

Analiza eficienței semnalelor din domeniul timp-frecvență

1. Introducere

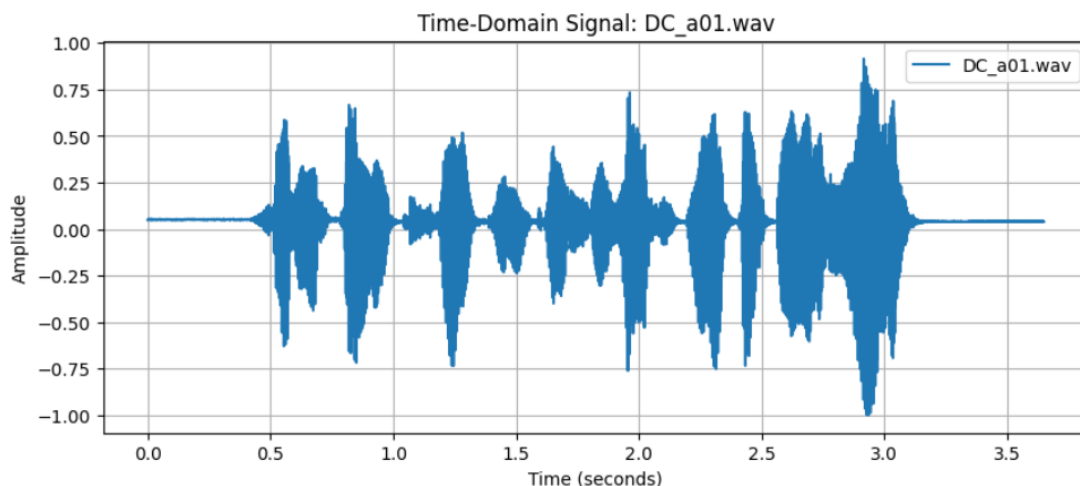
Proiectul de față pornește de la lucrarea “Time-Frequency Signal Processing” scrisă de Vladimir Klochko, Department of Automation and Information Systems in Management, Ryazan State Radio Engineering University named after V.F. Utkin, Ryazan, Russia. Lucrarea prezintă metode de procesare a semnalelor în domeniile timp-frecvență, pentru detectarea semnalelor care au frecvențe Doppler apropiate sau puncte de sosire similare. Sunt analizate trei abordări: procesarea în domeniul timpului, în domeniul frecvenței și o metodă care combină domeniile timpului și al frecvenței. Pornind de la metodele prezentate în lucrare, am aplicat transformările din domeniul timp în domeniile frecvență și timp-frecvență pe un set de date care conține mai multe semnale. Am făcut o analiză a energiei semnalelor și a factorului SNR din cele 3 domenii, urmând să utilizez semnalele din domeniul timp-frecvență pentru a antrena un algoritm de clasificare în funcție de un prag SNR.

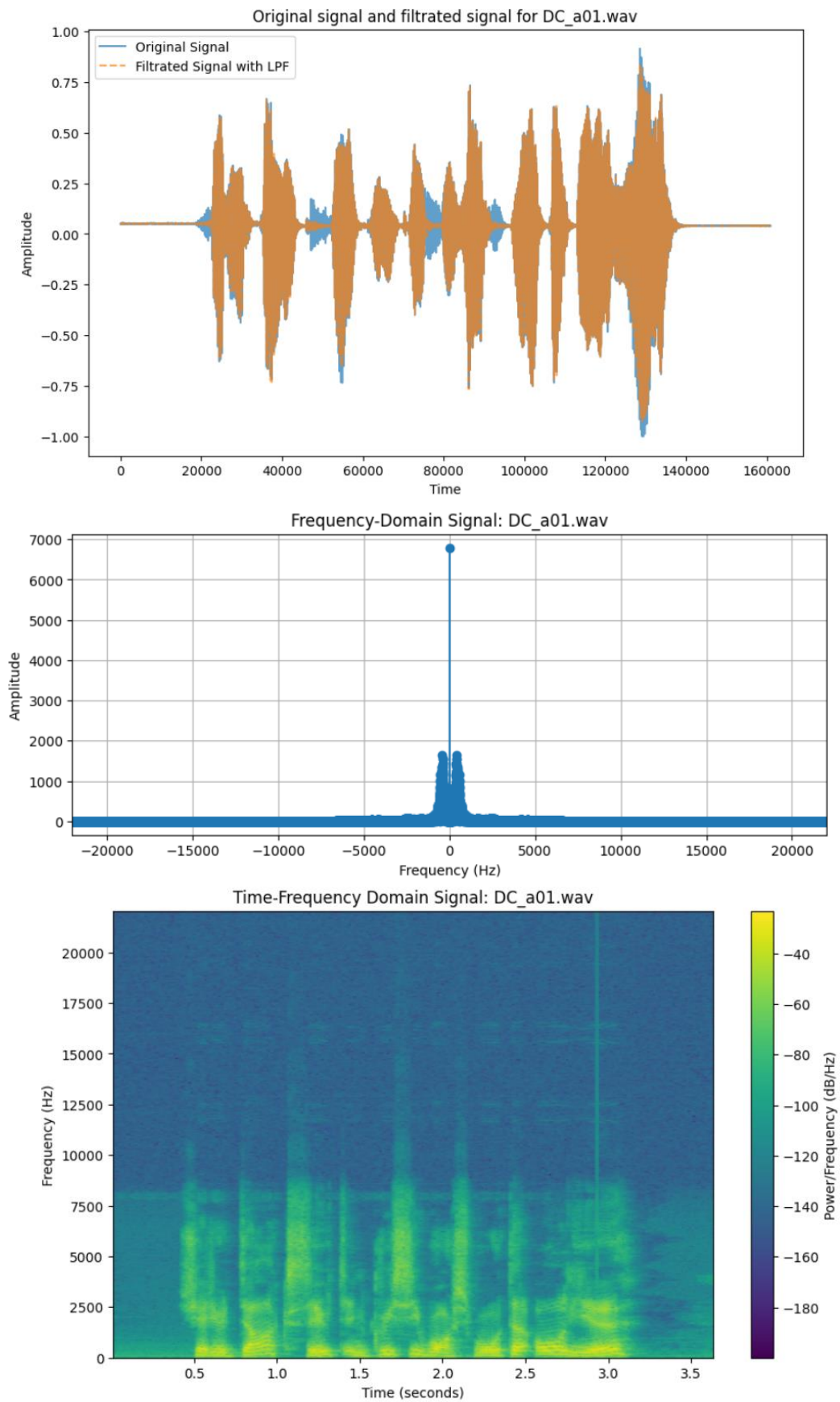
2. Alegerea setului de date

Am ales să folosesc setul de date SAVEE (Surrey Audio-Visual Expressed Emotion), care este un set de date creat pentru a studia recunoașterea emoțiilor umane pe baza semnalelor audio și vizuale. Acesta vrea să prezinte o analiză a expresiilor vocale și să ofere semnale audio reprezentative pentru diverse emoții. Semnalele sunt deja prelucrate minimal, neavând zgomot excesiv. Emoțiile rezultate din tonalitatea vocii pot fi detectate relativ ușor, prin schimbări bruște de tonalitate. Am ales acest set de date întrucât analiza semnalelor în domeniul timp-frecvență poate ușura considerabil detectarea inflexiunii vocilor și interpretarea acestora.

3. Transformările în domeniul timp-frecvență

În proiectul de față am aplicat transformările prezentate în lucrarea citată pentru a transforma semnalele din domeniul timpului în domeniul timp-frecvență. Acestea presupun utilizarea unei spectrograme pentru a descrie energia semnalelor în funcție de timp și frecvență. Se aplică o Transformare Fourier pe intervale scurte de timp (STFT) pentru a analiza semnalul în segmente succesive, iar rezultatul este o matrice 2D: frecvența pe axa verticală, timpul pe axa orizontală, iar amplitudinea este reprezentată prin intensitatea culorii. De menționat este faptul că am aplicat și un filtru trece-jos pe semnalul inițial din domeniul timp pentru a mă asigura că am eliminat cât mai mult posibil din semnalul de frecvență înaltă, păstrându-se doar frecvențele relevante pentru semnalul util, care este adesea concentrat în frecvențe mai joase. Mai jos se poate observa un exemplu de semnal afișat în domeniul timp, pe care este aplicat un filtru trece-jos, apoi reprezentările semnalului în domeniul frecvenței și în cel timp-frecvență.





4. Metrici de analiză a semnalelor

În continuare, am ales să utilizez două metrice pentru evaluarea semnalelor, și anume analiza energiei semnalelor și analiza factorului SNR. Mai departe voi explica rezultatele pentru semnalul prezentat și mai sus.

Results for DC_a01.wav:

Energy in Time Domain: 4318.4736

Energy in Frequency Domain: 370437024.0000

Energy in Time-Frequency Domain: 0.1038

Analiza energiei a returnat o valoare mare a energiei în domeniile timpului și cel al frecvenței, indicând un semnal clare și puternic, cu o concentrație bună în frecvență și timp. Rezultatul mic al energiei din domeniul timp-frecvență indică o dispersie a energiei pe intervale mari de timp și frecvență, lucru care se poate observa și din spectrograma semnalului, prezentată mai sus.

Results for DC_a01.wav:

SNR in the time domain: -1.73 dB

SNR in the frequency domain: 43.57 dB

SNR in the time-frequency domain: 48.67 dB

Analiza factorului SNR indică o valoare chiar negativă la analiza semnalului în domeniul timpului, lucru rezultat de variațiile mari de intensitate din semnalul audio (dat fiind și natura setului de date), având inflexiuni puternice (care pot fi cauzate de sentimentul de furie de exemplu) care trec la inflexiuni slabe (care pot fi datorate de sentimentul de tristețe), semnalul fiind mai greu de distins pe baza formei sale în timp. În schimb, în domeniul frecvenței, un SNR mai mare indică o definiție mult mai clară a semnalului pe anumite frecvențe, iar în domeniul timp-frecvență a rezultat un SNR și mai mare, deci o definiție și mai clară a inflexiunilor semnalului. Se poate observa, astfel, cum metoda trecerii semnalului în domeniul timp-frecvență contribuie la o interpretare mult mai ușoară a semnalelor din setul de date ales.

5. Aplicarea algoritmului de clasificare pe semnalele din domeniul timp-frecvență

Algoritmul utilizat este un model de clasificare care folosește caracteristici extrase din semnale audio pentru a le încadra în două categorii (semnal util sau zgomot). Acesta combină coeficienții MFCC cu factorii SNR în domeniul timp-frecvență pentru a antrena un model de clasificare (în cazul de față, a fost utilizat Random Forest). Ca și set de antrenare, am folosit setul pe care l-am analizat până acum, iar ca date de testare am utilizat un alt set de date cu semnale, parte a aceluiași set de date SAVEE, pe care am aplicat transformările prezentate anterior înainte de a îl utiliza în algoritm. Se disting două clase (0 sau 1) care identifică semnalul ca fiind parte dintr-o categorie cu un nivel mai ridicat de zgomot sau mai scăzut de zgomot. Am realizat și metricile de performanță pentru algoritm, iar rezultatele sunt cele de mai jos, un rezultat destul de bun, având acuratețea de 90%.

	precision	recall	f1-score	support
--	-----------	--------	----------	---------

0	1.00	0.80	0.89	5
---	------	------	------	---

1	0.83	1.00	0.91	5
---	------	------	------	---

accuracy	0.90			10
----------	------	--	--	----

6. Concluzie

Am utilizat metodele de procesare a semnalelor în domeniul timp-frecvență, pornind de la lucrarea “Time-Frequency Signal Processing” de Vladimir Klochko, analiza bazându-se pe transformarea semnalelor audio din setul de date SAVEE în diferite domenii – timp, frecvență și timp-frecvență – pentru a evalua energia semnalelor și factorul SNR. Rezultatele au demonstrat că reprezentarea semnalelor în domeniul timp-frecvență oferă o interpretare mai ușoară și mai detaliată asupra caracteristicilor acestora, îmbunătățind identificarea inflexiunilor și interpretarea emoțiilor reprezentate. Astfel, relevanța utilizării transformărilor în domeniul timp-frecvență pentru procesarea semnalelor audio, mai ales în aplicații complexe, precum recunoașterea emoțiilor este notabilă. Metoda propusă poate fi extinsă și în alte domenii, cum ar fi procesarea vorbirii, analiza biomedicală sau detecția semnalelor radar, având un potențial semnificativ în extragerea și interpretarea caracteristicilor complexe ale semnalelor.

Bibliografie

<https://ieeexplore.ieee.org/document/10510142>