SVEUČILIŠTE U ZAGREBU

**FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA**

DIPLOMSKI RAD br. 2806

**IZVEDBA METODE STROJNOG UČENJA ZA RAZLIKOVANJE IZVORA AKUSTIČKIH EMISIJA U BILJNIM TKIVIMA NA UGRADBENOM SUSTAVU NISKE POTROŠNJE**

Renato Gracin

SVEUČILIŠTE U ZAGREBU

**FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA**

DIPLOMSKI RAD br. 2806

**IZVEDBA METODE STROJNOG UČENJA ZA RAZLIKOVANJE IZVORA AKUSTIČKIH EMISIJA U BILJNIM TKIVIMA NA UGRADBENOM SUSTAVU NISKE POTROŠNJE**

Renato Gracin

**SVEUČILIŠTE U ZAGREBU**

**FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA**

**Zagreb, 11. ožujka 2022.**

DIPLOMSKI ZADATAK br. 2806

Pristupnik: **Renato Gracin (0036508945)**

Studij: Računarstvo

Profil: Računalno inženjerstvo

Mentor: doc. dr. sc. Dinko Oletić

Zadatak: **Izvedba metode strojnog učenja za razlikovanje izvora akustičkih emisija u biljnim tkivima na ugradbenom sustavu niske potrošnje**

Opis zadatka:

Istražiti metode nenadziranog i polu-nadziranog strojnog učenja, primjenjive za razlikovanje vrste izvora ultrazvučnih akustičkih emisija u biljnim tkivima za prijenos vode. Odabrati metodu odabira značajki prikladnu za primjenu na zadanom skupu podataka i predložiti skup značajki signala. Odabrati metodu klasifikacije koja omogućava izvedbu na ugradbenom sustavu, vodeći računa o programskoj složenosti i memorijskom zauzeću. Izvesti odabrani algoritam ekstrakcije značajki i klasifikacije na ugradbenom procesoru porodice Arm Cortex-M STM32L4. Odabrati metodologiju i izvesti ispitno okruženje. Provjeriti točnost algoritma na snimljenom skupu podataka i trajanje izvođenja na odabranom ugradbenom sustavu. Procijeniti doprinos algoritma ukupnoj potrošnji ugradbenogsustava.

Rok za predaju rada: 27. lipnja 2022.

Hvala svome mentoru na pomoći.

Sadržaj

[Uvod 1](#_Toc104971774)

[1. Pregled teme i prijašnjih doprinosa detekcije kavitacije 2](#_Toc104971775)

[1.1. Opis kavitacije kod biljaka 2](#_Toc104971776)

[1.2. Trenutna metodologija detekcije kavitacije 2](#_Toc104971777)

[2. Ekstrakcija značajki (Predobrada) akustičnih emisija 3](#_Toc104971778)

[2.1. Izbor i definiranje značajki akustičnih emisija 3](#_Toc104971779)

[2.2. Postupak izračuna matrice značajki akustičnih emisija 3](#_Toc104971780)

[2.3. Ekvilizacija sirovoj emisija 3](#_Toc104971781)

[2.4. Analiza i normalizacija skupa značajki 3](#_Toc104971782)

[3. Nenadzirano i polu-nadzirano strojno učenje 4](#_Toc104971783)

[3.1. Pregled metodologije strojnog učenja 4](#_Toc104971784)

[3.2. OPTICS algoritam i implementacija 4](#_Toc104971785)

[3.3. Analiza rezultata strojnog učenja 4](#_Toc104971786)

[4. Selekcija značajki 5](#_Toc104971787)

[4.1. Pregled metodologije selekcije značajki 5](#_Toc104971788)

[4.2. Analiza značajki pomoću PCA 5](#_Toc104971789)

[4.3. Implementacija selekcije značajki 5](#_Toc104971790)

[4.4. Indeksi validacije rezultata grupiranja (clustering-a) 5](#_Toc104971791)

[4.5. Analiza rezultata selekcije značajki 5](#_Toc104971792)

[5. Implementacija na ugradbenom računalnom sustavu 6](#_Toc104971793)

[5.1. Pregled ugradbenog računalnog sustava 6](#_Toc104971794)

[5.2. Nadogradnja algoritma klasifikacije akustičnih emisija 6](#_Toc104971795)

[5.3. Testiranje rada algoritma 6](#_Toc104971796)

[Zaključak 7](#_Toc104971797)

[Literatura 8](#_Toc104971798)

[Sažetak 9](#_Toc104971799)

[Summary 10](#_Toc104971800)

[Skraćenice 11](#_Toc104971801)

[Privitak 12](#_Toc104971802)

# Uvod

Preko 70% ukupne vode u svijetu koristi se u svrhu navodnjavanja agrikulture. Potreba za vodom raste porastom učestalosti i intenziteta suša u svijetu. Pronalaskom načina uštede vode u procesu navodnjavanja drastično bi utjecao na problem nestašice vode. Navedenom problematikom bavi se područje pametnog i preciznog navodnjavanja. Pametno navodnjavanje temelji se na proučavanju indikatora vodenog stresa biljke. Radi preciznije analize mijenja se pristup detekcije vodenog stresa od mjerenja deficita vode u tlu zbog velike ovisnosti o vrsti tla. Glavni indikatori vodenog stresa biljke je pojava kavitacija unutar ksilema biljke. Time fokus pametnog navodnjavanja prelazi na metode koje proučavaju hidrauličke, vizualne te auditivne indikatora kavitacije unutar biljke. Za detekciju kavitacije na terenu u stvarnom vremenu jedino su primjenjive akustične metode za koje nisu potrebni laboratorijski uvjeti. Akustične metode temelje se na analizi ultrazvučnih akustičnih emisija ili UAE koje biljka emitira u stanju vodenog stresa. Cilj akustičnih metoda je pronaći UAE uzrokovane kavitacijom. Unaprjeđivanje istraživanja UAE pokazuju na mogućnost prepoznavanje UAE nastala kavitacijom, međutim potrebna je dodatna analiza. Glavni temelji se rad je

Problem pametnog i preciznog navodnjavanja sve više raste na važnosti zbog povećanja nestašice vode. Istraživanja klimatskih promjena potvrđuju

kako Biljke emitiraju akustične emisije tijekom svih stadija svojega života koje daju informacijo o trenutnim fiziološkom stanju biljke.

U ovom radu analizirati će se akustične emisije biljaka emitirane prilikom dehidratacije u svrsi razvoja sustava pametnog navodnjavanja.

Glavni uzrok akustičnih emisija prilikom dehidratacije biljaka je proces embolizacije i kavitacije čijim djelovanjem nastaju akustične emisije. Prepoznavanjem akustičnih emisija uzrokovani kavitacijom omogućuje detekciju kada se biljka nalazi u procesu dehidratacije.

**Opiši ostale moguće uzroke u uvodu ili kasnije?**

Ovaj rad bavi se razvojem postupka klasifikacije uzroka akustičnih emisija kod biljaka prilikom dehidratacije pomoću metodologije strojnog učenja. Rad se nadovezuje na članak [2] u kojem se razvio senzorski čvor za akviziciju i jednostavnu klasifikaciju akustičnih emisija biljaka na terenu.

**Opis što će se odraditi u pojedinom poglavlju?**

**Mentor:**

1. **Motivacija – koji problem rješava i koji su izazovi za koje dajem rješenja**
2. **Jasno definiranje ciljeva – 2-3 rečenice**
3. **Moji doprinosi/pristup – 2-3 rečenice**
4. **Struktura paragrafa – što koje poglavlje opisuje**

# Pregled teme i prijašnjih doprinosa detekcije kavitacije

## Opis kavitacije kod biljaka

### Motivacija

Kod pametnog navodnjavanja biljaka potrebno je prepoznati kada je biljka u stanju vodenog stresa. Jedan od postupaka je određivanje vodnog potencijala tla iz kojeg biljka upija vodu. Zbog nepouzdanost Navedeni način je nepouzdan jer izrazito ovisi o okolinskim faktorima. Radi uštede vode i optimalnog uzgoja biljaka pametno navodnjavanje se bazira mjerenjem internih proces biljke.

### Prirodni rad biljke

Efikasna transpiracija vode kroz ksilem biljke glavni je faktor prirodnog rada i razvoja biljke. Proces transpiracije uključuje transport vode kroz biljku te evaporaciju vode na površinskim organima biljke kao što su listovi i stabljika. Kroz transport vode biljka prenosi hranjive tvari do svih svojih organa. Transpiracijsko-adhezijsko-kohezijsko-tenzijska (TAKT) teorija objašnjava proces transporta vode kroz ksilem. Transport vode kroz biljku postiže se negativnim hidrostatskim tlakom koji povlači stupac vode ksilema prema gore. Hidrostatski tlak ili tenzija uzrokovana je ubrzanom evaporacijom vode iz lista biljke u svrsi hlađenja prilikom zagrijavanja biljke i njezine okoline. Zbog navedenih tenzija voda se nalazi u fizikalno nestabilnom stanju te kada se hidrostatski tlak vode izjednači s tlakom pare voda prelazi u plinovito agregatno stanje. Unatoč tome kohezivna i adhezivna svojstva vode i ksilema održavaju vodu u tekućem stanju kao neprekinuti stupac vode.

### Procesa i uzrok kavitacije kod biljaka

Tijekom izrazito velikih tenzija uzrokovanih evaporacijom vode moguće je formiranje i širenje mjehurića plinova unutar ksilema. Navedeni proces naziva se kavitacija ili embolija. Kavitacija je reverzibilna te anatomija ksilema je napravljena za sprječavanje širenja kavitacije tijekom tipičnih uvjeta rasta biljke. Tijekom razdoblja suše brzina evaporacije stvara mjehuriće koji potpuno sprječavaju protok vode kroz ksilem i uzrokuju ugibanje biljke. Kavitacija je prvi znak stanja vodenog stresa biljke. Prepoznavanje nastanka kavitacije služi za pravovremeno i efikasno navodnjavanje biljaka. [0]

## Trenutna metodologija detekcije kavitacije

### Pregled postojećih metoda detekcije i mjerenja kavitacije te usporedba s metodom analize akustičnih emisija

Glavne metode detektiranja kavitacije mogu se podijeliti na hidrauličke metode, vizualne metode te akustične metode. Hidrauličke metode temelje se na mjerenju postotka gubitka hidrauličke vodljivosti biljke ili PLC. Metoda se sastoji se od dva mjerenja hidrauličke vodljivosti provedenih na odrezanom uzorku biljke. Prvo mjerenje hidrauličke vodljivosti provodi se na uzorku u prirodnom stanju. Prilikom mjerenja potrebno je paziti da se ne uvedu dodatni mjehurići zraka u uzorak. Drugo mjerenje mjeri maksimalnu hidrauličku vodljivosti uzorka nakon uklanjaju mjehurići iz uzorka. PLC se računa kao omjer dvaju mjerenja. Zbog destruktivne prirode mjerenja te nedostatka automatiziranosti hidrauličke metode nisu primjenjiva za terensku detekciju kavitacija u stvarnom vremenu. Nadalje vizualne metode temelje se na vizualnoj identifikaciji mjehurića zraka unutar ksilema biljke. Kavitacija može se ne destruktivno identificirati pomoću nuklearne magnetske rezonanca (NMR) i rendgenske kompjuterizirane tomografije (µCT). Kvalitetniju analizu daje µCT zbog bolje prostorne razlučivosti, iako pri mjerenju postoji mogućnost oštećenje stanica biljke. Vizualne metode su najpouzdanije za detekciju kavitacije, međutim nisu primjenjive kao terenski i rasprostranjeni alat detekcije kavitacije. Zbog svoje ne destruktivne i pouzdane prirode koriste se kao referentno mjerenje kavitacije prilikom korištenja akustičnih metoda. [1.1]

Završno akustične metode temelje se na mjerenjima akustičnih emisija (AE) unutar ksilema biljke. Početno AE bile su smatrane uzrokom puknuća kavitacijskih mjehurića zraka unutar ksilema te time kao optimalno sredstvo prepoznavanja kavitacije. [] Daljnjim istraživanjima pokazano je kako emitiranje AE je aktivno tijekom prirodnog stanja biljke te u stanju nakon totalnog gubitka hidrauličke vodljivosti biljke. Oba navedena stanja biljke isključuju kavitaciju kao jedini uzrok AE. Time su identificirani dodatni uzroci AE: sužavanje stanične stijenke, formiranje pukotina, kapilarno djelovanje vlakna biljke (Rosner 2012.) i skupljanje tkiva kore. Navedeni uzroci AE proizvode mehaničke valove propagirane kroz tkivo biljke koje je moguće mjeriti piezo električnim senzorima postavljenim na površinu stabljike.

Za detekciju kavitacije na terenu u stvarnom vremenu jedino su primjenjive akustične metode za koje nisu potrebni laboratorijski uvjeti.

Ne destruktivnosti i prilagodljivost akustične metode predstavlja potencijalnu terensku metodu za ranu detekciju kavitacije kod biljaka.

Zaključci istraživanja:

Danjim istraživanjima identificiran je da je frekvencijski pojas pojavljivanja AE između 100 kHz i 1000 kHz.

Pokazana je korelacija između nastanka AE emisija i smanjenja hidrauličke vodljivosti biljke te vodenog potencijala biljke (Mapiranje AE na VC). Isto tako određene AE su uspoređene s direktnom vizualizacijom embolizma metodom µ-CT.

[0]

Indirektno mjerenje hidrauličke vodljivosti postiže se mjerenjem vodnog potencijala ksilema pomoću krivulja ranjivosti (VC) definiranih na temelju sorte ili vrste biljaka nad kojom se vrše mjerenja.

Alati za mjerenja vodnog potencijala:

1. Scholander tlačna komora
2. Psihorometri (mjere vlažnost zraka) za tlak pare stabljike – ne invazivno i ovisno o temperaturi okoline

Navedene metode nisu primjenjive za proizvodnju velikih razmjera na terenu.

In-situ -> mjerenja nad organizmom u prirodnoj okolini i stanju, u divljini

### Pregled napretka i eksperimenta u području klasifikacije akustičnih emisija prilikom dehidratacije biljke. Uključuje i kombinaciju ostalih metoda uz analizu AE.

[1]

### Detaljan opis rada [2] na koji se nadovezuje ovaj diplomski tad te metodologije kojom će se nadograditi. Kakvi senzori trebaju?

Korišteni skupovi podataka dobiveni su kontinuiranom akvizicijom UAE u eksperimentima sušenja vinove loze. Radi ubrzanja i kontrole procesa sušenja u većini eksperimenta korištena je klima komora, dok u eksperimentu 6 sušenje je odrađeno prirodnim putem. Proces akvizicije UAE sastojao se od akvizicije UAE piezo električnim senzorom Vallen VS600-Z1, obrade analognog signala Vallen pojačalom AEP3N, digitalizacije signala analogno digitalnim pretvornikom National Instruments USB-DAQ 6366 ADC te pohrana rezultata pomoću Matlab skripte. ADC je digitalizirao UAE s frekvencijom uzorkovanja 2 MHz te 16-bitnom rezolucijom. Odrađena je analiza dobivenih UAE na temelju ekstrakcije značajki vremena pojave pojedine UAE te frekvencije rezonantnih vrhova pojedine UAE. Unutar svih skupova podataka identificirano je slično grupiranje frekvencije rezonantnih vrhova pojedinih UAE u određenim frekvencijskim pojasevima. Izdvojene su 3 karakteristične grupe UAE pomoću značajke broja rezonantnih vrhova pojedinih UAE. U prosjeku oko 25% UAE spada u grupu bez rezonantnih vrhova ili široko pojasne UAE. Navedene UAE grupiraju se u frekvencijskom pojasu između 100 i 200 kHz. U prosjeku oko 15% UAE spada u grupu UAE s jednim rezonantnim vrhom. Navedene UAE grupiraju su frekvencijskom pojasu oko 200 kHz. U prosjeku oko 60% UAE spada u grupu UAE s više rezonantnih vrhova. Za navedene UAE vrh s najvećom magnitudom tipično se grupira oko 200 kHz, dok ostali vrhovi između 200 i 600 kHz. [4]

Ovaj radi želio bi raditi dodatnu analizu navedenih saznanja pomoću metoda strojnog učenja te metoda selekcije značajki u svrsi pronalaska grupacija na temelju kojih je moguće kalsificirati UAE po izvoru nastajanja.

U tablici 1 prikazana su vremena trajanja, korišteni kultivar vinove loze te opis procesa sušenja za svaki eksperiment korištenih skupova podataka.

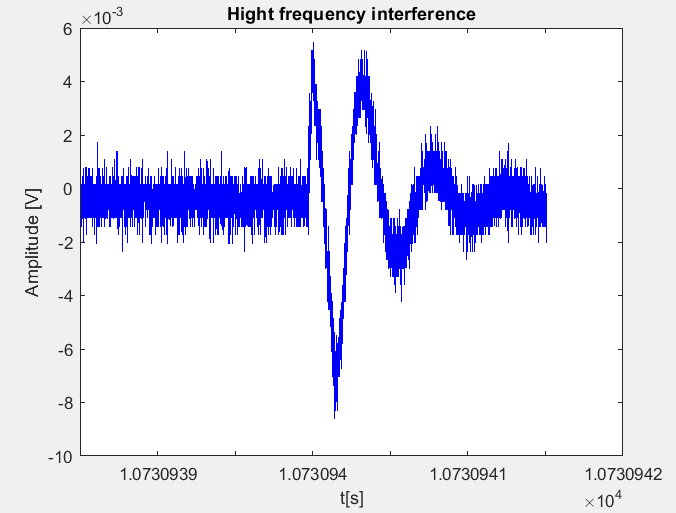
# Ekstrakcija značajki UAE

Ekstrakcija značajki (*engl. feature extraction*) je proces obrade sirovih podataka i pretvorbe u numeričke značajke uz očuvanje informacija o izvornom skupu podataka. Numeričke značajke se dovode na ulaz algoritma strojnog učenja koji klasificira skup podataka na temelju značajki. Zbog toga kvaliteta klasifikacije podataka iznimno ovisi o kvaliteti ekstrakcije značajki. Kvalitetna ekstrakcija značajki podrazumijeva izbor značajki relevantnih za domenu i problem koji se namjerava riješiti. [15] Skupovi podataka nad kojima se radi ekstrakcija značajki dobiveni su eksperimentima opisanim u [4]. Izabrano je 9 reprezentativnih skupova podataka prikazani u tablici 1. Ekstrakcija značajki odrađeno je jednako na svim skupovima podataka te podrazumijeva nadogradnju na postojeću predobradu iz [4]. Glavni koraci ekstrakcije značajki su izdvajanje validnih emisija iz skupa podataka, ekvilizacija emisija te izračun vrijednost značajki za svaku emisiju. Navedeni koraci prikazani su na slici 1. U slijedećim poglavljima opisuje se implementacija ekstrakcije značajki u Matlabu koja će kasnije biti u grubo preslikana na ugradbeni računalni sustav.

## Izdvajanje validnih emisija iz skupova podataka

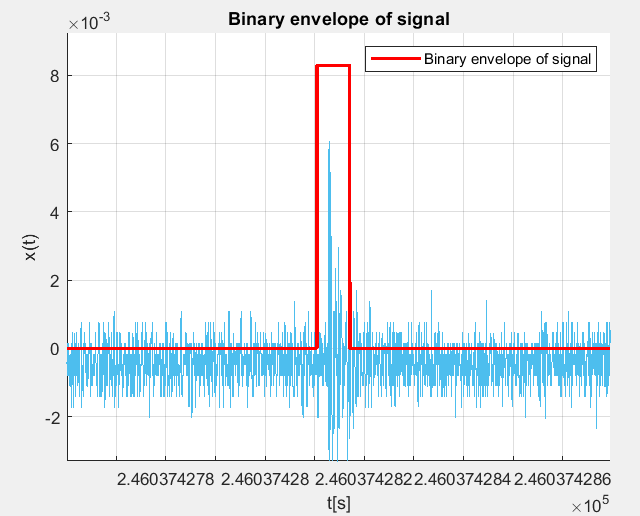
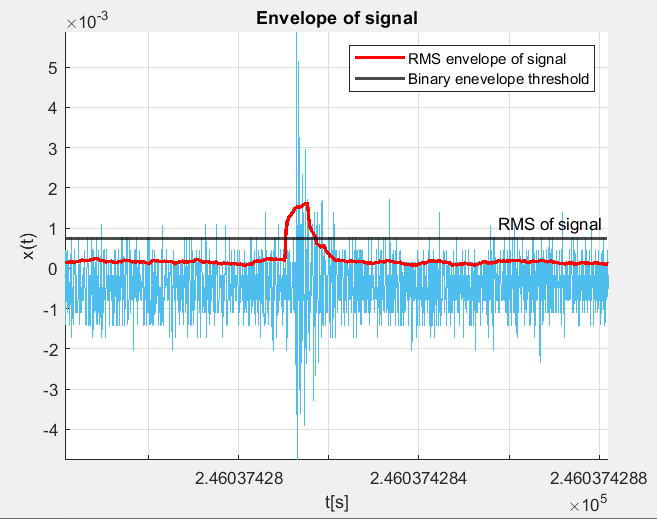
Izdvajanje potencijalnih UAE iz ukupnog skupa podataka na ugradbenom računalnom sustavu odrađuje akvizicijski sustav s postavljenim pragom od 2 mV za početak pohrane pojedine UAE. Međutim navedeni skupovi podataka dobiveni su kontinuiranom pohranom UAE te emisije su izdvojene po intervalima od 3 ms. Zbog toga prvi korak predobrade je odvojiti validne emisije od šuma za svaki vremenski prozoru od 3 ms.

Validacija emisija započinje izbacivanjem nisko frekvencijskih smetnji iz daljnje analize. Nisko frekvencijske smetnje karakterizirane su s maksimalnom magnitudom amplitudnog spektru ispod 100 kHz te u vremenskoj domeni vidimo vidljiva izobličenja spektra prikazana na slici 1.



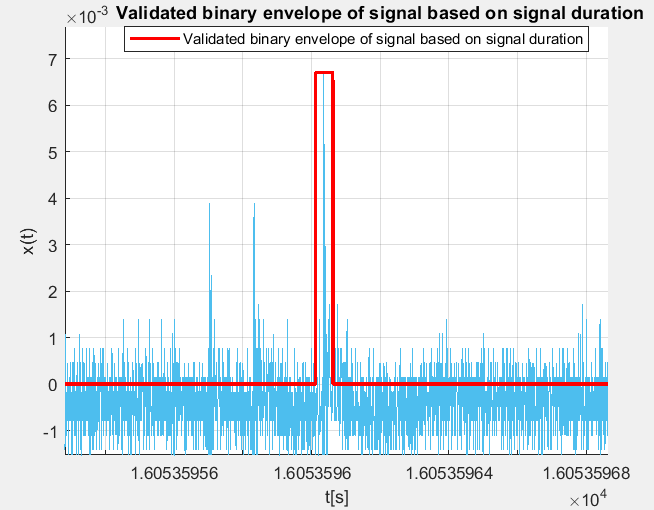
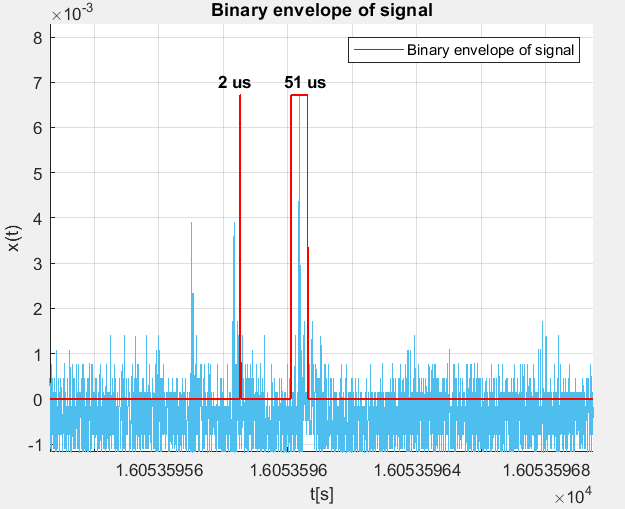
Slika 1 Emisija uzrokovana nisko frekvencijskom smetnjom

Slijedeće odvajaju se potencijalne emisije od šuma računanjem RMS omotnice signala s prozorom od 100 točaka tj. 50 us. Izračun RMS omotnice daje najveći uvid u amplitudu signala jer daje mjeru snage signala, dok stvara valni oblik koji je lako analizirati. [11] Nad dobivenim valnim oblikom, crvena krivulja na slici 2 lijevo, izdvajamo emisiju na temelju amplitudnog praga određen RMS-om čitave emisije, crni pravac na slici 2 lijevo. Time dobivamo binarnu omotnicu emisije prikazanu na slici 2 desno koja ima vrijednosti jedan u amplitudama potencijalne emisija te vrijednosti nula u amplitudama šuma.



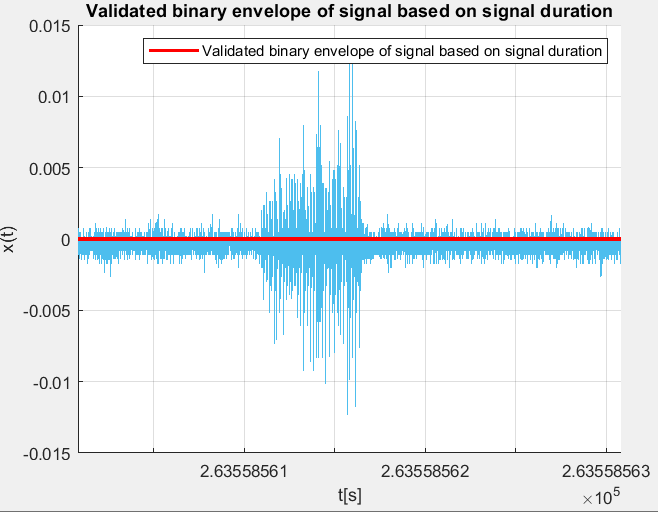
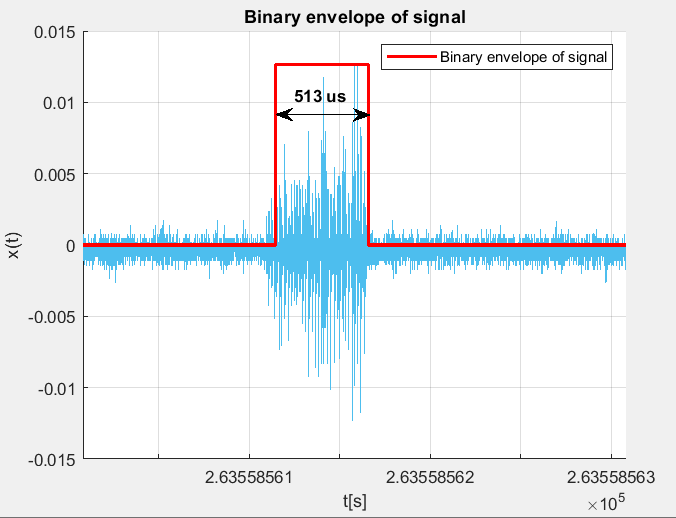
Slika 2 Prikaz izračuna binarne i RMS omotnice signala

Nadalje radi se vremenska provjera ispravnosti emisija izdvojenih binarnom omotnicom. Emisije su identificirane kao prekratke ako imaju trajanje ispod 15 µs. Na slici 3 lijevo vidimo primjer izdvajanja prekratke emisije s trajanjem o 2 µs. Emisije su identificirane kao preduge ako imaju trajanje iznad 500 µs. Na slici 3 lijevo vidimo primjer izdvajanja preduge emisije s trajanjem od 513 µs. Preduge i prekratke emisije smatraju se uzrokovane smetnjama prilikom akvizicije. Navedene emisije se zanemaruju te formiraju se nove binarne omotnice prikazana desno na slikama 3 i 4. koje uključuje samo validne emisije.

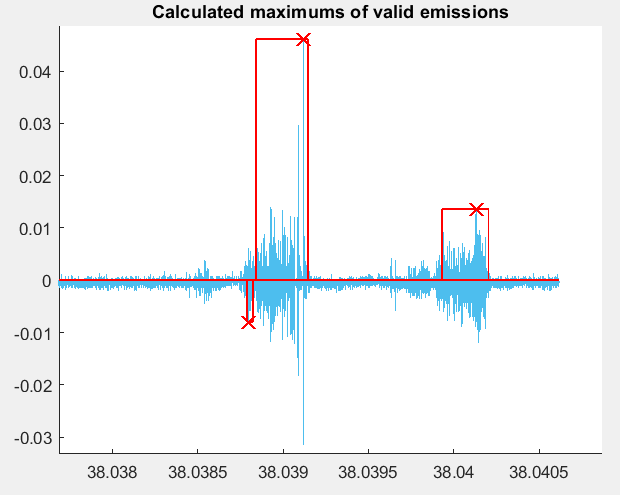


Slika 3 Prikaz izdvajanja prekratke UE

Slika 4 Prikaz izdvajanja preduge UAE

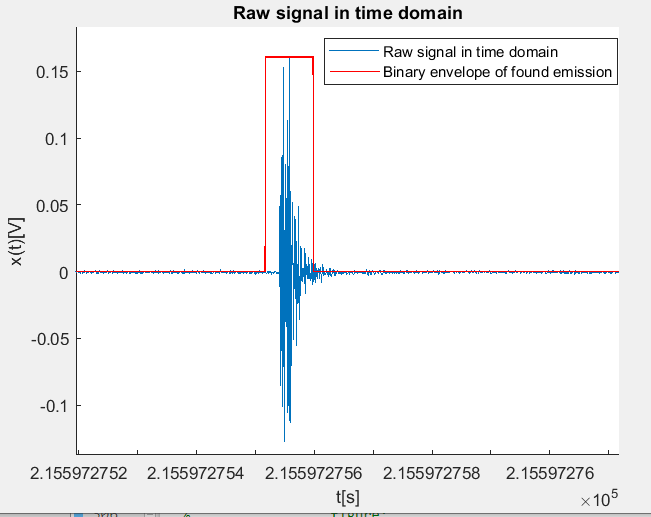


Amplitudno izbacivanje emisije radi se na temelju akvizicijskog praga od 2 mV. Gdje se emisije s maksimalnom amplitudom ispod 2 mV prikazano crveno na slici 5 lijevo zanemaruju u daljnjoj analizi. Na slici 5 desno vidimo binarnu omotnicu nakon izbacivanja emisije na temelju maksimalne amplitude.



Slika 5 Izbacivanje emisije s maksimalnom amplitudom ispod 2 mV

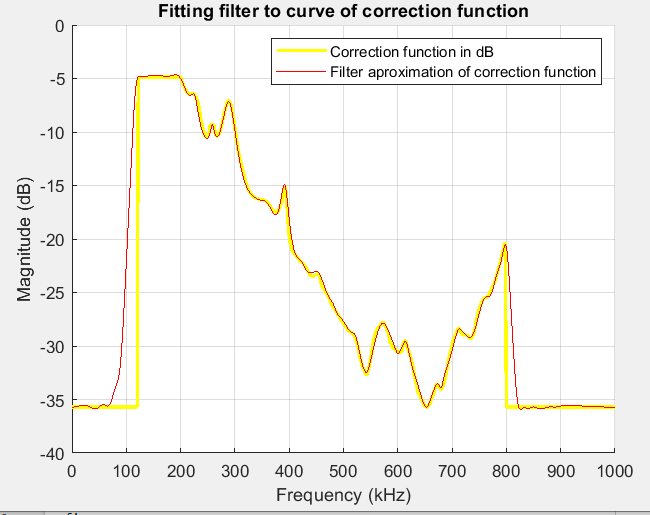
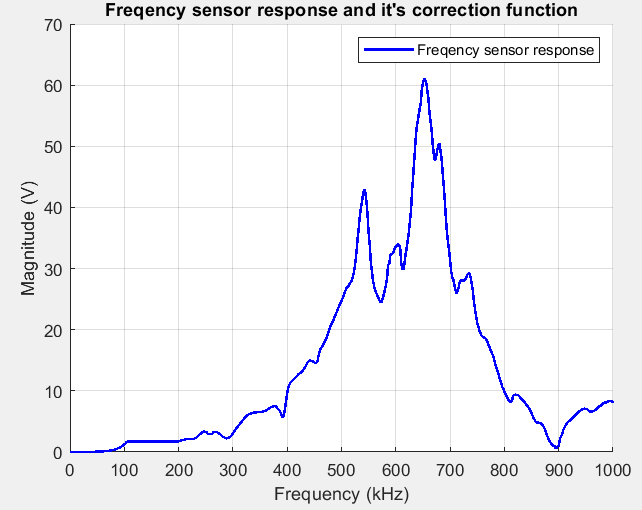
Završni emisije se izdvajaju pomoću validirane binarne omotnice te se proširuju nulama oko maksimalne amplitude emisije na fiksnu veličinu od 1024 točke ili 512 µs. Broj točaka je izabran kao potencija broja 2 kako bi računanje frekvencijskog spektra emisije pomoću brze fourierove transformacije (FFT) bilo efikasnije. [16] Isto tako kako bi uključio emisije svih trajanja ispod maksimalnog dozvoljenog od 500 µs. Prikaz tipične validne emisije izdvojene iz skupa podataka u vremenu vidi se na slici 5.



Slika 6 Tipična validna emisija u vremenskoj domeni

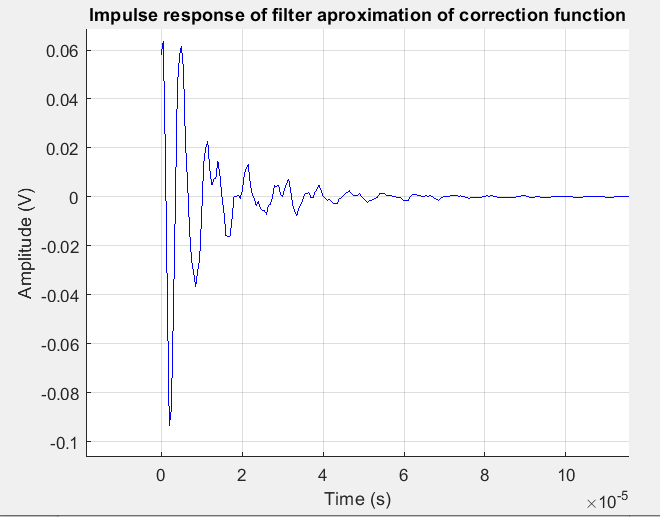
## Ekvilizacija sirovoj emisija

Jedan od najvažnijih koraka unutar navedenog postupka ekstrakcije značajki UAE je ekvilizacija sirovih ultrazvučnih emisija. Ekvilizacija predstavlja metodu kojom ispravljamo nelinearne promjene koji unosi senzor i obrada analognog signala prilikom akvizicije UAE. Za senzor korišteni pri akviziciji podataka definirani je amplitudni spektar frekvencijskog odziva senzora (slika 6 lijevo). Ekvilizacija amplitudnog spektra UAE postiže se množenjem s korekcijskom funkcijom definiranom recipročnim magnitudama frekvencijskog odziva senzora. Kako bi ekvilizirali UAE u vremenskoj domeni potrebno je formirati fazni spektar frekvencijskog odziva senzora u Matlabu. Pomoću alata za procesiranje signala u Matlabu [17] kreiranjem filtra određenog amplitudnog spektra u decibelima automatski se kreira prikladni fazni spektar. Na slici 6 desno prikazan je amplitudni spektar filtra u decibelima (crvena krivulja) koji aproksimira magnitude korekcijske funkcije (žuta krivulja).



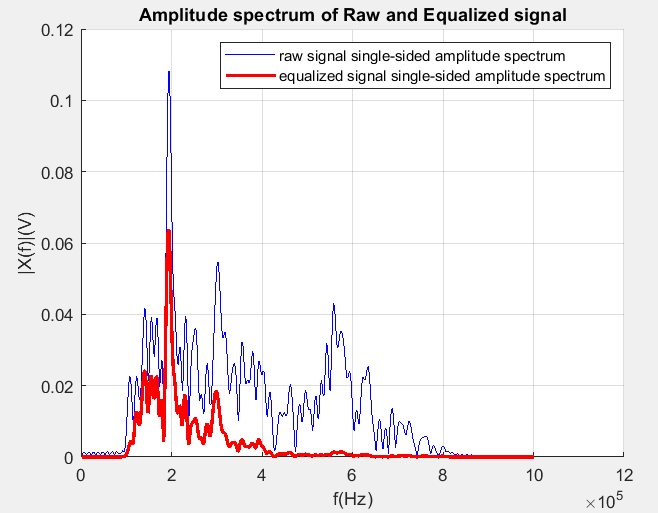
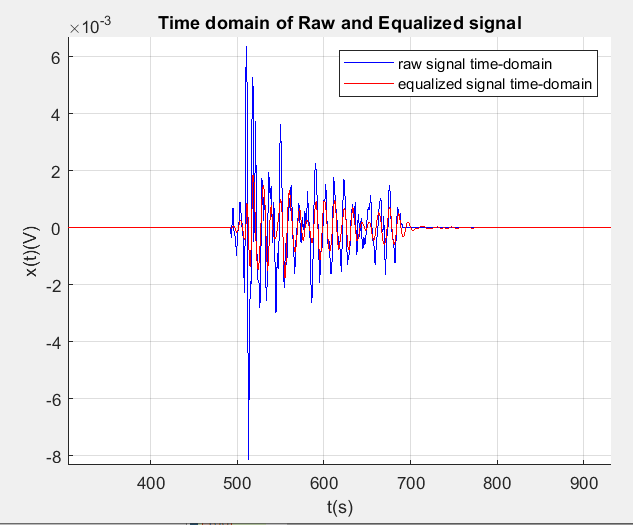
Slika 7 Amplitudni spektar frekvencijskog odziva senzora (desno) i njegova korekcijska funkcija (lijevo žuto) aproksimirana filtrom u Matlabu (lijevo crveno))

Iz navedenog filtra računa se impulsni odziv koji se koristi pri ekvilizaciji. Po teoremu konvolucije postupak konvolucije izvornog signala u vremenskoj domeni s impulsnim odzivom filtra ekvivalentan je množenju izvornog signala u frekvencijskoj domeni s magnitudama filtra u frekvencijskoj domeni. [18] Zbog toga postupak ekvilizacije obuhvaća konvoluciju amplituda UAE u vremenskoj domeni s impulsnim odzivom filtra koji aproksimira korekcijsku funkciju u decibelima prikazan na slici 7.



Slika 8 Impulsni odziv filtra koji aproksimira korekcijsku funkciju frekvencijskog odziva senzora

Ekvilizacija primjenjuje se na svakoj dobivenoj izdvojenoj validnoj emisiji. Odnos tipične sirove UAE (plava krivulja) i ekvilizirane UAE (crvena krivulja) prikazan je u vremenskoj domeni (lijevo) i amplitudnom spektru (desno) na slici 8.



Slika 9 Sirova (plavo) i ekvilizirana (crveno) emisija UAE u vremenskom (lijevo) i amplitudnom spektru (desno)

## Izbor i definiranje značajki akustičnih emisija

Izbor reprezentativnih značajki UAE je iznimno bitan kako se čitavi postupak klasifikacije uzroka UAE temelji na grupiranju vrijednosti značajki. Zbog nedostatka informacija o značajkama prikladnim za distinkciju izvora UAE implementirani je izračun većine značajki korištenih u prijašnjim eksperimentima. Primarni skup značajki preuzet je iz literature [9] koja za računanje značajki koristi program AEwin [10]. Dodatno preuzeta je značajka WPC kao pokazano bitna značajka kod razlikovanja uzroka UAE. [5] Završno izabrana je značajka broja rezonantnih vrhova u amplitudnom spektru UAE koja potencijalno izdvaja UAE izvore opisana u poglavlju 1.4. [2] Navedene značajke uz prikladne mjerne jedinice i kratice prikazane su na slici 10. Računanje određenih značajki je modificirano po potrebi

Iz vremenske domene izabrane su značajke definirane na temelju amplituda UAE te trajanja karakterističnih dijelova UAE.

* **Maksimalna amplituda** (*engl. peak amplitude*)
  + Maksimalna apsolutna amplituda UAE izražena u voltima.
* **Vrijeme uzdizanja** (*engl. rise time*)
  + Vrijeme od početka UAE *(engl. start time*) do pojave maksimalne amplitude UAE u vremenskoj domeni izraženo u sekundama.
* **Vrijeme pada** *(engl. fall time*)
  + Vrijeme od pojave maksimalne amplitude UAE do vremena kraja UAE *(engl. end time*) u vremenskoj domeni izraženo u sekundama.
* **Trajanje** (*engl. duration*)
  + Vrijeme trajanja UAE izraženo u sekundama.
* **Ukupni broj vrhova** (*engl. total counts*)
  + Ukupni broj amplituda UAE koje prelazi predodređeni amplitudni prag.
* **Broj vrhova prije maksimalne amplitude** (*engl. counts to peak*)
  + Broj vrhova UAE u vremenu prije maksimalne amplitude UAE uključujući maksimalnu amplitudu.
* **Broj vrhova poslije maksimalne amplitude** (*engl counts from peak*)
  + Broj vrhova UAE u vremenu poslije maksimalne amplitude UAE.
* **Srednja frekvencija** (*engl. average frequency*)
  + Frekvencija pojavljivanja vrhova u ukupnom UAE. Računa se po formuli (1) te izražena je u herzima.

|  |  |  |
| --- | --- | --- |
|  |  | (1) |
|  |  |  |

* **Inicijacijska frekvencija** (*engl. initiaton frequency*)
  + Frekvencija pojavljivanja vrhova u vremenu uzdizanja UAE. Računa se po formuli (2) te izražena je u herzima.

|  |  |  |
| --- | --- | --- |
|  |  | (2) |
|  |  |  |

* **Frekvencija odjeka** (*engl. reverberation frequency*)
  + Frekvencija pojavljivanja vrhova u vremenu pada UAE. Računa se po formuli (3) te izražena je u herzima.

|  |  |  |
| --- | --- | --- |
|  |  | (3) |

* **Snaga signala** (*engl. signal strength*)
  + Snaga apsolutnih amplituda UAE. Računa se po formuli (4) te izražena je u voltima po sekundi.

|  |  |  |
| --- | --- | --- |
|  |  | (4) |

* **Apsolutna energija** (*engl. absolute energy*)
  + Prava energija UAE izražena u džulima. Računa se po formuli (5) gdje 10 kΩ predstavlja referentni otpor tijekom trajanja UAE.

|  |  |  |
| --- | --- | --- |
|  |  | (5) |

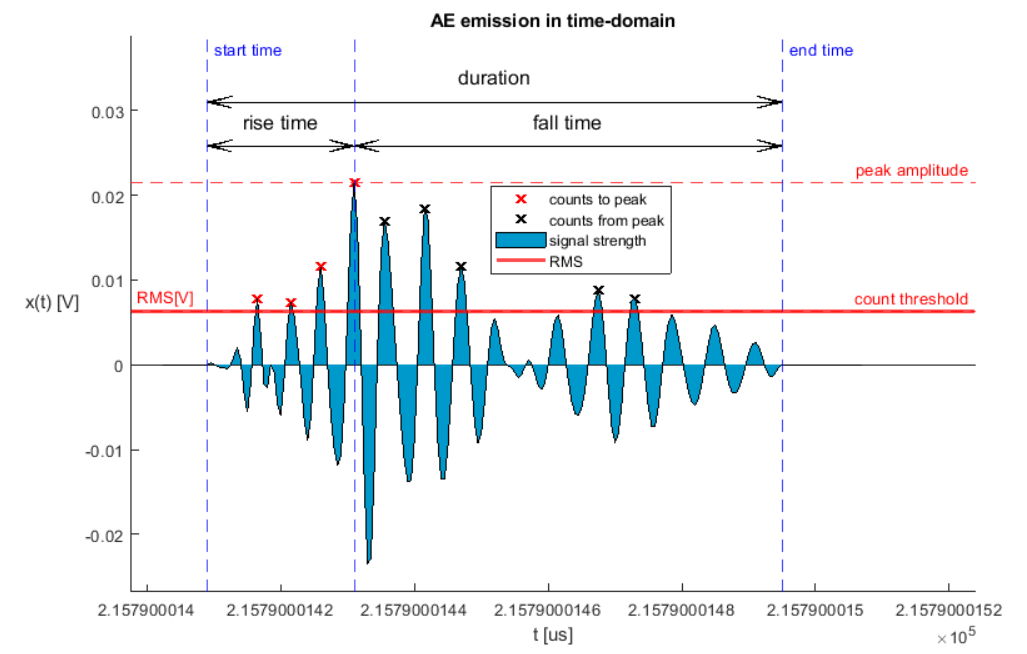
* **RMS**
  + Kontinuirane promjene amplituda UAE. Računa se po formuli (6) te izražena je u voltima.

|  |  |  |
| --- | --- | --- |
|  |  | (6) |

* **Srednja razina signala** (*engl. average signal level*) ili **ASL**
  + Kontinuirane promjene srednje vrijednosti amplituda UAE. Računa se po formuli (7) te izražena je u decibelima.

|  |  |  |
| --- | --- | --- |
|  |  | (7) |

Prikaz vremenskih značajka na tipičnoj UAE može vidjeti na slici 1.



Slika 10 Prikaz vremenskih značajki na tipičnoj UAE

Iz frekvencijske domene značajke se računaju na jednostranom amplitudnom spektru UAE s oznakom *UAE\_fft\_abs(f)* koji se računa pomoću brze fourierova transformacije nad UAE. Frekvencijske značajke su definirane na temelju magnituda amplitudnog spektra UAE te frekvencija karakterističnih dijelova UAE:

* **Parcijalna snaga** (*engl. partial power*) ili **PP**
  + Omjer snage signala u određenom frekvencijskom pojasu i ukupne snagu frekvencijskog spektra izražena u postotcima. Računa se po formuli (8) gdje je frekvencijski pojas određen frekvencijama *f1* i *f2*, a

|  |  |  |
| --- | --- | --- |
|  |  | (8) |

* **Maksimalna magnituda** (*engl. peak magnitude*)
  + Maksimalna apsolutna magnituda amplitudnog spektra UAE izražena u voltima.
* **Frekvencija maksimalna magnitude** (*engl. peak frequency*) ili **PF**
  + Frekvencija maksimalne apsolutne magnitude amplitudnog amplitudnog spektra UAE izražena u herzima.
* **Frekvencija centroida** (*engl. centroid frequency*) ili **CF**
  + Frekvencija centroida amplitudnog spektra UAE izražen u herzima. Računa se po formuli (9).

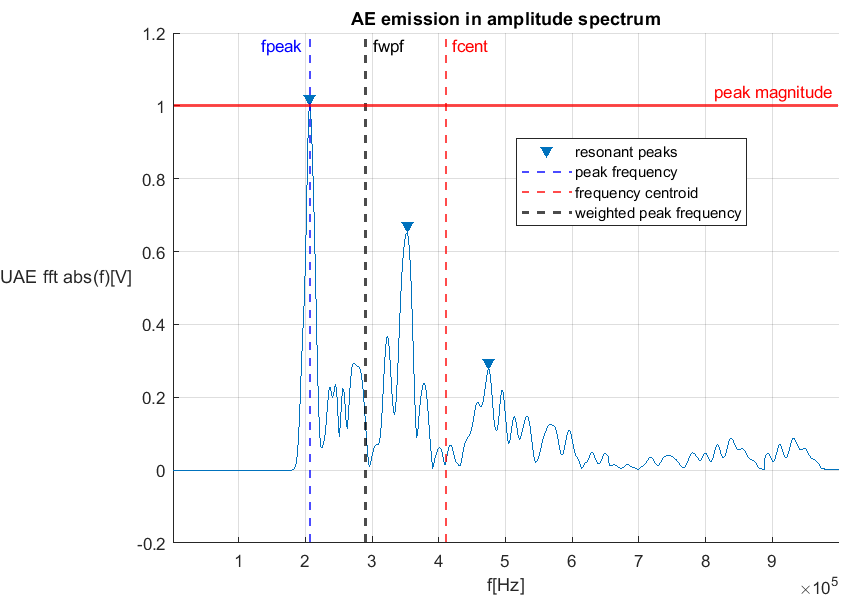
|  |  |  |
| --- | --- | --- |
|  |  | (9) |

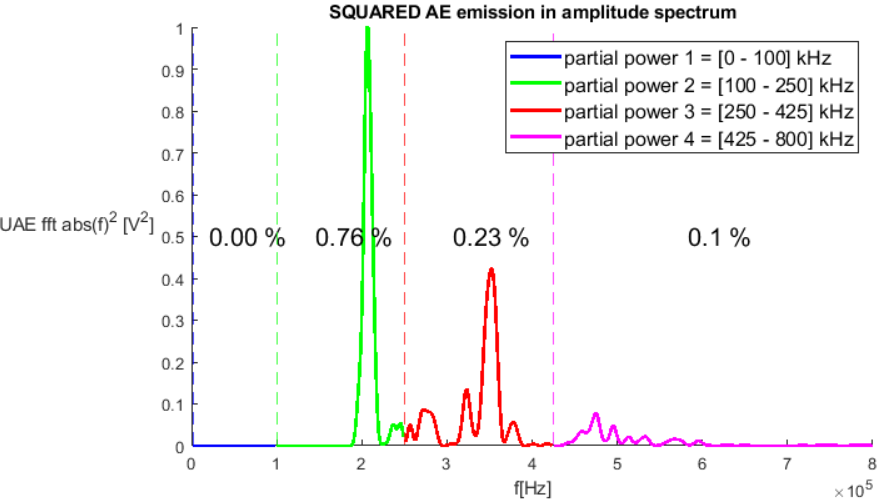
* **Težinska maksimalna magnituda** (*engl. weighted peak frequency*) ili **WPC**
  + Težinska maksimalna magnituda amplitudnog spektra UAE koristi se kombinacija informacija značajki PF i CF. Računa se po formuli (10) te izražena je u herzima.

|  |  |  |
| --- | --- | --- |
|  |  | (10) |

* **Broj rezonantnih frekvencijskih vrhova** (*engl. number of frequency peaks*)
  + Broj rezonantnih vrhova u amplitudnom spektru UAE određene širine te razlike u vrijednosti od susjednih magnituda.

Prikaz frekvencijskih značajka na tipičnoj UAE može vidjeti na slici 3 i 4.





Računanje integrala na diskretnim vrijednostima UAE računa se zbrojem površina pravokutnika između parova susjednih amplituda ili magnituda UAE ovisno o spektru. Visina pravokutnika je određen srednjom vrijednosti između amplituda, dok širina je određena vremenom između amplituda.

|  |  |  |
| --- | --- | --- |
| **Značajka UAE** | **Kratica** | **Mjerna jedinica** |
| Rise time | RISE TIME | s |
| Counts to peak amplitude | CNT TO | # |
| Counts from peak amplitude | CNT FROM | # |
| Duration | DUR | s |
| Peak amplitude | PEAK AMP | V |
| Average frequency | AVG FREQ | Hz |
| Root mean square | RMS | V |
| Average signal level | ASL | dB |
| Initiation frequency | INIT FREQ | Hz |
| Reverbation frequency | REVB FREQ | Hz |
| Signal strength | SIG STR | Vs |
| Absolute energy | AENG | J |
| Partial power 1 | PP1 | % |
| Partial power 2 | PP2 | % |
| Partial power 3 | PP3 | % |
| Partial power 4 | PP4 | % |
| Frequency centroid | FC | Hz |
| Peak frequency | PF | Hz |
| Weighted peak frequency | WPF | Hz |
| Peak magnitude | PEAK MAG | V |
| Number of resonant frequency peaks | RES PEAK CNT | # |
| Total counts | TOTAL CNT | # |
| Fall time | FALL TIME | s |

## Analiza distribucija skupa značajki

Prikazuje i analizira rezultate ekstrakcije značajki kao Matlab figure te opisuje razlog i proces normalizacije značajki.

## Normalizacija skupa značajki

Prikaz distribucije normaliziranih vrijednosti reprezentativnih skupova podataka.

# Nenadzirano i polu-nadzirano strojno učenje

## Pregled metodologije strojnog učenja

Pogledaj str 7 [download (psu.edu)](http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.295.8115&rep=rep1&type=pdf)

Nenadzirano strojno učenje koristi metode strojnog učenja za analizu ili grupiranje neoznačenog skupa podataka. Nenadzirano učenje po vrsti zadatka može se podijeliti na:

1. **Grupiranje** (*engl.* clustering) – metoda grupiranja skupa podatka tako da podaci unutar iste grupe su sličniji jedni drugima, nego podacima iz drugih grupa;
2. **Procjena gustoće** (engl. *density estimation*)– metoda nalaženja gustoće vjerojatnosti koja će opisivati distribuciju skupa podataka;
3. **Otkrivanje novih ili stršećih vrijednosti** (engl. *novelty/outlier detection*) – metoda otkrivanja podataka koji po vrijednosti značajki se bitno razlikuju od većine ostalih podataka u skupu podataka;
4. **Smanjenje dimenzionalnosti** (engl*. dimensionality reduction*) – metoda preslikavanja početnog skupa značajki u manji skup zadržavajući temeljna svojstva izvornog skupa značajki. [3.1]

U ovom radu koristiti će se smanjenje dimenzionalnosti radi početne analize skupa značajki na temelju koje će se raditi grupiranje. Izabrana metoda smanjenja dimenzionalnosti je analiza principalnih komponenti ili PCA na temelju literature [3].

…

Nakon početne analize skupa značajki PCA potrebno je izabrati metodu grupiranja prikladnu zadanom skup podataka te vrsti podataka. Podaci ne temelju značajki skupa podatka pozicioniraju se u p-dimenzionalni prostor unutar kojeg se odrađuje grupiranje na temelju izabrane metrike sličnosti.

Metode grupiranja možemo podijeliti na temelju zadanih svojstva grupa podataka:

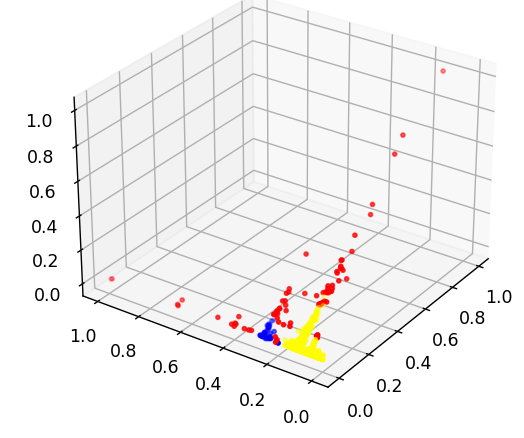
1. **Interna struktura grupe podataka** – postojanja interne strukture grupa podataka gdje grupe sadrže podgrupe. Kod particijskog grupiranja svaka grupa podataka je zasebna, dok kod hijerahijskog grupiranje svaka grupa može sadržavati podgrupe;
2. **Pripadnost podataka grupama** - kod čvrstog grupiranja svaki podataka pripada samo jednoj grupi, dok kod mekog grupiranja podaci mogu pripadati više grupa gdje pripadnost određenoj grupi je određena vjerojatnošću;
3. **Sličnosti podataka unutar grupa:**
   1. **Kvalitetno odvojene grupe** – svaki podatak u grupi je sličniji svakom drugom podatku unutar te grupe, nego podatcima izvan te grupe;
   2. **Bazirane na središtu grupe** – svaki podatak u grupi je sličniji središtu te grupe, nego središtima drugih grupa;
   3. **Bazirane na susjednim podacima unutar grupe** – svaki podatak u grupi je sličniji jednom ili više podataka unutar te grupe, nego podatcima izvan te grupe;
   4. **Bazirane na gustoću podataka** – raspoređenost podataka unutar grupe je veće gustoće, nego raspoređenost podataka između grupa. Primjenjuje se kod grupa arbitrarnog oblika i veličina te kod prisutnosti šuma ili stršećih vrijednosti;
   5. **Bazirane na konceptualnom svojstvu grupe** – grupe dijele zajedničko konceptualno svojstvo;
   6. **Bazirane na objektivnoj funkciji grupe** – grupe maksimiziraju ili minimiziraju određenu objektivnu funkciju. [3.2]

Svrha grupiranja je nalaženje prirodnih grupa u skupu podataka na temelju. Za navedeni skup podataka grupe će odgovarati izvorima AE emisija kod vodenog stresa biljaka. Zbog toga potrebno je grupirati u podatke u zasebne grupe gdje svaki podataka pripada jednoj grupi te grupe mogu biti arbitrarnog oblika, veličine i gustoće. Isto tako potrebno je prepoznati i odvojiti emisije od šuma. Iz navedenih zahtjeva vidi se kako je potrebno koristiti meke particijske metode grupiranja bazirane na gustoći podataka (DBCLA). DBCLA ovisi o određivanju gustoće određenih područja skupa podataka. Cilj DBCLA je pronaći grupe na različitim razinama gustoće uz odgovarajuće filtriranje šuma.

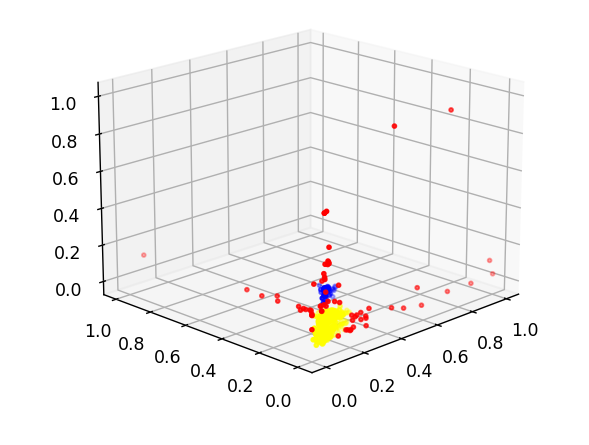
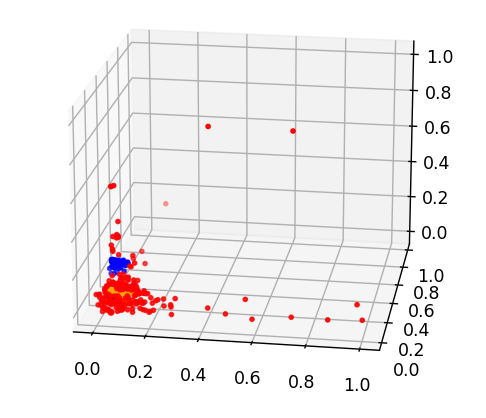
Pregled metoda grupiranja baziranih na gustoći podataka: [3.3]

HDBSCAN služi za smanjenje broja ulaznih parametara u DBSCAN i mogućnost clusteriranja clustera različite gustoće. Iako za većinu slučajeva HDBSCAN relativno dobro clusterira, za određene slučajeve je potrebno manualno podesiti min\_points i epsilon parametre što ga čini ekvivalentom OPTICS-algoritmu što se tiče korištenja u svrsi automatskog clusteriranja.

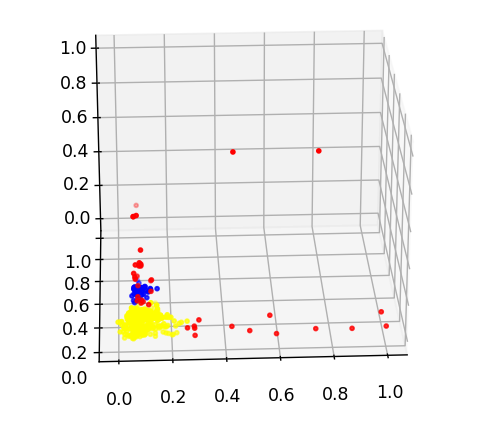
Na donjoj slici HDBSCAN je kvalitetno clusteriro bez podešavanja epsilon parametra.



Na slici dolje lijevo prikazano je clusteriranje bez podešenog epsilona već je HDBSCAN sam izračunao optimalni, dok na slici dolje desno prikazano je clusteriranje s podešenim epsilonom. Vidi se kako HDBSCAN nije prikladno odredio epsilon za prikazani podskup značajki. Zaključujemo da HDBSCAN nije prikladan za naš skupo podataka kako unosi dodatnu složenost, a daje približno jednake rezultate kao OPTICS algoritam. Isto tako mora primiti jednake parametre na ulazu kako bih došao do optimalnog rezultata što ga čini jednakim OPTICS-u prilikom selekcije značajki.



Ipak daje dobre rezultate ako mijenjamo odnos min\_points i min\_cluster\_size.



Opisuje općeniti proces ne-nadziranog strojnog učenja te uspoređuje ga s polu-nadziranim strojnim učenjem. Testiraj implementaciju u cpp-u [rohanmohapatra/hdbscan-cpp: Fast and Efficient Implementation of HDBSCAN in C++ using STL (github.com)](https://github.com/rohanmohapatra/hdbscan-cpp)

**Polu-nadzirano strojno učenje nije primjenjivo real-time već samo za selekciju značajki, zbog toga je potrebno fokusirati se na metode polu-nadzirane selekcije značajki.**

Opisuje pregled općenitih postojećih metoda ne nadziranog strojnog učenja te koje su primjenjive za naš skup podataka.

Opisuje problematiku polu-nadziranog strojnog učenja na temelju našeg skupa podataka.

## OPTICS algoritam i implementacija

Opisuje OPTICS algoritam strojnog učenja na temelju [4] te implementaciju u Matlabu.

Nakon OPTICS izračun računa se gradient clustering na temelju OPTICS reachability plot. Prvo reachability plot sortira po dobivenom redoslijedu točaka te postavljaju se vrijednosti svih outlier točaka na 1024. Time će te točke sigurno biti točke infleksije te će odvajati zasebne clustere. Samo zbog toga unutar gradient clusteringa OPTICS otkriva sve clustere manje gustoće od epsilon i Nmin što mi se čini jednako kao DBSCAN. W parametar udaljenosti između točaka je optimalno da bude epsilon koljeno kako bi infleksija točke bila izračunato ispravno s obzirom na vrijednosti reachability distance-a. Zbog toga moguće je postaviti epsilon na Inf i možda otkriti clustere. Međutim gradient clustering onda otkriva previše manjih clustera unutar velikih koje spaja u jedan cluster. Isto tako teško prepoznaje kraj clustera ako je povećanje reachability distance-a podjednako i naglo padne, već to gleda samo kao početak novog clustera.

## Analiza rezultata strojnog učenja

Analizira rezultate strojnog učenja za optimalni podskup značajki na temelju znanja iz fiziologije bilja i procesa kavitacije.

# Selekcija značajki

## Pregled metodologije selekcije značajki

Definira pojam selekcije značajki unutar metodologije strojnog učenja.

Opisuje pregled općenitih postojećih metoda selekcije značajki podijeljenih na:

1. Filter - uključuje metode redukcije dimenzionalnosti kao **PCA**
2. Wrapper
3. Embedded
4. Hybrid

Spominje koja selekcija značajki će se koristiti u ovom radu i na temelju koje literature.

## Analiza značajki pomoću PCA

Opisuje analizu značajki pomoću PCA,

## Implementacija selekcije značajki

Detaljno opisuje metodu selekcije značajki koja se koristi u ovom radu na temelju [3].

## Indeksi validacije rezultata grupiranja (clustering-a)

Kratko opisuje općenite indekse i opće density-based indeksa.

Analizira detaljno izabrani indeks DBCV. [4]

Pokazuje rezultate zašto je dobar DBCV za naš skup podataka.

## Analiza rezultata selekcije značajki

Analiza rezultate selekcije značajki na temelju kojih bira podskup značajki koji dobro grupira podatke u distinktne skupine.

# Implementacija na ugradbenom računalnom sustavu

## Pregled ugradbenog računalnog sustava

Opisuje koji se URS koristi za implementaciju iz [2] .

Opisuje postojeći algoritam na URS-u te potrebna nadogradnju.

## Nadogradnja algoritma klasifikacije akustičnih emisija

Opisuje kako je implementirana nadogradnja i koji su bili problemi prilikom implementacije.

## Testiranje rada algoritma

Provjera točnost algoritma na snimljenom skupu podataka i trajanje izvođenja na odabranom ugradbenom sustavu.

Procjena doprinosa algoritma ukupnoj potrošnji ugradbenog sustava.

# Zaključak

Na kraju rada piše se kratak zaključak, duljine do najviše jedne stranice.

# Literatura

[0] B. Lazarević i M. Poljak, Fiziologija bilja. Zagreb, Sveučilište u Zagrebu Agronomski fakultet, 2019., str. 28 [Online] Dostupno na: <https://urn.nsk.hr/urn:nbn:hr:204:366622>

[1] L. Vergeynst, “Investigation and application of the acoustic emission technique to measure drought-induced cavitation in woody plants,” Ghent University. Faculty of Bioscience Engineering, Ghent, Belgium, 2015. [**http://hdl.handle.net/1854/LU-6925217**](http://hdl.handle.net/1854/LU-6925217)

[2] D. Oletić , S. Rosner , M. Zovko, V. Bilas, 2020., Time-frequency features of grapevine’s xylem acoustic emissions for detection of drought stress. Computers and Electronics in Agriculture. Volume 178, Studeni 2020, 105797

[3] Lidewei L. Vergeynst, Markus G.R. Sause, Niels J.F. De Baerdemaeker, Linus De Roo, Kathy Steppe, Clustering reveals cavitation-related acoustic emission signals from dehydrating branches, *Tree Physiology*, Volume 36, Issue 6, June 2016, Pages 786–796, <https://doi.org/10.1093/treephys/tpw023>

[4] [DBCV.pdf (lmu.de)](https://www.dbs.ifi.lmu.de/~zimek/publications/SDM2014/DBCV.pdf)

[5] Ankerst, M., Breunig, M. M., Kriegel, H.-P., & Sander, J. (1999). OPTICS. ACM SIGMOD Record, 28(2), 49–60. doi:10.1145/304181.304187

(Rosner 2012) Rosner, Sabine. (2012). Acoustic Emission Related to Drought Stress Response of Four Deciduous Broad-Leaved Woody Species. Journal of Acoustic Emission. 30. 11-20.

[2.1] – De Baerdemaeker, N.J., Stock, M., Van den Bulcke, J., De Baets, B., Van Hoorebeke, L., Steppe, K., 2019. X-ray microtomography and linear discriminant analysis enable detection of embolism-related acoustic emissions.

[2.2] P.A. Corporation, AEwin PCI-2 Based AE System User’s Manual. Mistras Group Inc., REV 3 (April), (2007) 1–312.

[2.3] [Amplitude Analysis: Root-mean-square EMG Envelope - Delsys](https://delsys.com/amplitude-analysis-root-mean-square-emg-envelope/)

[3.1] [SU-2020-19-Grupiranje[1].pdf (unizg.hr)](https://www.fer.unizg.hr/_download/repository/SU-2020-19-Grupiranje%5b1%5d.pdf)

[3.2] [Microsoft PowerPoint - ClusteringAnalysis.pptx (wmich.edu)](https://cs.wmich.edu/alfuqaha/summer14/cs6530/lectures/ClusteringAnalysis.pdf)

[3.3] [Sci-Hub | A survey of density based clustering algorithms. Frontiers of Computer Science, 15(1) | 10.1007/s11704-019-9059-3](https://sci-hub.se/10.1007/s11704-019-9059-3)

# Sažetak

# Summary

# Skraćenice

# Privitak

Privitak je također opcionalno poglavlje (u dogovoru s mentorom).

Sadržaj koji se stavlja u privitak je, općenito, nešto što je, kao cjelinu, prikladno izdvojiti iz sadržaja samog rada.

Mogući primjer je tehnička dokumentacija vezana uz diplomski rad - npr. električka i položajna shema sklopa, sastavnica, predložak tiskane veze, plan bušenja, ispis programa s detaljnim opisom.

Drugi primjer uključuju upute za korištenje rezultata rada (softvera ili hardvera), detaljni ispisi mjerenja čiji su rezultati sažeto ili grafički prikazani u radu. Ako se radi o softveru, uobičajeno je navesti podatke o platformi na kojoj se izvodi (npr., karakteristike uređaja i operacijskog sustava te pomoćnog softvera), kao i upute za instalaciju.

U privitku nemojte koristiti stilove razine Heading, već samo (nenumerirani) stil Podnaslov.

Na primjer:

Instalacija programske podrške

Upute za korištenje programske podrške

.

.

.

Ostali savjeti

U izborniku Tools - opcija Language postavite Croatian kao *default* jezik. Na kraju provedite strojnu provjeru teksta (*spell checking*, ako ga imate ugrađenoga), ali svakako pažljivo i pročitajte vlastiti tekst.

Ako nemate ugrađeni *spelling checker* za hrvatski jezik, možete se poslužiti Hascheckom (izgovara se Hašek, a ime dolazi od kratice za Hrvatski akademski *spelling checker*), dostupan putem poveznice <https://ispravi.me/>.