposoban da prepozna zvučni izvor i locira ga u prostoru, u uslovima prisustva poremećaja. U

slučaju da bude realizovan ovakav sistem bi mogao da bude primenjen kao deo upravljačkog sistema robota i mašina, internet stvari i ostalih uređaja koji zahtevaju interakciju sa okolinom a nemaju na raspolaganju veliki kapacitet izračunavanja.

## LITERATURA

- [1] S. J. Russel and J. Stuart, *Artificial intelligence: a modern approach*. New York: Prentice Hall, 2010.
- [2] S. A. Bunge and I. Kahn, "Cognition : An Overview of Neuroimaging Techniques," in *Encyclopedia of the Neuroscience.*: Elsevier, 2009, p. 10500.
- [3] Geeks for Geeks. [Online]. <https://www.geeksforgeeks.org/turing-test-artificial-intelligence/>
- [4] N. J. Nilsson, "The Physical Symbol System Hypothesis: Status and Prospects," in *50 Years of Artificial Intelligence. Lecture Notes in Computer Science*. Berlin: Springer, 2007.
- [5] L. Gugerty, "Newell and Simon's Logic Theorist: Historical Background and Impact on Cognitive Modeling," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 50, no. 9, p. 880, 2006.
- [6] H. An, J. Shi, X. Wang, and L. Lyu, "Application of Depth-First Search Method in Finding Recirculation in Mine Ventilation System," *Civil Engineering Journal*, no. 3, pp. 286-295, 2017.
- [7] A. Bastian, "Solution Search Simulation The Shortest Step On Chess Horse Using Breadth-First Search Algorithm," *International Journal of Artificial Intelligence Research*, vol. 2, no. 2, 2018.
- [8] T. Imai and A. Kishimoto, "A Novel Technique for Avoiding Plateaus of Greedy Best-First Search in Satisficing Planning," in *Proceedings Of The National Conference On Artificial Intelligence*, 2011, pp. 985-991.
- [9] N. H. Barnouti, S. S. M. Al-Dabbagh, and M. A. S. Naser, "Pathfinding in Strategy Games and Maze Solving Using A\* Search Algorithm," *Journal of Communications and Computer Engineering*, vol. 4, no. 11, pp. 15-26, September 2016.
- [10] S. J. White, "Intention and Prediction in Means-End Reasoning," *American philosophical quarterly*, vol. 55, no. 3, pp. 251-266, 2018.
- [11] (2016, December) GitHub. [Online]. <https://github.com/orsenthil/coursedocs/blob/master/gatech/cs7637/05---means-end->



[analysis.rst](#)

- [12] E. Ivokhin, V. Navrodskiy, and D. Apanasenko, "About Production-Transport Problem Reduction to the Two-Level Problem of Discrete Optimization and Its Application," *Bulletin of Taras Shevchenko National University of Kyiv Economics*, vol. 3, pp. 48-53, June 2018.
- [13] S. Muggleton and H. Watanabe, *Latest Advances In Inductive Logic Programming*. New York: Imperial College Press, 2014.
- [14] N. Savage, "Using Functions for Easier Programming: Functional programming languages automate many of the details underlying specific operations," *Communications of the ACM*, vol. 61, no. 5, pp. 29-30, 2018.
- [15] B. G. Bauchanan and O. R. Duda, "Principles of Rule-Based Expert Systems," *Advances in Computers*, vol. 22, pp. 163-216, Dec. 1983.
- [16] V. G. Sasmito, B. Surarso, and A. Sugiharo, "Application Expert System of Forward Chaining and The Rule Based Reasoning For Simulation Diagnose Pest and Disease Red Onion and Chili Plant," in *Proceedings of The 1st International Conference on Information Systems For Business Competitiveness*, 2011, pp. 392-398.
- [17] M. Flasiniski, *Introduction to Artificial Intelligence.*: Springer, 2016.
- [18] D. H. Jonassen, *Structural knowledge : techniques for representing, conveying, and acquiring structural knowledge*. Hillsdale: Erlbaum, 1993.
- [19] H. Zhuge, "Inheritance rules for flexible model retrieval," *Decision Support Systems*, vol. 22, pp. 379–390, 1998.
- [20] D. D'Souza and P. Shankar, *Modern Applications Of Automata Theory.*: ingapore: World Scientific / Indian Inst Of Science, India, 2012.
- [21] (2018) Tutorialspoint. [Online].  
[https://www.tutorialspoint.com/automata\\_theory/deterministic\\_finite\\_automaton.htm](https://www.tutorialspoint.com/automata_theory/deterministic_finite_automaton.htm)
- [22] P. A. Engelbrecht, *Computational Inteligence*, 2nd ed. Pretoria, South Africa: John Wiley & Sons, 2007.
- [23] I. L. Kuncheva and C. L. Jain, "Nearest neighbor classifier: Simultaneous editing and feature selection," *Pattern Recognition Letters*, vol. 20, no. 11-13, pp. 1149-1156,

November 1999.

- [24] G. Guo, H. Wang, D. Bell, Y. Bi, and K. Greer, "KNN Model-Based Approach in Classification," in *Lecture Notes in Computer Science*, 2003, pp. 986-996.
- [25] R. A. Webb, *Statistical Pattern Recognition*. Chichester: John Wiley & Sons, 2002.
- [26] S. Abe, *Support Vector Machines for Pattern Classification*. London: Springer-Verlag, 2005.
- [27] L. Morissette and S. Chartier, "The k-means clustering technique: General considerations and implementation in Mathematica," *Tutorials in Quantitative Methods for Psychology*, vol. 9, no. 1, pp. 15-24, 2013.
- [28] V. Kumar, "Cluster Analysis: Basic Concepts and Algorithms," in *Introduction to data mining*. New York: Pearson Education, 2004, ch. 8, pp. 487–568.
- [29] I. Ben-Gal, "Bayesian Networks," in *Encyclopedia of Statistics in Quality & Reliability*.: John Wiley & Sons, 2007.
- [30] Q. Chen, A. Whitbrook, U. Aickelin, and C. Roadknight, "Data classification using the Dempster–Shafer method," *Journal of Experimental & Theoretical Artificial Intelligence*, vol. 26, no. 4, pp. 493–517, 2014.
- [31] T. Lukaszewski, "Updating the Evidence in the Dempster-Shafer Theory," *Informatica*, vol. 10, no. 1, p. 127, 1999.
- [32] B. Gerhard, I. Niemelä, and M. Truszczyński, "Foundations of Artificial Intelligence," in *Handbook of Knowledge Representation*.: Elsevier, 2009, pp. 239-284.
- [33] S. Kreutzer, *Pure and Applied Fixed-Point Logics*., 2002.
- [34] A. K. J. F. Benthem, *Preference logic, conditionals and solution concepts in games*.: Inst. for Logic, Language and Computation, 2005.
- [35] J. H. Zimmermann, *Fuzzy Set Theory and its Applications*.: Springer, 2001.
- [36] A. Bárdossy, *Fuzzy rule-based modeling with applications to geophysical, biological, and engineering systems*.: CRC Press, 1995.
- [37] D. Kriesel, *A Brief Introduction to Neural Networks*.: <http://www.dkriesel.com>, 2005.
- [38] R. Rojas, *Neural networks : A systematic introduction*.: Springer, 2000.

- [39] S. S. Haykin, *Neural networks: A comprehensive foundation.*: Prentice Hall, 1999.
- [40] M. Simonović, *Primena Veštačkih Neuronskih Mreža Za Kratkoročno Predviđanje I Analizu Sistema Daljinskog Grejanja.*: Mašinski fakultet Niš, 2016.
- [41] S. Hochreiter, "The Vanishing Gradient Problem During Learning Recurrent Neural Nets and Problem Solutions," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 6, pp. 107-116, 1998.
- [42] M. Črepinšek, S.-H. Liu, and M. Mernik, "Exploration and exploitation in evolutionary algorithms: a survey," vol. 45, no. 3, pp. 35-66, June 2013.
- [43] D. Ashlock, *Evolutionary Computation for Modeling and Optimization.*: Springer, 2005.
- [44] E. Kita, *Evolutionary Algorithms*, E. Kita, Ed.: InTech, 2011.
- [45] O. Kramer, *Machine Learning for Evolution Strategies.*: Springer, 2016.
- [46] B. Thomas, R. Günter, and S. Hans-Paul, "Evolutionary Programming and Evolution Strategies: Similarities and Differences," in *Proceedings of the Second Annual Conference on Evolutionary Programming*, 1997.
- [47] R. Poli, W.B. Langdon, and N.F. McPhee, *A Field Guide to Genetic Programming.*: Creative Commons AttributionNoncommercial-No Derivative Works 2.0 UK, 2008.
- [48] S. McAdams, "Recognition of sound sources and events," in *Thinking in Sound : The Cognitive Psychology of Human Audition.*: Oxford University Press, 1993.
- [49] A. Van Hirtum and D. & Berckmans, "Objective recognition of cough sound as biomarker for aerial pollutants," *Indoor Air*, vol. 14, no. 1, pp. 10-15, 2004.
- [50] J. Turan, "An Acoustic Object Recognition System," *Radioengineering*, vol. 1, no. 1, 1992.
- [51] J. Bosch, J. Janer, F. Fuhrmann, and P. & Herrera, "A Comparison Of Sound Segregation Techniques For Predominant Instrument Recognition In Musical Audio Signals," in *Proceedings of the 13th International Society for Music Information Retrieval Conference, Ismir, 2012*, pp. 559-564.
- [52] T. Heittola, A Klapuri, and T. Virtanen, "Musical Instrument Recognition in Polyphonic Audio Using Missing Feature Approach," in *Proceedings of the 10th International Society for Music Information* ,

- Kobe, 210, pp. 327-332.
- [53] C. Kim and M., R. Stern, "Automatic Threshold Selection Algorithm for Sound Source Separation Based on Inter-microphone Time Difference," *Journal of Latex Class Files*, vol. 6, no. 1, 2007.
- [54] I., B. Yildiz, K. von Kriegstein, and S., J. Kiebel, "From Birdsong to Human Speech Recognition: Bayesian Inference on a Hierarchy of Nonlinear Dynamical Systems," *Plos Computational Biology*, vol. 9, no. 9, pp. 1-16, 2009.
- [55] F. Camastra and A. Vinciarelli, *Machine Learning for Audio, Image and Video Analysis: Theory and Applications.*: Springer, 2015.
- [56] BBC News. (2014, March) Helpful Colin. [Online]. <https://helpfulcolin.com/tag/ear/>
- [57] B. Owsinski, *The recording engineer's handbook*. Boston: Course Technology PTR, 2009.
- [58] I. Osamu and T., Masafumi, N. Tetsuya, "Sound Source Localization Using a Pinna-based Profile Fitting Method," *Ieice Transactions - IEICE*, pp. 263-266, January 2003.
- [59] J., S. Lederman, "Auditory Texture Perception," *Perception*, vol. 8, no. 1, pp. 93-103, 1979.
- [60] A., W. Richards, *Natural Computation.*: MIT Press, 1988.
- [61] X. Li and R., J., Pastore, R., E. Logan, "Perception of acoustic source characteristics: Walking sounds," *The Journal of the Acoustical Society of America*, vol. 90, no. 6, pp. 3036-3049, 1991.
- [62] C. Carello, K., L. Anderson, and A. J. Kunkler-Peck, "Perception of object length by sound. , 9(3), ," *Psychological Science*, vol. 9, no. 3, pp. 211-214, 1998.
- [63] A. J. Kunkler-Peck and M. T. Turvey, "Hearing shape," *Journal Of Experimental Psychology: Human Perception And Performance*, vol. 26, no. 1, pp. 279-294, 2000.
- [64] R., A. Lutfi, "Auditory detection of hollowness," *The Journal Of The Acoustical Society Of America*, vol. 110, no. 2, pp. 1010-1019, 2001.
- [65] B. Kollmeier, T. Brand, and B. Meyer, "Perception of Speech and Sound," in *Springer Handbook of Speech Procesing*. Berlin, Heidelberg: Springer, 2008.
- [66] M. Foley. (2015) KMUW. [Online]. <http://www.kmuw.org/post/loudness-and-fletcher->

munson-curve

- [67] L. Chen, S. Gunduz, and M., T. Ozsu, "Mixed Type Audio Classification with Support Vector Machine," in 2006 IEEE International Conference on Multimedia and Expo, 2006, pp. 781-784.
- [68] R. Radhakrishnan, A. Divakaran, and A. Smaragdis, "Audio analysis for surveillance applications," IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 2005.
- [69] H. Wu, M. Siegel, and P. Khosla, "Vehicle sound signature recognition by frequency vector principal component analysis," in IMTC/98 Conference Proceedings. IEEE Instrumentation and Measurement Technology Conference, 1998, pp. 429-434.
- [70] F., A. Bobick, Natural Object Categorization.: Massachusetts Institute of Technology, 1987.
- [71] M. Keith and V. Barry, Sound-Source Recognition: A Theory and Computational Model.: Massachusetts Institute of Technology, 2000.
- [72] Yu N. C. (2010) Semantic Scholar. [Online].  
<https://www.semanticscholar.org/paper/Improved-Learning-of-Structural-Support-Vector-%3A/4d6df32bb3d7ea7db49041a379984a55b0ecc64c>
- [73] T. Zhang and J., C. C. Kuo, "Hierarchical system for content-based audio classification and retrieval," in Proc.SPIE, 1998, pp. 3527 - 3527.
- [74] L. Lie, Z. Hong-Jiang, and J. Hao, "Content Analysis for Audio Classification and Segmentation," IEEE Transactions On Speech And Audio Processing, vol. 10, no. 7, pp. 504-518, October 2002.
- [75] C. Bishop, Neural networks for pattern recognition. Oxford: Oxford University Press, 1995.
- [76] N. Mhatrea, F. Montalgre-Za, R. Balakrishnanb, and D. Robert, "Sound reception and radiation in a small insect," in Proceedings of the Acoustics 2012 Nantes Conference, 2012, pp. 2287-2292.
- [77] S. Gerlach, S. Goetze, and S. Doclo, "2D Audio-Visual Localization in Home Environments using a Particle Filter," ITG FACHBERICHT, vol. 236, pp. 75-78, 2012.

- [78] S. Costiner et al., "Asymmetric Acoustic Scattering for Structural Health Monitoring," *Journal of the American Helicopter Society*, vol. 59, no. 2, 2014.
- [79] E. Vozáriková, M. Pleva, and J., Cižmár, A. Juhár, "Surveillance System based on the Acoustic Events Detection," *Journal of Electrical and Electronics Engineering*, vol. 4, no. 1, pp. 255-258, January 2011.
- [80] S. U. Bokhari and M., Z. Khan, "Implementation of Signal Processing in Stereo-Scopic Active Sonar Using Heterodyne System," *Journal of Electrical and Electronics Engineering*, vol. 5, no. 2, pp. 31-36, February 2012.
- [81] M. L. Johnson, "Systems and methods of processing information regarding weapon fire location using projectile shockwave and muzzle blast times of arrival data," 2015.
- [82] D. Macho et al., "Automatic Speech Activity Detection, Source Localization, and Speech Recognition on the CHIL SeminarCorpus," in *IEEE International Conference On Multimedia And Expo 2*, 2005, pp. 877-880.
- [83] C. R. Rowell et al., "Three-Dimensional Volcano-Acoustic Source Localization at Karymsky Volcano, Kamchatka, Russia," *Journal of Volcanology and Geothermal Research*, vol. 283, pp. 101-115, August 2014.
- [84] S. R. Martín, M. Genescà, J. Romeu, and A. Clot, "Aircraft localization using a passive acoustic method. Experimental test," *Aerospace Science and Technology*, vol. 48, pp. 246-253, 2016.
- [85] K. Grabowski et al., "Time–distance domain transformation for Acoustic Emission source localization in thin metallic plates," *Ultrasonics*, vol. 68, pp. 142–149, 2016.
- [86] C. Tan et al., "A low-cost centimeter-level acoustic localization system without time synchronization," *Measurement*, vol. 78, pp. 73–82, 2016.
- [87] D. Nikolić, S. Y. Kim, and R. Allen, "Identification of sound-localization cues in the HRTF of the bat-head model," *Biomedical Signal Processing and Control*, vol. 7, pp. 270–277, 2012.
- [88] S. Carlile, "The Physical and Psychophysical Basis of Sound Localization," in *Virtual Auditory Space: Generation and Applications.*: R.G. Landes Company, 2002, ch. 2.
- [89] Z. Lin, Z. Xiao-Yan, C. Xu, and W. Zhen-Yang, "Binaural sound source localization

- based on sub-band SNR estimation," *International Journal Of Multimedia And Ubiquitous Engineering*, vol. 10, no. 5, pp. 303-314, 2015.
- [90] E. A. Macpherson and A. T. Sabin, "Vertical-plane sound localization with distorted spectral cues," *Hearing Research*, vol. 306, pp. 76–92, 2013.
- [91] F. Georgiou and B. Fazenda. (2012, September) Research Gate. [Online].  
[https://www.researchgate.net/publication/272327880\\_RELATIVE\\_DISTANCE\\_PERCEPTION\\_OF\\_SOUND\\_SOURCES\\_IN\\_CRITICAL\\_LISTENING\\_ENVIRONMENT\\_VIA\\_BINAURAL\\_REPRODUCTION](https://www.researchgate.net/publication/272327880_RELATIVE_DISTANCE_PERCEPTION_OF_SOUND_SOURCES_IN_CRITICAL_LISTENING_ENVIRONMENT_VIA_BINAURAL_REPRODUCTION)
- [92] A. Saxena and A. Y. Ng, "Learning sound location from a single microphone," in *2009 IEEE International Conference on Robotics and Automation*, 2009, pp. 1737-1742.
- [93] S. Argentieri, P. Danes, and P. Soueres, "A survey on sound source localization in robotics: From binaural to array processing methods," *Computer Speech & Language*, vol. 34, no. 1, pp. 87-112, November 2015.
- [94] J. A. Belloch, A. Gonzalez, A. M. Vidal, and M. Cobos, "On the performance of multi-GPU-based expert systems for acoustic localization involving massive microphone arrays," *Expert Systems With Applications*, vol. 42, no. 13, pp. 5607-5620, 2015.
- [95] X. Wu, G. Wang, D. Dai, and M. Tong, "Accurate acoustic energy-based localization with beacon position uncertainty in wireless sensor networks," *Journal Of Network And Computer Applications*, vol. 43, pp. 76-83, 2014.
- [96] G. Athanasopoulos, H. Brouckxon, and W. Verhelst, "Sound Source Localization for Real-World Humanoid Robots," in *11th International Conference on Signal Processing*, 2012, pp. 131-136.
- [97] H. Nakashima and T. Mukai, "3D sound source localization system based on learning of binaural hearing," in *2005 IEEE International Conference on Systems, Man and Cybernetics*, 2005, pp. 3534-3539.
- [98] L. A. Seewald, J. L. Gonzaga, M. R. Veronez, V. P. Minotto, and C. R. Jung, "Combining SRP-PHAT and two Kinects for 3D Sound Source Localization," *Expert Systems With Applications*, vol. 41, pp. 7106-7113, November 2014.
- [99] Markus V. S. L., Wallace A. M., Leonardo O. N., Luiz W. P. B, and Tadeu N. F,

"Efficient Steered-Response Power Methods for Sound Source Localization Using Microphone Arrays," Computing Research Repository, 2014.

- [100] gfai tech GmbH. (2014, May) Youtube. [Online].  
<https://www.youtube.com/watch?v=obrkJZoGXc0Y>
- [101] M., B. Dehkordi, "Sound Source Localization with CS Based Compressed Neural Network," American Journal of Signal Processing, vol. 1, pp. 1-5, August 2012.
- [102] Q. Huang and T. Wang, "Acoustic source localization in mixed field using spherical microphone arrays," EURASIP Journal on Applied Signal Processing, vol. 1, 2014.
- [103] M. B. Dehkordi, H. R. Abutalebi, and H. Ghanei, "A compressive sensing based compressed neural network for sound source localization," in 2011 International Symposium on Artificial Intelligence and Signal Processing (AISP), 2011, pp. 6-10.
- [104] S. Paulose, E. Sebastian, and B. Paul, "Acoustic Source Localization," International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, vol. 2, no. 2, pp. 933-939, February 2013.
- [105] R. Bucher and D. Misra, "A Synthesizable VHDL Model of the Exact Solution for," VLSI Design, vol. 15, no. 2, pp. 507-520, 2002.
- [106] C. Park, J. Jeon, and Y. Kim, "Localization of a sound source in a noisy environment by hyperbolic curves in quefreny domain," Journal Of Sound And Vibration, vol. 333, pp. 5630-5640, 2014.
- [107] B. T. Polyak, "Newton-Kantorovich Method and Its Global Convergence," Journal Of Mathematical Sciences, vol. 133, no. 4, pp. 1513-1523, 2006.
- [108] N.I.M. Gould and Leyffer S., "An Introduction to Algorithms for Nonlinear Optimization," in Frontiers in Numerical Analysis. Berlin: Springer, 2003, p. 79.
- [109] L. Chen, "A high-order modified Levenberg–Marquardt method for systems of nonlinear equations with fourth-order convergence," Applied Mathematics and Computation, vol. 285, pp. 79-93, July 2016.
- [110] T. B. Fang, "Simple Solutions of Hyperbolic and Related Position Fixes," Aerospace and Electronic Systems, vol. 26, pp. 748-753, October 1990.
- [111] Hing C. S., "A comparative study of two discrete-time phase delay estimators," IEEE



- Transactions on Instrumentation and Measurement, vol. 54, pp. 2501-2504, 2005.
- [112] A. Hero and S. C. Schwartz, "Hero, Alfred and SCHWARTZ, SC," *Acoustics, Speech and Signal Processing*, vol. 33, pp. 38-45, March 1985.
- [113] P. Roth, "Effective measurements using digital signal analysis," *IEEE Spectrum*, vol. 8, pp. 62-70, 1971.
- [114] H. Khaddour, "A Comparison of Algorithms of Sound Source Localization Based on Time Delay Estimation," *Elektrorevue*, vol. 2, no. 1, pp. 31-37, April 2011.
- [115] G. C. Carter, H. A. Nuttall, and Cable G. P., "Carter, G.Clifford and H. Nuttall, Albert and G. Cable, Peter," in *Proceedings of the IEEE*, 1973, pp. 1497-1498.
- [116] R. S. Marinescu, A. Buzo, H. Cucu, and C. Burileanu, "Extensive Evaluation Experiments for the Accumulated Cross-Power Spectrum Methods for Time Delay Estimation," in *2013 7th Conference on Speech Technology and Human - Computer Dialogue*, December 2013, pp. 1-6.
- [117] L. Chen, Y Liu, F. Kong, and N. He, "Acoustic Source Localization Based on Generalized Crosscorrelation Time-delay Estimation ," *Advanced in Control Engineering and Information Science*, vol. 15, pp. 4912 – 4919, 2011.
- [118] C. L. Aguilar, I. Mohino, and L. A. Perez, "A comparative study of time-delay estimation techniques for convolutive speech mixtures," in *Advances in Computer Science*, 2011, pp. 291-296.
- [119] M. Kovandžić, V. Nikolić, A. Al-Noori, I. Ćirić, and M. Simonović, "Near field acoustic localization under unfavorable conditions using feedforward neural network for procesing time difference of arrival," *Expert Systems with Applications*, vol. 71, no. 1, pp. 138-146, April 2017.
- [120] S. Kalafat and M. G. Sause, "Acoustic emission source localization by artificial neural networks," *Structural Health Monitoring*, vol. 14, no. 6, pp. 633–647, 2015.
- [121] C., M. Killion, "New Thinking on Hearing in Noise: A Generalized Articulation Index," *Seminars in Hearing*, vol. 23, no. 1, pp. 57-76, January 2002.
- [122] C. Panagiotakis and G. Tziritas, "A speech/music discriminator using RMS and zero-crossings," in *2002 11th European Signal Processing Conference*, 2002, pp. 1-4.

- [123] J. Y. Choi, E. R. Hu, and T. K. Perrachione, "Varying acoustic-phonemic ambiguity reveals that talker normalization is obligatory in speech processing," *Attention, Perception & Psychophysics*, vol. 80, no. 3, pp. 784–797, 2018.
- [124] S. Haykin, *Adaptive Filter Theory*. New York, USA: Prentice-Hall, 1996.
- [125] S. Dwivedi and P. Shrivastav, "Survey Paper on Digital IIR Filter Using Evolutionary Algorithms," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 3, no. 6, pp. 5106-5111, June 2015.
- [126] S. Dhull, S. Arya, and O. P. Sahu, "Comparison of time-delay estimation techniques in acoustic environment," *International Journal of Computer Applications*, vol. 8, no. 9, pp. 29–31, 2010.
- [127] K., V. Madisetti, *The Digital Signal Processing Handbook.*: CRC Press, 2009.

## BIOGRAFIJA

Kandidat Marko Kovandžić rođen je 16.1.1974 u Aleksincu. Završio je osnovnu školu “Aca Milojević” u Aleksincu kao nosilac diplome “Vuk Karadžić” i diplome “Mihajlo Petrović Alas”. Prve dve godine srednje škole pohađao je vojnu gimnaziju “Bratstvo jedinstvo” u Beogradu a srednjoškolsko obrazovanje okončao je u gimnaziji “Drakče Milovanović” u Aleksincu. Diplomirao je, 2003.god., na Katedri za Precizno mašinstvo i automatiku Mašinskog fakulteta Univerziteta u Nišu, na temu pod naslovom: “Projektovanje zakona upravljanja za multivarijabilne sisteme primenom klasičnih i savremenih tehnika”.

Profesionalnu karijeru započinje u fabrici cevi i profila “Alpos” iz Aleksinca, od 2004. do 2011. god, na samom početku kao komercijalista a nakon toga kao šef informatičkog odeljenja. U toku ovog perioda, kao rezultat pohađanja velikog broja kurseva u zemlji i inostranstvu i učestvovanja u realnim projektima, specijalizuje se za poslovni softver i baze podataka. Kao konsultant za ERP sistem i poslovne procese angažovan je, između 2011. i 2013. god, u fabrici silosa, profila i opreme “Utva Silosi” iz Kovina. U toku 2014. radi, kao ERP programer, u preduzeću “Atenic commerce” iz Čačka, na projektu otvaranja nove fabrike cevi. U periodu između 2017. i 2018. god. rukovodi projektom implementacije ERP sistema u novoosnovanoj fabrici za proizvodnju automobilskih delova, delova motora i opštih industrijskih delova, “Mei Ta Europe” iz Obrenovca. Trenutno je zaposlen u geodetskoj agenciji “Geokovandžić” iz Aleksinca.

Do sada je publikovao nekoliko naučnih radova, u domaćim i internacionalnim časopisima, iz oblasti upravljanja, robotike i veštačke inteligencije.

## ИЗВЕШТАЈ О ОЦЕНИ ДОКТОРСКЕ ДИСЕРТАЦИЈЕ

### ПОДАЦИ О КАНДИДАТУ

Презиме, име једног родитеља и име	Кованцић Небојша Марко
Датум и место рођења	16.01.1974, Алексинац, Србија

### Основне студије

Универзитет	Универзитет у Нишу
Факултет	Машински факултет
Студијски програм	Машинско инжењерство
Звање	Дипломирани инжењер машинства
Година уписа	1993
Година завршетка	2003
Просечна оцена	7,54

### Магистарске студије

Универзитет	
Факултет	
Студијски програм	
Звање	
Година уписа	
Година завршетка	
Просечна оцена	
Научна област	
Наслов завршног рада	

### Докторске студије

Универзитет	Универзитет у Нишу
Факултет	Машински факултет
Студијски програм	Машинско инжењерство 2014
Година уписа	2010
Остварен број ЕСПБ бодова	150
Просечна оцена	10.00

### НАСЛОВ ТЕМЕ ДОКТОРСКЕ ДИСЕРТАЦИЈЕ

Наслов теме докторске дисертације	Оптимално препознавање и локализација извора звука применом метода вештачке интелигенције
Име и презиме ментора, звање	др Властимир Николић, редовни професор
Број и датум добијања сагласности за тему докторске дисертације	8/20-01-002/19-021 од 25. 02.2019.

### ПРЕГЛЕД ДОКТОРСКЕ ДИСЕРТАЦИЈЕ

Број страна	164
Број поглавља	9
Број слика (шема, графикона)	85
Број табела	3
Број прилога	-

МАШИНСКИ ФАКУЛТЕТ У НИШУ			
Примљено: 06.11.2019.			
Орг.јед.	Број	Прилог	Вредности
	6080	392	19



**ПРИКАЗ НАУЧНИХ И СТРУЧНИХ РАДОВА КАНДИДАТА  
који садрже резултате истраживања у оквиру докторске дисертације**

Р. бр.	Аутор-и, наслов, часопис, година, број волумена, странице	Категорија
1	<p><b>Kovandžić, M., Nikolić, V., Al-Noori, A., Ćirić, I., Simonović, M., Near field acoustic localization under unfavorable conditions using feedforward neural network for processing time difference of arrival</b>, Expert Systems With Applications, 2017, 71, 138-146. doi: 10.1016/j.eswa.2016.11.030</p> <p><i>У овом раду истражени су потенцијали вештачких неуронских мрежа у решавању проблема хиперболичног позиционирања, који настаје као резултат одређивања просторне позиције објекта на основу кашњења међу сигнаlima прикупљеним на различитим просторним позицијама у околини извора звука. Проблем додатно компликују поремећаји, различитог порекла који, у реалним условима, не могу бити избегнути. Уз помоћ великог броја акустичних узорака, испитане су перформансе система у погледу поставке сензора, конфигурације мреже и параметара за њено тренирање, ради проналажења оптималне конфигурације. Експеримент даје корисне смернице за практичну реализацију вештачких система за локализацију извора звука. Поступак не захтева експертско знање због чега је прихватљив широком кругу корисника.</i></p>	M21
2	<p><b>Simonović M., Kovandžić M., Nikolić V., Stojčić M., Knežević D., Artificial Neural Network Application for Temporal Properties of Acoustic Perception</b>, Facta Universitatis, Series: Mechanical Engineering, 2019, accepted</p> <p><i>У овом раду је приказана примена вештачке неуронске мреже за одређивање темпоралних карактеристика код акустичке перцепције. Акустичко перцепција је прилично добро објашњена у литератури али је недовољно примењена у пракси.. Коришћење вештачких неуронских мрежа је јако погодно за процесирање акустичких сигнала али је потребно реализовати процедуру која подразумева прикупљање односно аквизицију података, филтрирање, издвајање карактеристика и категоризацију акустичних сигнала. Представљен је експеримент који потврђује теоријске претпоставке да темпорална резолуција акустичке перцепције, коришћењем вештачких неуронских мрежа, веома зависи од процедуре издвајања карактеристика акустичких сигнала. Експеримент истражује два елементарна процеса акустичке перцепције: акустичко препознавање и акустичку локализацију. У оба случаја се уводи вештачка неуронска мрежа као алат за процесирање сигнала због своје једноставности, универзалности и одличних перформанси.</i></p>	M24
3	<p><b>Kovandžić, M., Nikolić, V., Simonović, M., Ćirić, I., Al-Noori, A., Soft robot positioning using artificial neural network</b>, Facta Universitatis, Series: Automatic Control and Robotics, 2019, 18 (1), 19 – 30. doi:10.22190/FUACR1901019K</p> <p><i>У овом раду је приказан експеримент који истражује перформансе вештачких неуронских мрежа у решавању инверзног кинематичког проблема код меког (еластичног) робота. За ову потребу, конструисан је једноставни меки робот, од градивних блокова који су нанизани на три гумена црева и покретачког система за обезбеђивање хидрауличног притиска. Аксијална деформација једног од црева, изазвана повећањем притиска, док су друга два црева у опуштеном стању, доводи до савијања робота. За моделирање понашања робота примењена је вештачка неуронска мрежа која је тренирана помоћу скупа улазно-излазних података. Показало се да је грешка позиционирања у оквиру прихватљивих граница чиме је оправдана примена вештачких неуронских мрежа за управљање меким роботима.</i></p>	M51
4	<p><b>Kovandžić, M., Miltenović, A., Application Of Triz Method For Improvement Of Black Welded Tubes Classification Process</b>, Machine Design, 2016, 8, 27-32. <a href="http://www.mdesign.ftn.uns.ac.rs/download/v8n1/p5.pdf">http://www.mdesign.ftn.uns.ac.rs/download/v8n1/p5.pdf</a></p> <p><i>Овај рад презентује примену систематске процедуре за решавање техничких проблема (руски Теорија решенија изобретатељских задатака-ТРИЗ) за унапређење поступка класификације црних шавних цеви. Полазно стање подразумевало је класификацију цеви од стране људског оператора на бази непосредног опажања. Ефикасност поступка је често нарушавана због опадања форме и концентрације оператора. Применом ТРИЗ методе, одлучено је о предузимању неколико акција које су водиле олакшавању процеса и смањењу умора.</i></p>	M52



	<i>Ове акције су имале за последицу очување радних способности оператора у дужем временском интервалу. На крају је процес класификације комплетно аутоматизован чиме је људски фактор елиминисан из процеса.</i>	
5	<p><b>M. Kovandžić, V. Nikolić, R. Nikolić, M. Stojičić, D. Knežević: Temporal Properties of Acoustic Perception by Artificial Neural Networks using Standard Computer Equipment, Proceedings of XIV International Conference on Systems, Automatic Control and Measurements, ISBN 978-86-6125-205-1 (FEE), SAUM 2018, Niš, November 14-16, 2018, 215-219</b></p> <p><i>Акустична перцепција је добро описана у литератури али је недовољно примењена и потврђена у пракси. Експериментални резултати су одлични али многа питања остају отворена када дође до примене у реалним условима. Коришћење неуронских мрежа омогућава да процесирање акустичних сигнала буде веома добро са математичког становишта. Међутим, велики посао треба урадити да би то било могуће. Та процедура укључује аквизицију података, филтрирање, избор узорака. Ове технике захтевају много више ресурса у односу на неуронску мрежу и представљају ограничавајући фактор за имплементацију. Овај рад истражује и описује потпуну процедуру акустичне перцепције у односу на време и идентификацију ограничења.</i></p>	M33

### ИСПУЊЕНОСТ УСЛОВА ЗА ОДБРАНУ ДОКТОРСКЕ ДИСЕРТАЦИЈЕ

Кандидат испуњава услове за оцену и одбрану докторске дисертације који су предвиђени Законом о високом образовању, Статутом Универзитета и Статутом Факултета.

ДА НЕ

Кандидат је дипломирани машински инжењер, има одобрену тему докторске дисертације, објавио је већи број научних и стручних радова и поднео докторску дисертацију одговарајуће садржине, обима и квалитета, у складу са одобреном темом докторске дисертације.

### ВРЕДНОВАЊЕ ПОЈЕДИНИХ ДЕЛОВА ДОКТОРСКЕ ДИСЕРТАЦИЈЕ

Кратак опис појединих делова дисертације

Разматрана теза састоји се из 9 поглавља којима претходе резиме на српском и енглеском језику, списак кључних речи, садржај и листе табела и слика. На крају се налази списак са 127 референци.

Уводно поглавље објашњава мотив предузетог истраживања, његов предмет и циљ и даје уводне напомене о примењеним методама. На крају овог дела дат је кратак преглед садржаја дисертације.

Друго поглавље посвећено је теоретској анализи вештачке интелигенције, са акцентом на рачунској интелигенцији због њеног значаја на предмет дисертације. Објашњени су основни методи вештачке интелигенције и дата њихова општа класификација, на основу приступа решавању проблема. За различите класе вештачке интелигенције представљени су репрезентативни алгоритми.

Након теоретске анализе, као најефикасније средство за обраду података, у поступцима акустичне опсервације, изабране су вештачке неуронске мреже. Због тога треће поглавље дисертације објашњава структуру и принцип функционисања неуронских мрежа, даје њихову класификацију и објашњава основне методе за њихово обучавање на основу скупа експерименталних података.

Четврто поглавље посвећено је еволуционом рачуну пошто је примењен за конфигурацију поступка предобраде сигнала приликом акустичне опсервације. Дат је преглед основних оператора и смернице за успешно извођење еволуционих алгоритама. Објашњена је подела, на 4 опште категорије, у зависности од примењених оператора и намене еволуционих алгоритама.

Препознавање акустичних извора разматрано је у петом поглављу. У првом делу су разматрани основни перцептивни квалитети и механизми којима се користе жива бића у сврху акустичног препознавања. У наставку су објашњена техничка средства и методе акустичног препознавања, по правилу инспирисане способностима живих бића. На крају поглавља описан је експеримент акустичног препознавања, предузет у оквиру научноистраживачког рада, презентовани резултати примене вештачких неуронских мрежа и одговарајући закључци.

За потребе овог експеримента снимљено је и прикупљено 685 дигиталних узорака звука, на стандардној фреквенцији од 44.1 kHz у трајању од 1s. Ради симулирања различитих нивоа апстракције узорци звука су припадали трима различитим категоријама – цврчак као један од симбола акустичне комуникације међу инсектима, мува као један од најчешћих инсеката које човек сусреће па постоје потенцијали практичне примене њихове детекције и узорци звука различитог порекла (људски глас, животиње, природни феномени, возила, машине,...)



Шесто поглавље посвећено је акустичној локализацији. Описани су основни физички феномени и методе локализација извора звука од стране живих бића и вештачких направа. Преглед базичних метода за локализацију извора звука праћен је прегледом поступака за предобраду акустичних сигнала намењених овој сврси. На крају поглавља описан је експеримент акустичне локализације предузет у оквиру истраживачког рада, презентовани су резултати експеримента са применом вештачких неуронских мрежа, применом генетског алгорита за оптимизацију путање извора звука и на основу њих, донети закључци.

За потребе овог експеримента коришћена су два модела.

Први модел је сачињен од 12 паралелно везаних минијатурних звучника, сваки смештен у центру странице правилног додекахедрона.

Други модел је састављен од пинг-понг лоптице у којој је смештен минијатурни вибрациони мотор тако да, ударајући у њене зидове, производи акустичне вибрације. Снимљено је преко 2000 узорака, уз помоћ 8 просторно распоређених микрофона, и сваком од њих придружене су одговарајуће просторне координате добијене ласерским мерењем.

Утицај поремећаја као и репрезентативне методе предобrade сигнала које за циљ имају смањење овог утицаја на тачност акустичне опсервације разматрани су у седмом поглављу. Подаци сакупљени у експериментима акустичне опсервације, представљеним у два претходна поглавља, употребљени су за извођење експеримента са методама предобrade акустичног сигнала.

Предложен је општи метод за конфигурисање поступка предобrade акустичних сигнала и презентовани позитивни ефекти његове примене.

Други део експеримента описаног у петом поглављу је овде презентован са циљем испитивања робустности постигнутог решења у односу на присуство поремећаја и мерама намењених превенцији њиховог утицаја. Употребом дигиталног филтера другог реда са неограниченим импулсним одзивом приметно је смањена грешка препознавања како у зонама корелисаног утицаја тако и у зонама утицаја случајног шума.

У овом поглављу је урађен још један додатни експеримент са циљем конфигурисања комплетног поступка предобrade сигнала намењених акустичкој локализацији у блиском пољу. Показано је да правилан избор конфигурационих параметара обезбеђује драстично смањење грешке израчунатих кашњења међу сигнаlima и тиме омогућује локализацију извора у блиском пољу.

У осмом поглављу предложен је модел хибридног ситета за акустичну опсервацију пројектован у складу са принципима ефикасности, једноставности и универзалности. Модел остварује стратегију опсервације комбиновањем различитих метода вештачке интелигенције са циљем прикупљања што веће количине корисних информација о посматраном објекту. На крају је предвиђена и пост обрада резултата акустичке локализације применом Калмановог филтра.

Применом репрезентације знања уз помоћ правила омогућено је ефикасно контролисање поступка акустичке перцепције каоја укључује интеракцију између различитих метода. Смештањем различитих стратегија акустичке перцепције и одговарајућих конфигурационих параметара у базу података остварена је универзалност поступка.

Последње поглавље резервисано је за извођење општих закључака истраживачког рада и представљање планова даљег истраживања на пољу акустичне опсервације околине применом метода вештачке интелигенције.



## ВРЕДНОВАЊЕ РЕЗУЛТАТА ДОКТОРСКЕ ДИСЕРТАЦИЈЕ

### Ниво остваривања постављених циљева из пријаве докторске дисертације

Циљеви постављени у пријави докторске дисертације су остварени, уз поштовање предложеног оквирног садржаја.

Истраживања су по садржају обухватила више актуелних научноистраживачких праваца. Систематизовано је сазнање о феномену акустичне опсервације које потиче из најразличитијих научних дисциплина. Осим саме интерпретације сигнала обрађене су и пратеће теме као што су прикупљање, предобрада и постобрада сигнала. Анализа интелигентних алгоритама истакла је ефикасност вештачких неуронских мрежа у интерпретацији акустичних сигнала док је за одређивање адекватних вредности конфигурационих параметара, примењен еволуциони рачун. Због тога је, у зрелој фази истраживања, фокус усмерен ка овим методама рачунске интелигенције.

У складу са резултатима анализе осмишљени су поступци акустичне опсервације намењени препознавању извора звука и његовој локализацији у непосредној околини. Робустност опсервације, у реалним условима, обезбеђена је јединственим поступком предобраде акустичних сигнала. Примена овог поступка предвиђена је у оквиру алгоритма стохастичке оптимизације који се обавља на основу скупа расположивих узорака.

Тестирање поступака извршено је независно у експерименталном окружењу креираном за ову намену. Резултати су доказали да се применом вештачке интелигенције може извршити успешно препознавање и локализација извора звука, у реалним условима, уз утрошак скромних ресурса. На крају је демонстрирана могућност реализације хибродног система, оствареног комбиновањем предложених поступака помоћу елемената логичког расуђивања, са циљем повећања количине информација о предмету акустичне опсервације.

### Вредновање значаја и научног доприноса резултата дисертације

Обрађивана тема докторске дисертације је веома значајна и актуелна, како у научном смислу, тако и смислу практичне применљивости. Поднета докторска дисертација представља оригиналан и вредан научни и стручни допринос кандидата. Научни допринос разматраног рукописа и објављених радова се пре свега огледа у следећем:

- Развој система за аквизицију и обраду акустичних сигнала намењених препознавању
- Развој система за аквизицију и обраду акустичних сигнала намењених локализацији у блиском пољу
- Развој алгоритма за препознавање (категоризацију) звука на основу амплитудно-фреквентног спектра
- Развој система за локализацију извора звука, у блиском пољу, на основу кашњења међу сигналима регистрованим на различитим просторним позицијама
- Развој адаптивног алгоритма, универзалне примене, за конфигурисање предобраде акустичних сигнала са циљем унапређења тачности акустичне опсервације.
- Развој хибродног система са циљем проширења спектра информација о објекту опсервације.

### Оцена самосталности научног рада кандидата

Кандидат дипл.инг Марко Кованџић је показао значајно теоријско и практично знање, као и висок ниво самосталности, систематичности и креативности у бављењу научно-истраживачким радом. Кандидат је приказао детаљну, свеобухватну и квалитетну анализу постојеће научне литературе из области теме докторске дисертације. Познавање литературе и стечена знања из више области је искористио да на креативан начин осмисли, формулише и примени научни приступ за развој система вештачке интелигенције намењених опсервацији акустичних сигнала. Креирао је одговарајуће симулационо окружење које омогућава развој и примену предложених метода.

Неки од научних резултата представљених у разматраној тези презентирани су у оквиру научних радова који су штампани у часописима и представљени на конференцијама.



## ЗАКЉУЧАК

На основу прегледа поднете радне верзије докторске дисертације и увидом у објављене научне радове кандидата, чланови Комисије за оцену и одбрану докторске дисертације закључују следеће:



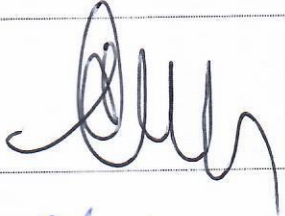


- Поднети рукопис одговара теми докторске дисертације одобреној од стране Наставно научног већа Машинског факултета у Нишу и Научно стручног већа Универзитета у Нишу;
- Докторска дисертација представља оригиналан и вредан научни и стручни допринос веома актуелној и значајној проблематици опсервације звука у реалним условима применом метода вештачке интелигенције
- Научни допринос и оригиналност дисертације показани су објављивањем већег броја радова;
- Докторска дисертација је адекватно конципирана и технички квалитетно урађена;
- Резултати приказаног научног рада имају висок степен општости и применљивости;
- Кандидат поседује висок ниво теоријских и практичних знања из више области потребних за решавање комплексних проблема опсервације звука применом метода вештачке интелигенције и добро је упознат са досадашњим научним достигнућима;
- Кандидат је показао висок ниво самосталности и систематичности у бављењу научно-истраживачким радом, као и креативан приступ формулацији и решавању разматраних проблема.

Имајући у виду напред наведено, Комисија предлаже Наставно научно већу Машинског факултета у Нишу да се поднети рукопис кандидата **Марка Кованџића**, дипломираног машинског инжењера, под називом:

### „ОПТИМАЛНО ПРЕПОЗНАВАЊЕ И ЛОКАЛИЗАЦИЈА ИЗВОРА ЗВУКА ПРИМЕНОМ МЕТОДА ВЕШТАЧКЕ ИНТЕЛИГЕНЦИЈЕ“

прихвати као докторска дисертација, а кандидат позове на усмену јавну одбрану.

## КОМИСИЈА

Број одлуке ННВ о именовану Комисије		НСВ број 8/20-01-007/19-022	
Датум именовања Комисије		30.10.2019.	
Р. бр.	Име и презиме, звање		Потпис
1.	др Милош Симоновић, доцент	председник	
	Аутоматско управљање и роботика (Научна област)	Машински факултет Универзитета у Нишу (Установа у којој је запослен)	
2.	др Властимир Николић, редовни професор	ментор	
	Аутоматско управљање и роботика (Научна област)	Машински факултет Универзитета у Нишу (Установа у којој је запослен)	
3.	др Драган Антић, редовни професор	члан	
	Аутоматика (Научна област)	Електронски факултет Универзитета у Нишу (Установа у којој је запослен)	
4.	др Иван Ћирић, доцент	члан	
	Аутоматско управљање и роботика (Научна област)	Машински факултет Универзитета у Нишу (Установа у којој је запослен)	
5.	др Жарко Ђојбашић, редовни професор	члан	
	Аутоматско управљање и роботика (Научна област)	Машински факултет Универзитета у Нишу (Установа у којој је запослен)	

Датум и место:  
У Нишу, новембра 2019.



## ИЗЈАВА О КОРИШЋЕЊУ

Овлашћујем Универзитетску библиотеку „Никола Тесла“ да, у Дигитални репозиторијум Универзитета у Нишу, унесе моју докторску дисертацију, под насловом:

Оптимално препознавање и локализација извора звука применом метода вештачке интелигенције

Дисертацију са свим прилозима предао/ла сам у електронском облику, погодном за трајно архивирање.

Моју докторску дисертацију, унету у Дигитални репозиторијум Универзитета у Нишу, могу користити сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons), за коју сам се одлучио/ла.

1. Ауторство (CC BY)
2. Ауторство – некомерцијално (CC BY-NC)
3. Ауторство – некомерцијално – без прераде (CC BY-NC-ND)
4. Ауторство – некомерцијално – делили под истим условима (CC BY-NC-SA)
5. Ауторство – без прераде (CC BY-ND)
6. Ауторство – делили под истим условима (CC BY-SA)

(Молимо да подвучете само једну од шест понуђених лиценци; опис лиценци дат је у наставку текста).

У Нишу, 5.2.2020.

Аутор дисертације: Марко Кованџић

Потпис аутора дисертације:

**ИЗЈАВА О ИСТОВЕТНОСТИ ПШТАМПАНОГ И ЕЛЕКТРОНСКОГ ОБЛИКА  
ДОКТОРСКЕ ДИСЕРТАЦИЈЕ**

Име и презиме аутора: Марко Кованџић

Наслов дисертације: Оптимално препознавање и локализација извора звука применом метода  
вештачке интелигенције

Ментор: prof. dr Vlastimir Nikolić

Изјављујем да је штампани облик моје докторске дисертације истоветан електронском облику,  
који сам предао/ла за уношење у **Дигитални репозиторијум Универзитета у Нишу**.

У Нишу, 5.2.2020.

Потпис аутора дисертације:



## ИЗЈАВА О АУТОРСТВУ

Изјављујем да је докторска дисертација, под насловом:

Оптимално препознавање и локализација извора звука применом метода вештачке интелигенције која је одбрањена на Машинском факултету Универзитета у Нишу:

- резултат сопственог истраживачког рада;
- да ову дисертацију, ни у целини, нити у деловима, нисам пријављивао/ла на другим факултетима, нити универзитетима;
- да нисам повредио/ла ауторска права, нити злоупотребио/ла интелектуалну својину других лица.

Дозвољавам да се објаве моји лични подаци, који су у вези са ауторством и добијањем академског звања доктора наука, као што су име и презиме, година и место рођења и датум одбране рада, и то у каталогу Библиотеке, Дигиталном репозиторијуму Универзитета у Нишу, као и у публикацијама Универзитета у Нишу.

У Нишу, 5.2.2020

Аутор дисертације: Марко Кованџић

Потпис аутора дисертације:

