

SINTEZA SUNETELOR SPECIALE ȘI MESAJE CU VOCE PENTRU CERCETĂRI BIOACUSTICE

S. Tincovan, Iu. Sorocean
s_tincovan@mail.md, iuis@mail.md

Universitatea Tehnică a Moldovei

Abstract: *In the full cycle of technological process of incubation and growing in fowling for increasing of efficiency is used the bioacoustic effects and spectral analysis of the sound reaction from biological object. In the work are considered structures of devices for analysis and syntheses of sound in bioacoustics that can be applied not only in fowling, but also in the other area such as syntheses of many language speech messages for diagnostics, signalling and training.*

Cuvinte-chee: *Boacustică, delta-modulare, analiză spectrală, fonemă, formanta frecvențelor, sunet vocal, sunet consonant.*

1. Introducere

În procesele biotehnologice pentru sporirea eficienței se utilizează influențarea bioacustică a obiectului biologic, unde este necesar de sintetizat sunete speciale de stimulare pentru ecloziune sincronă a puilor și selectarea lor după puterea biologică după ieșire din ou (pentru faza de incubatie), stimularea lor ulterioară în proces de creștere cu scopul măririi gradului de asimilare a nutrețurilor (pentru faza de creștere). Utilizarea influențării bioacustice permite de scurtat durata ciclului tehnologic fără a pierde din calitatea producției finite sau (și) mărirea a unor parametri de calitate.

În cadrul analizei spectrului semnalului bioacustic prevede separarea după sex a puilor în prima zi de viață, imitarea dialogului „părinte-copil” (influențare bioacustică pe baza rezultatelor analizei spectrului), diagnosticarea bolilor după parametrii vocii obiectului biologic și depistarea situațiilor de alarmă din partea răpitorilor (păsări și animale) și catastrofe (calamități) naturale.

În calitate de funcție suplimentară de generare a sunetelor speciale pot servi mesajele cu voce a situațiilor critice pentru toate etapele ciclului tehnologic în avicultură, unul din momentele specifice este menținerea sintezei mesajelor cu voce în mai multe limbi concomitent. Aceste posibilități de sinteză a mesajelor cu voce pot fi utilizate nu numai în procesele tehnologice în avicultură, ci și în procesele de instruire, educație, studierea limbilor moderne, în sistemele de diagnostică, sisteme de alarmă și etc.

2. Formularea sarcinii

În lucrarea dată se analizează structura și algoritme de analiză și sinteza sunetelor speciale a dispozitivelor universale pentru bioacustică și sinteza mesajelor cu voce.

Pentru analiza spectrală și sinteza sunetelor este necesar de analizat două module aparte, unul pentru analiză pe baza transformărilor Fourier discrete (DFT) sau transformări Fourier rapide (FFT), alt modul pentru sinteza sunetului.

3. Formularea problemei

Pentru mărirea eficienței proceselor biotehnologice este necesar de soluționat următoarele sarcini:

1) De obținut o variantă optimală a structurii și algoritmului de funcționare pentru modului de analiză spectrală a semnalului bioacustic cu ajutorul DFT sau FFT cu posibilitatea prelucrării mai profunde la calculator.

2) De obținut structura și algoritmul de funcționare optimal a modului de sinteză a sunetelor și mesajelor cu voce, unde dicționarul (vocabularul) de sunete și foneme este nelimitat.

Pe parcursul soluționării este necesar de ținut cont de repartizarea funcțiilor executate între partea HARD și SOFT a modului de analiză spectrală și sinteză a semnalelor bioacustice și mesajelor cu voce.

4. Căile de soluționare și implementare

Pentru implementarea modului de analiză spectrală au fost elaborată schema de structură (fig.1), în care este implementat DFT (FFT) la nivel HARD cu posibilitatea restabilirii semnalului inițial și transmiterea torentului de date la calculator [1]. Ca o alternativă poate fi implementarea FFT în varianta SOFT pe microprocesor RISC.

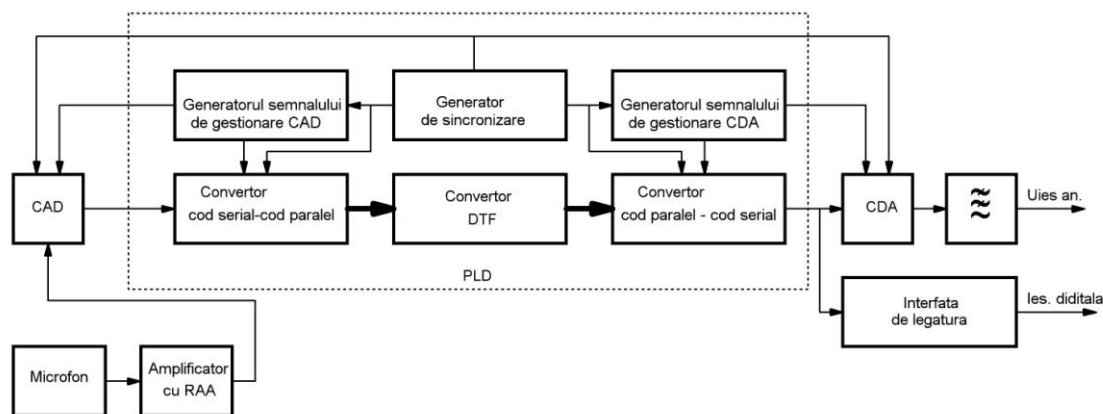


Fig. 1. Schema de structură a modului universal cu DFT implementat.

Analiza științifică și implementarea structurii HARD-ului și SOFT-ului pentru modului universal de sinteză a sunetelor speciale și mesajelor cu voce în mai multe limbi. Variantele posibile de realizare a sintezei sunetelor speciale și vocii pot utiliza următoarele tehnologii:

- 1) Metoda directă de codare-restabilire a semnalelor sunetelor și vocii;
- 2) Modelare digitală a tractului vocii;
- 3) Sinteza analogică a formantelor de frecvență;
- 4) Sinteza digitală a fonemelor;
- 5) Codarea vocii cu ajutorul coeficienților liniari de predicție (CLP)

Pentru metoda directă de codare-restabilire a semnalelor sunetelor și vocii codarea semnalelor vocii se efectuează cu ajutorul PCM (Pulse Code Modulation) sau delta-modulare [2].

Informația despre semnal vine în formă succesivă de eșanționări cu frecvența de discretizare f_{discr} , Conform teoremei Kotelnicov este necesar de respectat $2 \cdot f_{max} \leq f_{discr}$, unde f_{max} este frecvența maximă a spectrului semnalului vocii. Această relație este valabilă pentru filtrare ideală a funcției restabilite, pentru condițiile reale este acceptată condiția $f_{discr} = 45 \cdot f_{max}$.

Pentru intervalul de frecvențe, unde valoarea maximă nu depășește 8-10 KHz în tractul preliminar de prelucrare a semnalului se include filtru trece jos cu scopul de a elimina frecvențele mai înalte $f_{discr}/2$ pentru intrarea CAD. Pentru imprimarea și redarea vocii este necesar să fie asigurată compatibilitatea CAD și DAC după binaritate și frecvența discretizării. Rata de transfer a datelor constituie circa 96 Kbit/s pentru o secundă de sonorizare, care permite de simplificat partea HARD a echipamentului și asigură calitate suficientă a sintezei sunetului. În calitate de alternativă a metodei directe de codare-restabilire a semnalelor sunetelor și vocii poate servi delta-modularea, care permite de redus rata de transfer, care se bazează pe variația relativă a amplitudinii și nu pe valorile absolute. Semnalul vocii ca și pentru PCM este prelucrat preliminar, care apoi este supus delta-modulării.

Sinteza analogică a formantelor de frecvență se bazează pe cunoașterea detaliată a fonemelor și descompunerii fonetice a mesajelor, la temelia cărora stau două noțiuni: lingvistică – foneme, și acustică – formantă.

Pentru sinteza vocii sunt suficiente 2-4 frecvențe a formantelor, unde prima formantă are frecvența 200 Hz (prima formantă a vocii bărbătești) până la 2000 Hz (a treia formantă a vocii feminine). În procesul vorbirii frecvențele formantelor și obertoanelor sunt prezente simultan și se deplasează pe axa spectrului, care corespund particularităților cuvântului pronunțat. Deaceia mesajele vocale sau sunetele lumii animale se percep nu ca o singură frecvență, ci o mulțime de armonici, care se formează la filtrarea impulsurilor, formate la ieșirea tractului vocal (fig. 2).



Fig. 2. Modelul tractului vocal.

Schema de structură a modului permite sinteza a sunetelor vocale și consoanante, din care se poate de format sunete speciale și mesaje (fig. 3), unde sunetele sunt sintetizate în ordinea prescrisă de vocabular (dicționar) [3].

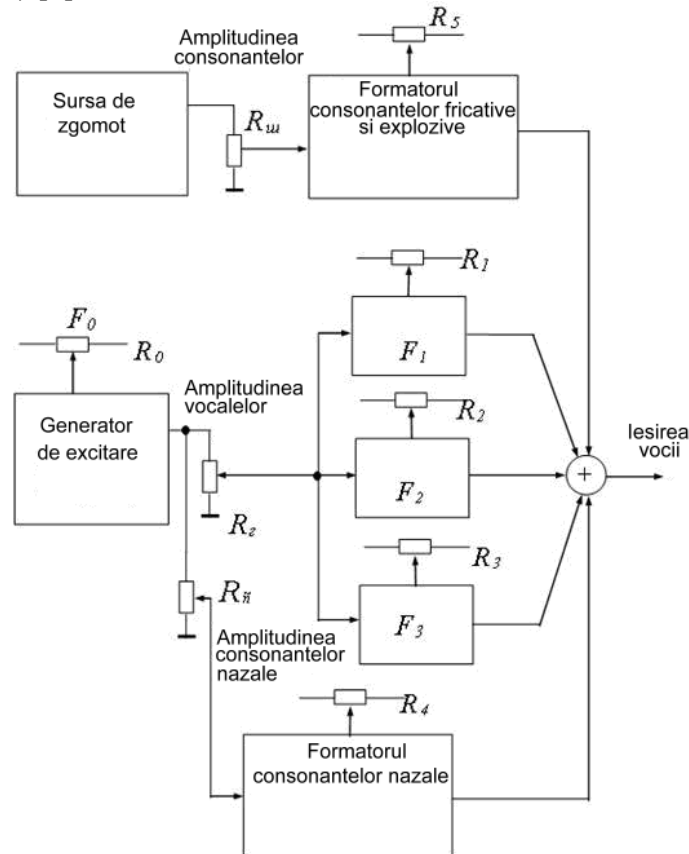


Fig. 3. Schema de structură a sintetizatorului sunetelor vocale și consoane.

Sinteza digitală a fonemelor se bazează pe generarea fonemelor și compilarea ulterioară din ele a sunetelor speciale, cuvinte, propoziții și fraze. Realizarea ei combină metode compacte de prelucrare digitale și flexibilitatea de gestionare cu parametrii principali a vocii, ce este specific modelelor de formante. Procesul codării vocabularului necesar este substituit cu compilarea mesajelor arbitrare dintr-un set de elemente a vocii, pregătit preliminar [4].

Sinteza fonemelor conține 3 nivele de prelucrare, unde la primul nivel se efectuează translatarea simbolurilor orfografice în codurile fonemelor, la nivelul doi se calculează setul parametrilor acustici, care servește pentru gestionarea nivelului trei – formarea semnalului vocii.

Pentru nivelul doi de prelucrare se calculează setul parametrilor de gestionare, care reprezintă frecvențele formantelor (F_1, F_2, F_3), banda de frecvențe ($\Delta F_1, \Delta F_2, \Delta F_3$), frecvența tonului principal și amplitudinea de vocalizare. Setul parametrilor regenerează cu interval de 6,4 ms, ce asigură de urmărit cele mai rapide variații între foneme, rata de transfer de la nivelul doi la nivelul trei constituie circa 45 Kbit/s. Sintezarea propriu zisă se efectuează la nivelul trei, unde semnalele de excitație (armonic și de zgomot) se filtrează cu ajutorul filtrelor de rezonanță a formantelor a tractului vocii.

Codarea vocii cu ajutorul coeficienților liniari de predicție (CLP) se bazează pe teoria analizei statistice a seriilor de timp. Seria de timp reprezintă o succesiune de observații (referințe) aranjate în timp. Esența acestei metode este următoare. Admitem succesiunea seriei a semnalului vocal x_1, x_2, \dots, x_t . Pentru această mulțime se calculează "valoare medie". Pentru un interval prestabilit proprietățile statistice a semnalului vocal se consideră neschimbate și acest interval se codează cu un set de coeficienți a_s , care minimizează eroarea medie patrică a predicției, (reduce la minim eroarea între seria inițială și "netezită").

5. Concluzii

Pentru implementarea modulului de analiză spectrală preferința a fost în folosul realizării DFT cu mijloace HARD (circuitul integrat EP2C35U484C6, firma Altera), restul funcțiilor auxiliare pot fi executate de microprocesor specializat. Pentru intervalul analizat de frecvențe până 10 kHz binaritatea CAD și ADC este de 16 biți, pentru frecvența discretizării 48 kHz circuitele integrate de tipul ADS8320 și DAC8830 satisfac condițiilor cerute.

Pentru implementarea modulului de sinteză din variantele menționate pentru cerințele de menținere dicționarului nelimitat cele mai reușite metode este analiza formantelor (poate sinteza orice sunet a lumii animale) și metoda fonemică (poate sinteza mesaje în mai multe limbi). Alt argument în folosul sintezei fonemice este cea mai joasă rată transfer pentru canalele de legătură (tabelul 1.), unde se păstrează calitatea sunetului sintetizat.

Tabelul 1.

Metoda de codare a vocii	Cheltuieli informaționale, bit/s
PCM	$(40-100) \cdot 10^3$
Delta-modulare	$(20-50) \cdot 10^3$
Delta-modulare adaptată	$(10-25) \cdot 10^3$
Analiza formantelor	$(2-4) \cdot 10^3$
CLP-analiză	$(1,2-2) \cdot 10^3$
Metodă fonemică	50-100

O variantă de compromis poate fi combinarea sintezei digitale fonemelor cu redarea directă PCM sau delta-modulare, unde sample-ul al dictatorului-donor pot fi prelucrate pe scara timpului real (modificare tonului, amplitudinii și tempoului redării).

6. Bibliografie

1. Г. Нуссбаумер. Быстрое преобразование Фурье и алгоритмы вычисления сверток.— М.: Радио и связь, 1985.
2. А. В. Фролов, Г. В. Фролов. Мультимедиа для Windows. Библиотека системного программиста, т. 15 М: Диалог-МИФИ, 1994
3. Л. Захаров. Проблемы создания аллофонной базы автоматического синтеза речи (<http://art.bdk.com.ru/govor/rasp.htm>).
4. Black A.W., Taylor P. Automatically clustering similar units for unit selection in speech synthesis // In Proceedings of Eurospeech 97. Rhodes, Greece, 1997. Vol.2, pp. 601-604.