

DISPOZITIV ELECTRONIC DE PREPROCESARE ÎN REGIM PARALEL A SPECTRULUI VOCAL

HULEA MIRCEA¹, UNTU ALINA²

¹*Universitatea Tehnică „Gheorghe Asachi”, Facultatea de Automatică și Calculatoare
Iași – România;*

²*Universitatea Tehnică „Gheorghe Asachi”, Facultatea de Electronică,
Telecomunicații și Tehnologia Informației, Iași – România;*

mhulea@tuiasi.ro, auntu@etti.tuiasi.ro

Rezumat

Lucrarea de față prezintă un dispozitiv hardware bazat pe amplificatoare operaționale a cărui funcție principală reprezintă discriminarea frecvențelor spectrului vocal în funcție de formații specifice vocalelor limbii române. Proiectarea acestui sistem s-a realizat în scopul generării simultane în regim paralel a trenurilor de impulsuri ce stimulează o rețea neuronală analogică utilizată în domeniul recunoașterii vocale independente de vorbitor. Astfel, prototipul prezentat în această lucrare poate fi utilizat până în acest moment pentru evidențierea canalelor de frecvență ce semnalizează recepția vocalelor. Având în vedere rezultatele obținute în acest sens, cercetările ulterioare vor avea ca scop testarea principiilor de funcționare ale acestui sistem și pentru evidențierea canalelor de frecvență activate la recepția consoanelor.

1. Introducere

Un domeniu de o importanță majoră pe plan științific reprezintă procesarea limbajului natural de către structurile de calcul artificiale. Punctul de plecare al activităților desfășurate în această direcție de cercetare constituie recunoașterea vocală a cuvintelor care se poate realiza în mod dependent sau în mod independent de vorbitor. Cea de-a doua modalitate de recunoaștere vocală constă în extragerea și apoi recunoașterea unui set de caracteristici spectrale specifice informației transmise prin vorbire. Extragerea proprietăților ce conduc la recunoașterea vocală independentă de vorbitor a cuvintelor poate fi realizată la nivel software prin elaborarea de algoritmi sau la nivel hardware prin proiectarea de circuite electronice. Scopul acestei lucrări constă în prezentarea principiilor structurale și funcționale ale unui dispozitiv hardware ce realizează preprocesarea semnalului audio în vederea recunoașterii vocalelor limbii române. Structura prototipului a fost elaborată în vederea utilizării acestuia ca element de intrare pentru rețelele neuronale analogice ce au ca unitate funcțională neuronul electronic de inspirație biologică. Funcționarea acestui sistem are la bază principiile de percepție a sunetelor evidențiate la nivelul urechii umane care realizează divizarea spectrului audio pe canale de frecvență (Stevens, 1999). În această direcție, literatura de specialitate prezintă aspecte privitoare la existența unor elemente de preprocesare care au fost proiectate ținând cont de caracteristicile limbii engleze (Hopfield, Brody, 2001), (Wills, 2004), (Loizou, Dorman, 2006). Sistemul prezentat în cadrul acestei lucrări păstrează principiul de divizare pe canale de frecvență a spectrului vocal cu deosebirea că, filtrarea semnalului audio se realizează la nivel hardware folosind un set de parametri ce

permite captarea caracteristicilor spectrale ale vocalelor limbii române. În faza inițială, cercetările proprii efectuate în acest domeniu au folosit un element de procesare a semnalului ce avea la bază filtre rezonante trece-bandă. Performanțele acestora în delimitarea canalelor de frecvență au fost reduse deoarece prezentau o atenuare scăzută a frecvențelor din afara benzii de trecere. Efectul acestui dezavantaj crește invers proporțional cu frecvența de rezonanță a filtrelor care la valori mai mici de 1 kHz necesitau creșterea substanțială a capacităților folosite. Mai mult decât atât, pentru creșterea factorului de calitate a etajelor de filtrare și păstrarea valorilor condensatoarelor în limite acceptabile, a fost necesară utilizarea de bobine speciale cu raportul L/R ridicat. De asemenea folosirea filtrelor rezonante pentru procesarea în regim paralel a semnalului audio are ca efect negativ inductanța mutuală dintre bobine.

Prin urmare, în vederea eliminării dezavantajelor induse de filtrele rezonante, pentru obținerea canalelor de frecvență sistemul actual folosește filtre active de tip Chebyshev a căror ieșire este adaptată pentru stimularea rețelelor neuronale analogice. Frecvențele de tăiere ale acestor filtre trece-bandă s-au calculat prin utilizarea rezultatelor obținute pentru formanții F1 și F2 ai vocalelor limbii române, publicate de către H.N. Teodorescu și M. Feraru pe sit-ul „Sunetele limbii române” (SRoL) prezentat în (Teodorescu et al., 2005-2007), (Teodorescu et al., 2007), (Teodorescu, Feraru, 2007). Dintre aceste date statistice s-au folosit pentru această lucrare valorile medii ale formanților pentru vocale izolate sau în context, extrase de către autori cu utilitarul PraatTM.

În continuare, secțiunea a doua a lucrării va oferi detalii privitoare la caracteristicile funcționale ale modulelor sistemului, precum și la modul de efectuare a statisticii ce a avut ca rezultat obținerea parametrilor elementelor de filtrare a semnalului vocal. Rezultatele obținute în urma testării comportamentului sistemului în cazul rostirii susținute a vocalelor sunt evidențiate în secțiunea a treia, iar concluziile referitoare la performanțele sistemului sunt prezentate în secțiunea a patra.

2. Structura sistemului

Rețelele neuronale analogice au la bază neuroni electronici ce realizează procesarea informațiilor în regim paralel folosind trenuri de impulsuri. Atât frecvența acestor impulsuri cât și energia generată de fiecare impuls pot varia în funcție de gradul de antrenare a rețelei. Pentru ca datele primite din exterior să poată fi procesate cu ajutorul structurilor neuronale de acest tip este necesară conversia în trenuri de impulsuri a semnalelor cu variație continuă. De asemenea, datorită faptului că neuronul artificial de inspirație biologică este prin esență un detector de coincidențe temporale, un interes deosebit trebuie acordat și momentelor apariției stimulilor. Astfel, rețeaua neuronală analogică va fi sensibilă atât la frecvența stimulilor de intrare cât și la gradul de concurență a acestora. Pentru stimularea unei astfel de rețele neuronale ce poate fi antrenată pentru recunoașterea vocală independentă de vorbitor este necesară o procesare prealabilă a semnalului audio. Figura 1 prezintă elementele funcționale ale sistemului de preprocesare a semnalului vocal recepționat de la un microfon. Elementul de preamplificare *PreAmp* realizează amplificarea și filtrarea semnalului audio cu scopul obținerii benzii de frecvențe corespunzătoare spectrului vocal. Semnalul obținut este normalizat de către modulul *DynAmp* prin adaptarea dinamică a amplitudinii

semnalului cu ajutorul circuitului MAX9756. Etapa următoare a procesării semnalului este îndeplinită de modulul *FTB* ce realizează divizarea pe canale de frecvență a spectrului vocal în funcție de formanții specifici vocalelor limbii române.

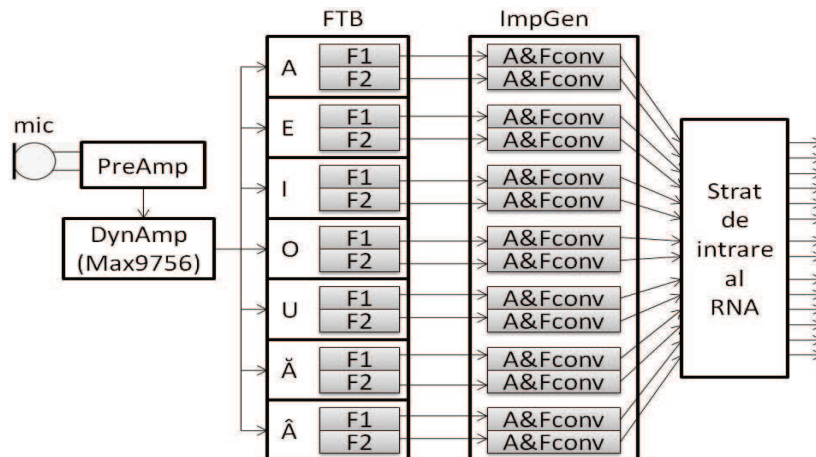



Figura 1: Structura dispozitivului de preprocesare a semnalului audio alcătuită din modulele *PreAmp* (preamplificarea semnalului), *DynAmp* (reglarea amplitudinii semnalului), *FTB* (divizarea pe canale de frecvență a spectrului vocal), *ImpGen* (conversia amplitudinii și frecvenței semnalului); Dispozitivul stimulează stratul de intrare al rețelei neuronale analogice (RNA).

Acest modul constă într-o serie de filtre active de ordin patru ce au fost proiectate folosind amplificatoare operaționale $\beta A741$. Astfel, $F1[v]$ și $F2[v]$ unde $v \in \{a, e, i, o, u, \check{a}, \hat{a}\}$ reprezintă filtre de tip Chebyshev acordate pe frecvențele specifice formanților $F1$ și respectiv $F2$ ai fiecărei vocale. Pentru a fi posibilă funcționarea normală a rețelei neuronale a fost necesară proiectarea unui modul auxiliar *ImpGen* care, cu ajutorul stratului de intrare a rețelei neuronale, realizează adaptarea amplitudinii și frecvenței impulsurilor generate de fiecare canal de frecvență al *FTB*.

2.1 Modulul *DynAmp*

Experimentele anterioare au evidențiat faptul că pentru obținerea unei acuități ridicate în recunoașterea vocală a fonemelor limbii române este necesară rostirea acestora cu o intensitate ridicată și aproximativ constantă. În vederea soluționării acestei probleme sistemul își propune normalizarea la nivel hardware a semnalului recepționat de la microfon după preamplificarea prealabilă a acestuia. Această adaptare de semnal se realizează cu ajutorul circuitului MAX9756 produs de firma , ce îndeplinește funcția de reglare automată a amplitudinii semnalului. Prin utilizarea acestui circuit se reduce considerabil sensibilitatea răspunsului sistemului la variația intensității semnalului vocal de intrare. În continuare se vor prezenta parametrii utilizați în proiectarea schemei electronice de control a circuitului pentru inițializarea funcției de control a amplitudinii semnalului ALC (*Automatic Level Control*). Prin această funcție se amplifică componentele de intensitate scăzută ale semnalului fără a distorsiona componentele de amplitudine ridicată (MAXIM Innovation Delivered™, 2006). Sesizarea depășirii unui anumit prag de putere pe ieșirea circuitului determină activarea ALC ce reduce amplificarea semnalului cu până la 6 dB. Figura 2 (a) ilustrează etapele

acțiunii ALC care sunt *creșterea* în timpul t_a a atenuării puterii semnalului, *blocarea* atenuării la valoarea curentă pe parcursul intervalului de timp t_h și *revenirea* treptată a amplificării la valoarea normală în timpul t_r .

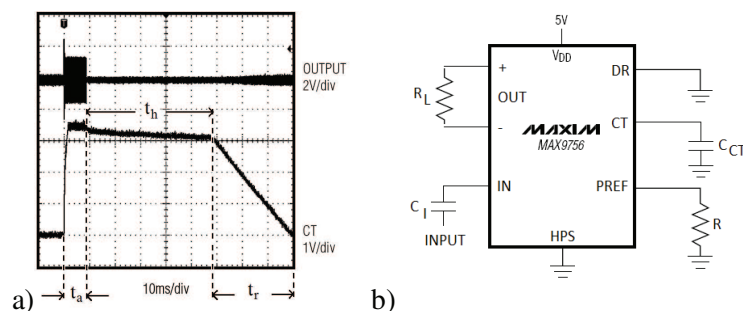


Figura 2. a) digrama de semnal obținută în urma testelor efectuate de producător asupra circuitului MAX9756 ce permite evidențierea etapelor reglării automate a amplitudinii semnalului; b) schema de conectare a pinilor prin care se definesc parametrii funcției ALC (DR, CT, P_REF), precum și a pinilor ce definesc intrarea, ieșirea și selecția ieșirii (IN, OUT și HPS).

Circuitul integrat oferă posibilitatea ajustării timpului de atac t_a și a timpului de revenire t_r prin modificarea parametrilor C_{CT} și DR . Având în vedere scopul pentru care este proiectat dispozitivul, s-a ales valoarea minimă pentru $C_{CT} = 10nF$ deoarece dezavantajele vitezei ridicate de răspuns a ALC (MAXIM Innovation DeliveredTM, 2006) nu sunt sesizabile în cazul recunoașterii vocale. De asemenea, s-a ales $t_r = 300ms$ (maxim) prin conectarea DR la GND pentru a preveni fluctuațiile rapide de intensitate a semnalului vocal pe parcursul cuvintelor.

2.2 Modulul FTB

Pentru a fi posibilă utilizarea rețelelor neuronale analogice în domeniul procesării artificiale a semnalului vocal este necesar ca receptorii specializați în detecția caracteristicilor spectrale ale fonemelor să aibă capacitatea de a fi activați simultan. Acest fapt prezintă o importanță majoră datorită faptului că în mod normal un fonem este caracterizat de un număr de frecvențe a căror prezență în spectrul vocal trebuie semnalizată rețelei neuronale în timp real. Recunoașterea fonemului respectiv se produce cu condiția ca apariția acestor frecvențe la intrarea rețelei neuronale să se producă în mod simultan. Plecând de la această idee cât și de la modelul urechii umane, modulul FTB realizează divizarea pe canale de frecvență a spectrului audio. Scopul acestei abordări reprezintă extragerea în timp real a frecvențelor ce conduc la recunoașterea informației transmise prin vorbire în mod independent de caracteristicile vocii.

2.2.1 Simularea soft a filtrelor Chebyshev

În vederea implementării etajelor de filtrare utilizate în recunoașterea vocală independentă de vorbitor, s-au proiectat cu ajutorul utilitarului *FilterLab* oferit gratuit de *Microchip*, câte două filtre trece-bandă de tip Chebyshev pentru fiecare vocală a limbii române. Având în vedere faptul că formații F1 și F2 (multipli ai frecvenței

fundamentale) sunt implicați în discriminarea vocalelor, pentru recunoașterea acestora s-a implementat câte un filtru ce evidențiază recepția fiecăruia dintre cei doi formați. Parametrii de proiectare utilizați sunt: tip filtru *Chebyshev*; riplu permis în banda de trecere -3 dB; numărul de poli 4; câștig 1. De asemenea frecvențele de tăiere pentru cele șapte vocale (*a, e, i, o, u, ă, â*) sunt prezentate în tabelul 1.

2.2.2 Obținerea frecvențelor de tăiere

Pe baza valorilor medii ale formațiilor preluate de pe sit-ul SRoL, s-a calculat media și deviația standard pentru formații F1 și F2 specifici celor șapte vocale. Pentru stabilirea lățimii benzii de frecvență a celor două filtre corespunzătoare fiecărei vocale, am ales drept frecvență de tăiere minimă f_{\min} și maximă f_{\max} după cum urmează:

$$f_{\min} = f_{\text{central}} - st_dev \quad (1)$$

$$f_{\max} = f_{\text{central}} + st_dev \quad (2)$$

unde f_{central} este frecvența centrală a benzii de trecere, iar st_dev este deviația standard. Atât frecvența centrală cât și deviația standard s-au obținut în urma statisticii efectuate pe un număr de persoane aflate în diverse stări emoționale.

Analiza statistică s-a efectuat în mod diferit pentru grupurile de vocale *a, e, i* și respectiv, *o, u, ă, â* deoarece am avut disponibile valori ale formațiilor specifice vocalelor rostite în contexte diferite. Pentru *a, e, i* au existat valori pentru vocale izolate, iar pentru *o, u, ă, â* au existat valori pentru vocale în context. Cele din urmă au fost extrase din propozițiile „Aseară” și „Vine mama” pentru patru stări emoționale: neutru, fericire, furie, tristețe. Frecvența centrală a filtrelor s-a obținut prin calculul mediilor a 30 de valori pentru vocalele izolate *a, e, i*, iar pentru vocalele în context *o, u, ă, â*, media a fost calculată folosind 40 de valori. La ultimul grup de vocale s-au considerat 10 vorbitori (cinci de gen feminin și cinci de gen masculin) pe patru stări emoționale. Rezultatele obținute sunt ilustrate în tabelul 1.

Tabel 1: Valorile frecvențelor medii, minime și maxime corespunzătoare formațiilor vocalelor

Vocală	a		e		i		o		u		ă		â	
Formant	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
f_{central}	751	1274	583	1889	399	2328	466	915	403	788	542	1433	389	1648
f_{\min}	618	1093	486	1671	338	2091	361	794	333	708	432	1192	325	1428
f_{\max}	883	1456	680	2107	460	2565	570	1036	473	868	651	1674	453	1867

Schema generală a unui filtru trece bandă de tip Chebyshev alcătuit din două amplificatoare operaționale este dată în figura 3.

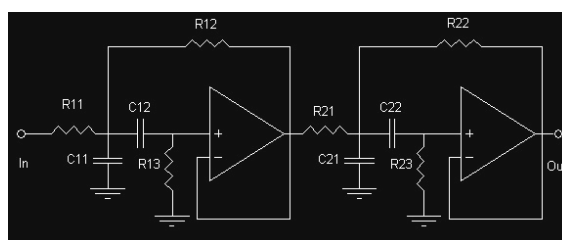


Figura 3: Schema generală a unui filtru trece bandă Chebyshev.

Pentru evidențierea comportamentului canalelor de frecvență implementate, în figura 4 este prezentată simularea în *FilterLab* a răspunsului în frecvență a filtrului trece-bandă pentru formantul F2 al vocalei *e*. Frecvențele de tăiere corespunzătoare la -3dB sunt $f_{\min} = 1671\text{Hz}$ și $f_{\max} = 2107\text{Hz}$, iar lățimea benzii de trecere este egală cu dublul deviației standard

$$\Delta f_b = 2 \cdot st_dev = 436\text{Hz} \quad (3)$$

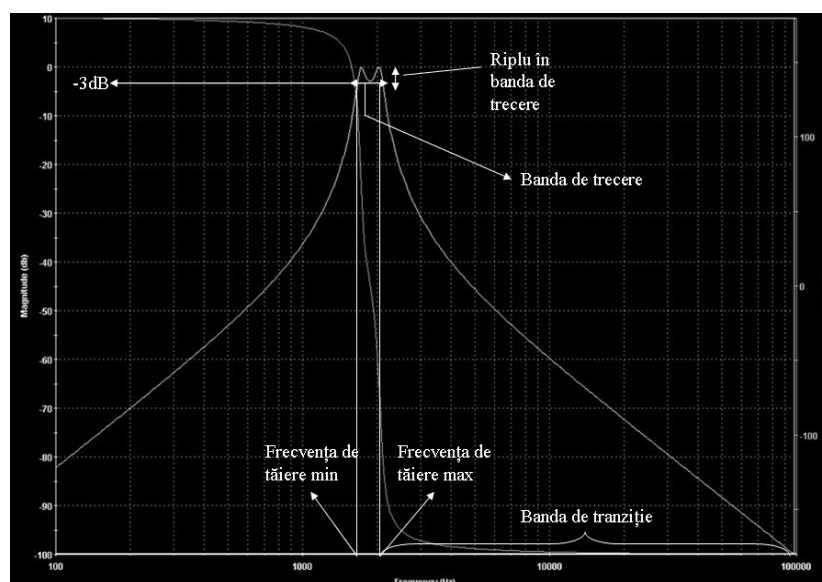


Figura 4: Răspunsul în frecvență a filtrului trece-bandă de tip Chebyshev.

Filtrele Chebyshev de tipul I sunt filtre polinomiale (numai cu poli, fără zerouri), având o caracteristică de modul cu ripluri egale în banda de trecere și monoton descrescătoare în banda de oprire. Dintre toate filtrele polinomiale de ordinul N , filtrele Chebyshev de tipul I au zona de tranziție cea mai îngustă. Performanțele filtrului sunt complet precizate de mărimea riplului în banda de trecere și de ordinul filtrului ce determină panta caracteristicii de modul. Utilizând programul *FilterLab* se pot proiecta trei tipuri de filtre: Bessel, Butterworth și Chebyshev. Filtrul Bessel se poate realiza doar pentru tipul de răspuns trece-jos, iar celelalte două pentru răspunsuri de tip trece-jos, trece-sus, și trece-bandă. În comparație cu filtrul Butterworth, filtrul Chebyshev realizează o mai bună atenuare a frecvențelor din afara benzii de trecere, acest fapt justificând alegerea făcută în procesul de proiectare a filtrelor.

2.2.3 Delimitarea canalelor de frecvență

Semnalul sinusoidal generat de filtrele trece-bandă prezentate anterior este convertit în impulsuri prin utilizarea unui comparator LM324 pentru fiecare formant. Acest etaj permite și reglarea amplitudinii de la care activitatea canalului de frecvență va produce stimularea rețelei neuronale. Prin urmare, ieșirea generată de modulul *FTB* se materializează într-un tren de impulsuri de amplitudine, durată și frecvență variabile ca în figura 5(a). Energia acestor impulsuri este convertită în trenuri de impulsuri de amplitudine și frecvență constantă de către modulul *ImpGen* ce constituie o interfață între modulul de filtrare și stratul de intrare al rețelei neuronale.

2.3 Modulul *ImpGen*

Experimentele realizate pe parcursul cercetărilor efectuate în domeniul recunoașterii vocale au arătat că pentru creșterea stabilității în funcționare a rețelei neuronale analogice este necesară o adaptare a amplitudinii și frecvenței impulsurilor ce stimulează intrările acestora (Hulea, 2009). Schema electrică a acestui modul conectată cu un neuron electronic specific stratului de intrare al rețelei este prezentată în figura 5 (b).

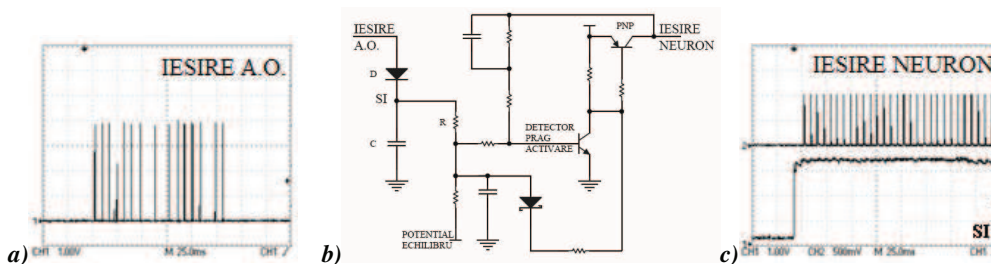


Figura 5. a) semnalul de ieșire a unui canal de frecvență al modulului *FTB*; b) schema electrică a ansamblului [modul *ImpGen* - neuron de intrare] pentru un canal de frecvență oarecare; c) ieșirea neuronului de intrare (semnalul IESIRE NEURON) și integrarea ieșirilor modulului *FTB* de către capacitatea *C* (Semnalul SI).

Conversia ieșirii modulului *FTB* constă în integrarea stimulilor generați și transformarea energiei obținute în impulsuri de amplitudine și frecvență constante. Prima etapă se realizează folosind capacitatea *C* ce se încarcă prin dioda *D*, iar cea de-a doua este asigurată de componentele electronice ce compun neuronul de intrare (Hulea, 2008). Frecvența impulsurilor de ieșire este invers proporțională cu valoarea rezistenței *R* prin care are loc descărcarea condensatorului *C*. În cazul sistemului descris în această lucrare, valoarea de 470 k Ω a acestei rezistențe a fost aleasă pentru a obține la ieșire o frecvență de aproximativ 160 Hz.

3. Rezultate experimentale

Răspunsurile în frecvență ale filtrelor de tip Chebyshev obținute în urma simulărilor efectuate în programul *FilterLab* au arătat că performanța acestui tip de filtru se

îmbunătățește proporțional cu ordinul acestuia. Prin urmare, odată cu creșterea numărului de etaje de filtrare utilizate, crește și panta caracteristicii de modul a filtrului. Pe de altă parte, panta maximă de tranziție se obține prin delimitarea strictă a spectrului util cu ajutorul comparatoarelor LM324 implementate pe ieșirea fiecărui canal de frecvență. Stabilitatea limitelor benzii de trecere obținute în acest mod este asigurată în cazul în care variația amplitudinii semnalului este redusă la minim. Această condiție este îndeplinită prin reglarea dinamică a amplitudinii semnalului de către modulul *DynAmp* conectat pe intrarea modului *FTB*. Pentru ilustrarea comportamentului descris anterior, figura 6 prezintă câteva aspecte legate de funcționarea canalului de frecvență centrală 542 Hz corespunzător formantului F1 al vocalei /ă/. Formele de undă prezentate au fost obținute în urma evidențierii cu ajutorul osciloscopului a semnalelor generate prin nodurile de conectare a celor patru module ale sistemului. Vocalele au fost rostite independent de context la o distanță de aproximativ 3cm de microfon.

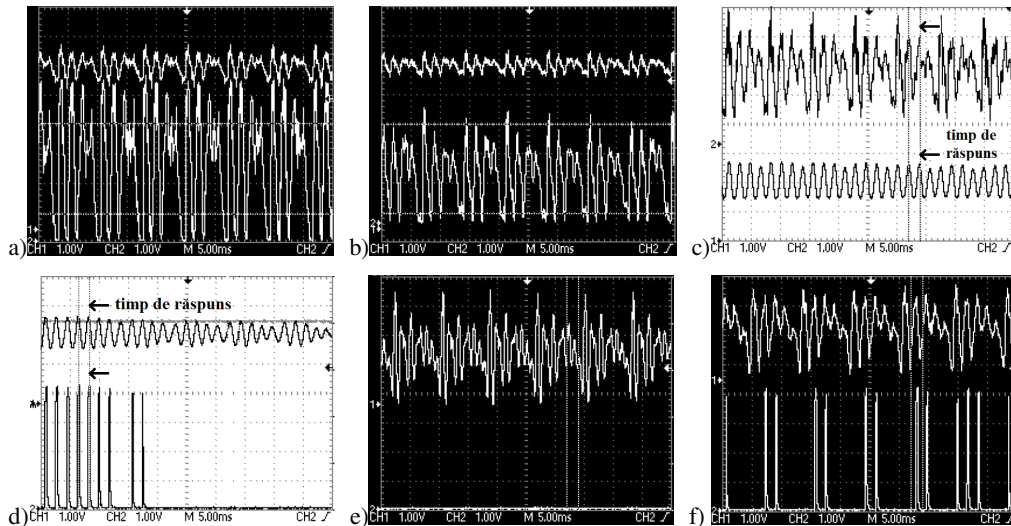


Figura 6. a) Intrarea (sus) și ieșirea (jos) a modului *DynAmp* când funcția de reglare automată a amplitudinii semnalului (ALC) este dezactivată; b) intrarea (sus) și ieșirea (jos) a modului *DynAmp* după activarea ALC; c) ieșirea *DynAmp* (sus) și ieșirea filtrului Chebyshev corespunzător formantului F1 (jos) în timpul rostirii vocalei susținute a vocalei /ă/ (cursorii verticale evidențiază timpul de răspuns a *FTB*); d) Semnalul generat de filtrul Chebyshev (sus) și ieșirea modului *FTB* corespunzătoare formantului F1 (jos) a vocalei /ă/ la rostirea vocalei /ă/; semnalul continuu reprezintă pragul de la care semnalul generat de canalul de frecvență este considerat util (cursorii verticale evidențiază timpul de răspuns a *ImpGen*); e) ieșirea circuitului MAX9756 la rostirea vocalei /ă/ și ieșirea *FTB* corespunzătoare formantului F1 al vocalei /ă/ f) semnalul generat de *DynAmp* la rostirea vocalei /ă/ și ieșirea *FTB* corespunzătoare formantului F1 al vocalei /ă/.

Așadar primele două diagrame de semnal (a) și (b) din figura 6, evidențiază importanța funcției ALC a circuitului MAX9756 în adaptarea amplitudinii semnalului generat de modulul *PreAmp*. Semnalele prezentate în figura 6 (c) prezintă efectul produs de filtrul Chebyshev de frecvență centrală 542 Hz corespunzător formantului F1 al vocalei /ă/ în momentul rostirii susținute a acesteia. Diagrama de semnal (d) ale aceleiași figuri prezintă pragul de la care activitatea canalului de frecvență va fi considerată utilă. În

continuare sunt prezentate rezultatele unui test efectuat asupra elementelor de filtrare corespunzătoare canalului de frecvență menționat. Pentru efectuarea acestui test s-a rostit vocala /a/ și s-a verificat activitatea ieșirii *FTB* corespunzătoare formantului F1 al vocalei /ă/. După cum se poate observa, rezultatele testului au fost cele așteptate în sensul că vocala /a/ nu a produs activarea canalului de 542 Hz (diagrama de semnal (e)), acesta fiind activat la recepția vocalei /ă/ (diagrama de semnal (f)).

Valorile obținute pentru frecvențelor de tăiere ale filtrelor (tabelul 1) au impus efectuarea unor ajustări în implementarea dispozitivului pentru eliminarea elementelor de filtrare redundante. În acest sens, pentru formantul F1 al vocalelor *i, u*, și *â* s-a utilizat același filtru de frecvență centrală 400 Hz, iar pentru formantul F1 al vocalelor *e* și *ă* s-a utilizat un filtru de frecvență centrală 560 Hz. Diferența de aproximativ 200 Hz dintre valorile medii obținute pentru formantul F2 fac posibilă discriminarea vocalelor limbii române folosind doar filtrele corespunzătoare acestui formant.

Un alt aspect important ce a fost luat în considerare în procesul de testare a sistemului reprezintă evaluarea timpului de răspuns a acestuia. Cursorii verticali de pe diagramele de semnal (c) și (d) ale figurii 6 evidențiază timpii de răspuns a submodulelor *FTB* și respectiv *ImpGen* care sunt neglijabili în comparație cu perioada semnalelor specifice spectrului vocal. De asemenea diagramele de semnal (a) și (b) ale aceleiași figuri evidențiază cu ușurință timpul de răspuns a modulului *DynAmp* care de asemenea este neglijabil. Prin urmare se poate deduce faptul că timpul total de răspuns a întregului sistem de preprocesare este neglijabil în comparație cu variația semnalelor specifice spectrului vocal.

4. Concluzii

Pentru a beneficia de avantajele utilizării rețelelor neuronale cu răspuns în timp real în domeniul recunoașterii vocale, s-a proiectat un sistem de preprocesare în regim paralel a semnalului audio. Plecând de la modelul aparatului auditiv uman ce realizează divizarea spectrului vocal pe canale de frecvență s-a implementat un set de filtre trece-bandă acordate pe frecvențele caracteristice formațiilor vocalelor limbii române. În vederea obținerii unei acuități ridicate în recunoașterea vocală independentă de vorbitor s-a urmărit creșterea performanțelor sistemului de filtrare. Aceasta s-a realizat atât prin utilizarea filtrelor de tip Chebyshev care prezintă cea mai bună atenuare a frecvențelor din afara benzii de trecere, cât și prin folosirea circuitului MAX9756 care realizează adaptarea dinamică a amplitudinii semnalului. De asemenea, procesul de selecție a frecvențelor de tăiere pentru aceste filtre s-a bazat pe date statistice ce cuprind o serie de voci de gen masculin și feminin precum și diferite stări emoționale.

Experimentele efectuate pentru studiul funcționării sistemului prezentat în această lucrare au arătat o superioritate evidentă a performanțelor filtrelor Chebyshev în comparație cu filtrele rezonante utilizate în cadrul cercetărilor anterioare descrise în (Hulea, 2008). Performanțele sistemului s-au îmbunătățit substanțial mai ales în cazul canalelor de frecvențe centrale mai mici de 1000 Hz. Un alt avantaj al filtrelor Chebyshev reprezintă creșterea eficienței filtrelor prin introducerea elementelor active (A.O.) și eliminarea componentelor inductive (L). De asemenea capacitățile folosite s-au redus cu trei ordine de mărime ceea ce reduce costurile implementării. De asemenea, îmbunătățirea performanțelor în filtrare a fost obținută fără a afecta răspunsul în timp

real a sistemului. Totuși o consecință a utilizării circuitului MAX9756 pentru adaptarea amplitudinii semnalului audio reprezintă creșterea consumului de curent al dispozitivului cu aproximativ 48 mA. Cu toate acestea, acest aspect trebuie luat în considerare doar în cazul utilizării unei surse portabile de energie pentru alimentarea sistemului. Având în vedere creșterea acuității în delimitarea benzilor de frecvență a sistemului de filtrare se preconizează pe de o parte îmbunătățirea performanțelor în domeniul recunoașterii vocale independente de vorbitor, cât și simplificarea prin reducerea numărului de neuroni a arhitecturii rețelei neuronale utilizate în detecția cuvintelor.

Referințe bibliografice

- Hopfield, J. J., Brody C.D. (2001). What is a moment? Transient synchrony as a collective mechanism for spatiotemporal integration, *Proceedings of the National Academy of Sciences of the USA*, 1282 – 1287.
- Hulea, M. (2008). A Model of Silicon Neurons Suitable for Speech Recognition, *Computer Engineering and Applied Informatics*, Vol.10, Nr. 4, 32-41, ISSN 1454-8658.
- Hulea, M. (2009). A New Method to Obtain Non-Volatile Memory for Networks of Spiking Neurons, *Memoirs of the Scientific Sections, Romanian Academy* în curs de apariție.
- Loizou P.C., Dorman, M. (2006). On the Number of Channels Needed to Understand Speech, *Journal of the Acoustical Society in America*.
- MAXIM Innovation Delivered™ (2006). 2.3W Stereo Speaker Amplifiers and DirectDrive Headphone Amplifiers with Automatic Level Control.
- Stevens, K. (1999). *Acoustic Phonetics*, Editura Cambridge MA: The MIT Press, 204 - 214.
- Teodorescu, H.N., Feraru, S.M., Trandabăț, D., Zbancioc, M., Luca, R., Verbuță, A., Hnatiuc, M., Ganea, R., Voroneanu, O., Pistol, L., Șcheianu, D. (2005-2007), situl Web Sunetele Limbii Române.
http://www.etc.tuiasi.ro/sibm/romanian_spoken_language/index.html
- Teodorescu, H.N., Trandabăț, D., Feraru, S.M., Zbancioc, M., Luca, R. (2007). A corpus of the sounds in the Romanian spoken language for language-related education, Chapter Six, pp. 73-90. În volumul Carlos Perrián Pascual (Editor), „Revisiting Language Learning Resources”, Cambridge Scholars Publishing (CSP), UK, ISBN 1-84718-156-2; ISBN 13: 9781847181565, 2007.
- Teodorescu, H.N., Feraru, S.M. (2007). A study on Speech with Manifest Emotions, *10th International Conference on Text, Speech and Dialogue*, TSD 2007, Pilsen, Czech Republic, September 3-7, 2007, *Lecture Notes in Computer Science*, Springer Verlag, vol. 4629/2007, 254-262, ISBN 978-3-540-74627-0.
- Wills, S. (2004). Computation with spiking neurons, *teză de dizertație trimisă la University of Cambridge pentru admiterea la doctorat*, 5-12.