

L2. Estimarea Frecvenței Fundamentale – în domeniul timp

Problema estimării frecvenței fundamentale constă în preluarea unei porțiuni din semnal și găsirea frecvenței dominante de repetiție. Dificultatea rezultă din faptul că:

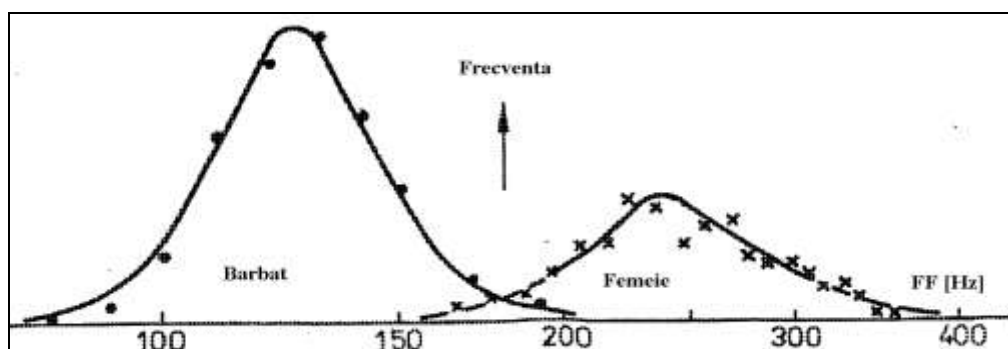
- nu toate semnalele sunt periodice;
- cele ce sunt periodice pot să-și modifice frecvența în zona de interes;
- semnalele pot fi contaminate de zgomot, chiar cu semnal periodic cu altă frecvență fundamentală;
- semnalele care sunt periodice cu intervalul T sunt de asemenea periodice cu $2T$, $3T$ etc, astfel trebuie găsit cel mai mic interval de periodicitate sau FF cea mai mare;
- chiar și semnalele de FF constantă se pot modifica în alte moduri pe domeniul de interes.

Există mai multe metode de determinare a FF în domeniul timp și frecvență.

1. Metode de determinare a frecvenței fundamentale a semnalului vocal în domeniul timp

Frecvența fundamentală (F_0) este mărimea care se definește pentru porțiunile sonore ale semnalului vocal și reprezintă frecvența de oscilație a coardelor vocale. Corespondentul acustic al frecvenței fundamentale este perioada fundamentală (T_0) sau *pitch*. Frecvența fundamentală poate lipsi din semnalul vocal (de exemplu în vorbirea telefonică). Perioada fundamentală este prezentă chiar și în semnalele vocale de bandă limitată.

Media și deviația standard ale frecvenței fundamentale pentru vocile masculine este de 125 Hz și respectiv 20.5 Hz, iar pentru cele vocile feminine aceste valori sunt duble. De asemenea s-a constatat că variația frecvenței fundamentale în timpul vorbirii este lentă, sub 10 Hz.



Distribuția frecvenței fundamentale pe sexe.

Pentru a adapta rezultatele la rezoluția sistemului auditiv uman, frecvența fundamentală este adesea reprezentată pe scară logaritmică astfel [Pic93]:

$$F(n) = \text{Log}_{10}(F_0(n))$$

Pentru porțiunile de vorbire vocalizată, frecvența fundamentală este cuprinsă între 50 Hz și respectiv 500 Hz, iar pentru porțiunile nevocalizate se consideră, prin convenție, că F_0 are valoarea 0.

Datorită complexității semnalului vocal, frecvența fundamentală nu poate fi determinată cu un algoritm simplu. În timpul sunetelor vocalizate, aerul este împins prin tot tractul vocal, inclusiv corzi vocale, gât și gură. Părțile articulatorii (limba, gâtul și buzele) filtrează semnalul generat de corzile vocale. Această filtrare poate ascunde informația legată de frecvența fundamentală prin scăderea energiei în zona frecvenței fundamentale și prin creșterea energiei în zona altor frecvențe.

Un semnal vocal nu are o perioadă fundamentală pe întreaga durată. Astfel pe timpul pauzelor în vorbire sau pentru sunetele nevocalizate, perioada fundamentală lipsește. Dintre sunetele nevocalizate, cele fricative prezintă un interes special, acestea fiind generate prin restrângerea curgerii aerului prin corzile vocale pe o perioadă de timp relativ lungă. Pentru a determina dacă un segment vocal anume are perioadă fundamentală se poate folosi energia medie. Pentru perioadele

de pauză sau pentru sunetele fricative, energia medie are valori mult mai mici decât pentru sunetele vocalizate.

Există mai multe modalități de calcul a frecvenței fundamentale : în domeniul timp, în domeniul frecvență sau prin metode timp-frecvență.

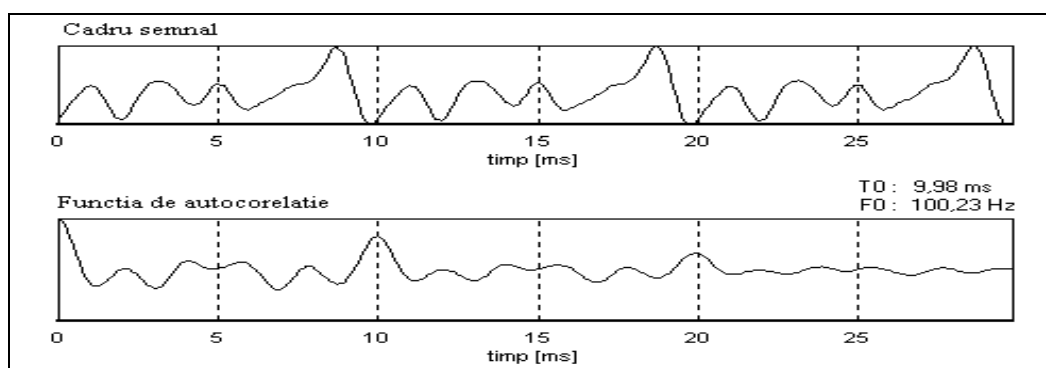
În domeniul timp majoritatea metodelor se bazează pe funcția de autocorelație. În continuare vor fi prezentate rezultate obținute prin metoda autocorelației, metoda funcției diferență de amplitudine medie și metoda cu limitare centrală.

1.1. Metoda autocorelației

Funcția de autocorelație este definită ca [Sto84]:

$$R_n(k) = \sum_{m=0}^{N-1-k} [x(m+n)w(m)][x(n+m+k)w(m+k)]$$

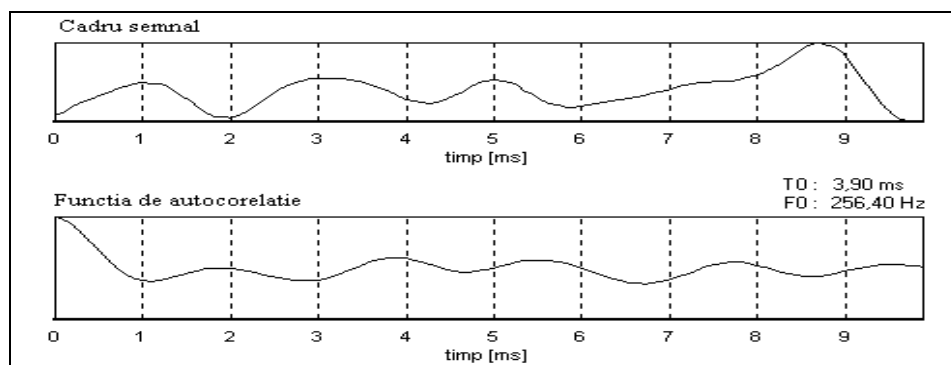
Pentru a evita efectele de ponderare introduse de alte ferestre (de exemplu fereastra Hamming), în acest caz se folosește o fereastră dreptunghiulară. Pentru un semnal periodic funcția de autocorelație prezintă o periodicitate cu aceeași perioadă ca cea a semnalului inițial .



Segment vocal și funcția de autocorelație asociată

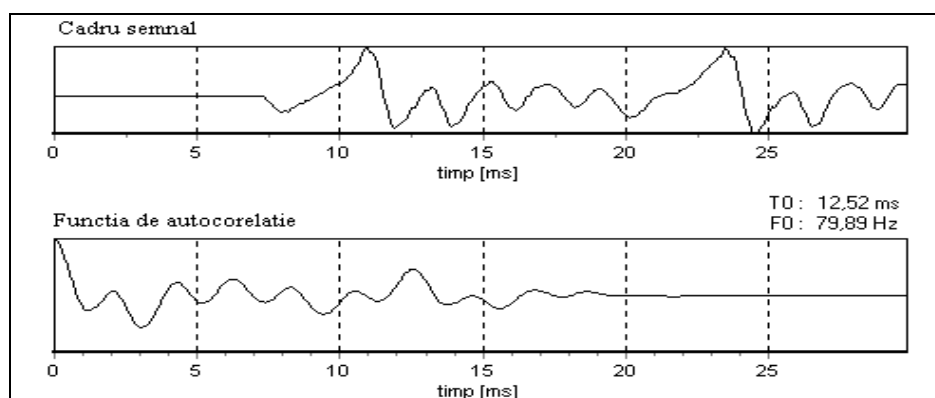
Funcția de autocorelație prezintă maxime distincte la intervale de timp cu lungimea $1/F_0$. Distanța dintre aceste maxime dă perioada fundamentală. În cazul funcției de autocorelație, aceste maxime sunt descrescătoare în amplitudine ceea ce ușurează detectarea acestora.

Lungimea ferestrei trebuie să îndeplinească anumite cerințe pentru a obține rezultate valide [Sto84][Mar91]. Pe de o parte lungimea ar trebui să fie mică datorită variabilității semnalului vocal, iar pe de altă parte, pentru a putea obține informația de periodicitate, este nevoie ca fereastra să cuprindă cel puțin două perioade ale formei de undă. Pentru valori mai mici pot apărea erori, mai ales în cazul vorbitorilor masculini. În acest caz maximele funcției de autocorelație nu sunt evidente și valorile determinate pentru F_0 și T_0 sunt eronate.



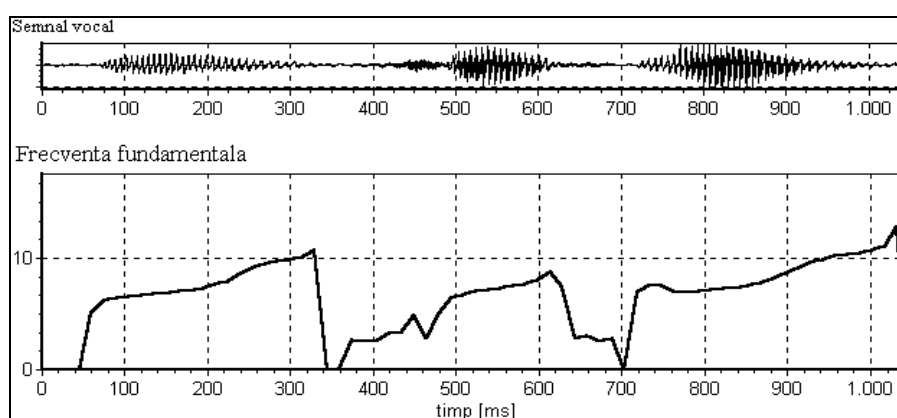
Segment vocal de 10 ms și funcția de autocorelație asociată

De asemenea apar rezultate deformate în zonele de trecere, la începutul segmentului vocal și la sfârșitul acestuia datorită informației insuficiente prezente în cadrul analizat .



Segment de trecere și funcția de autocorelație asociată

În urma aplicării acestei metode, pentru un semnal vocal, după eliminarea zonelor de liniște, s-a obținut conturul pentru frecvența fundamentală.



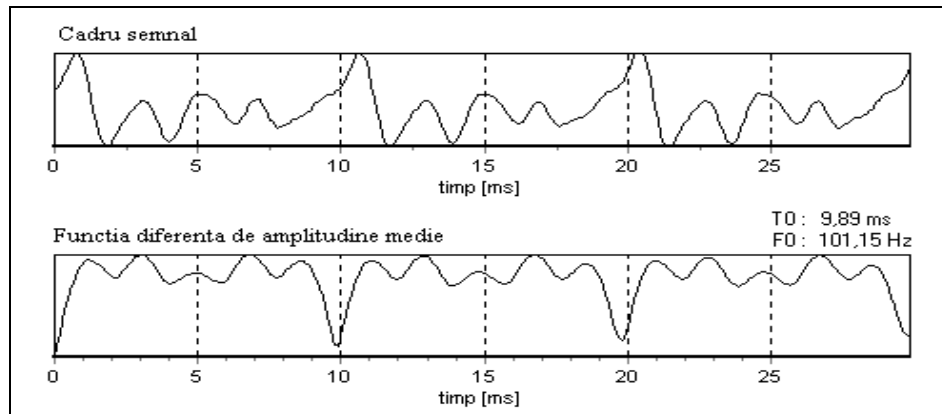
Semnal vocal și frecvența fundamentală asociată obținută prin metoda autocorelației

1.2. Metoda funcției diferență de amplitudine medie

Această metodă de determinare a frecvenței fundamentale se bazează pe funcția diferență de amplitudine medie (AMDF : Average Magnitude Difference Function), funcție ce are expresia [Sto84]:

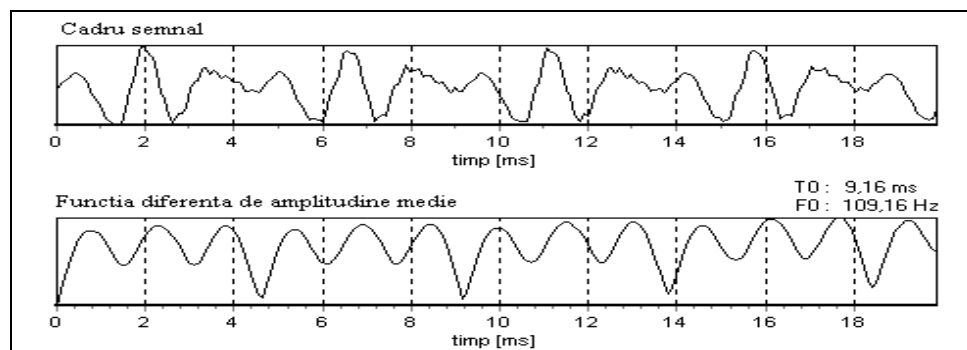
$$D(n) = \frac{1}{N} \sum_{k=0}^{N-1} |s_k - s_{k-n}| \quad , 0 \leq n \leq N-1.$$

Funcția $D(n)$ prezintă minime la intervale de timp de lungime $1/F_0$, minime care, în general, descresc în amplitudine.



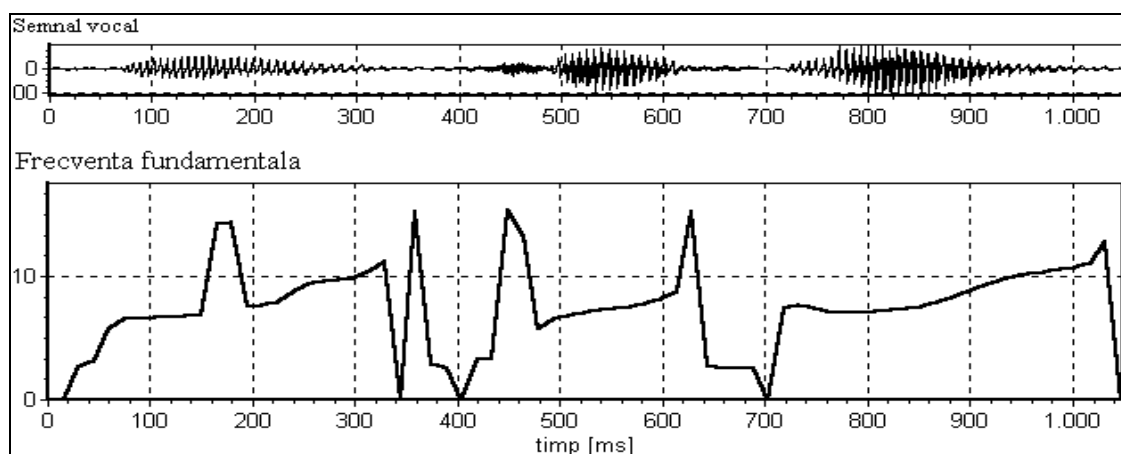
Funcția diferență de amplitudine medie (AMDF)

Apar însă situații când aceste minime nu sunt descrescătoare în amplitudine, ceea ce conduce la obținerea unor valori duble ale perioadei. Aceste situații apar în cazul când lungimea segmentului vocal luat în considerație nu este suficient de mare, chiar dacă acoperă 2-3 perioade de semnal.



Perioadă dublă obținută prin metoda AMDF

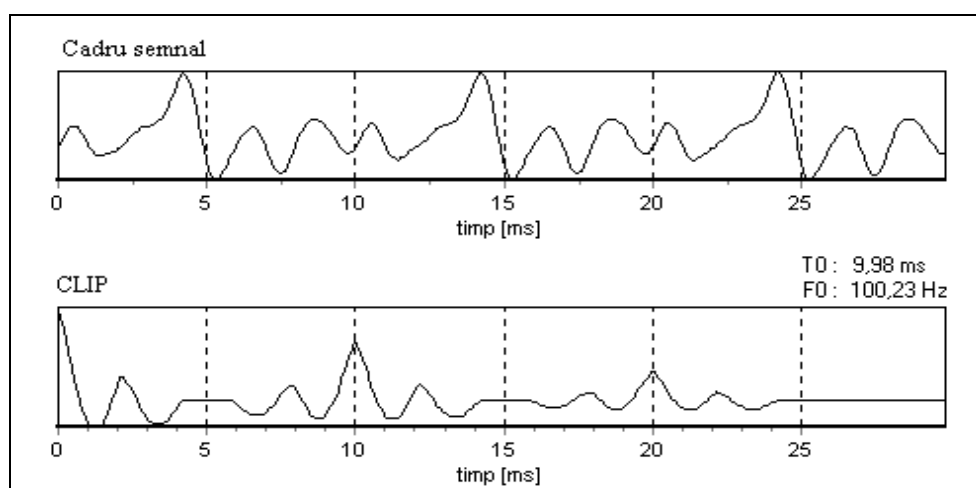
Și această metodă furnizează rezultate eronate în zonele de trecere, la începutul și sfârșitul unui segment vocal. Conturul pentru frecvența fundamentală este prezentat în fig.



Semnal vocal și frecvența fundamentală asociată obținută prin metoda AMDF

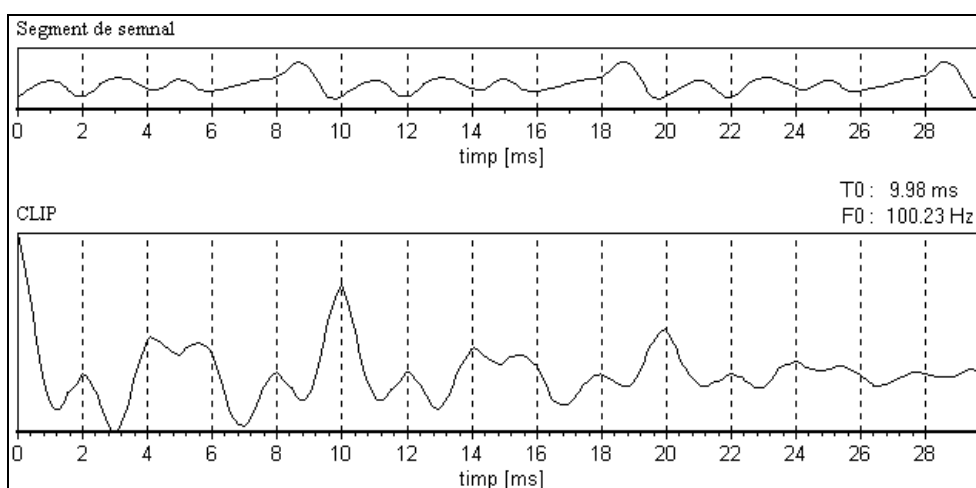
1.3. Metoda cu limitare centrală (CLIP : Center Clipping Pitch Detector)

Această metodă [Sto84] folosește un prag C_L și se rețin numai acele elemente din semnalul vocal a căror valoare absolută depășește pragul $|C_L|$. Pentru valorile rămase se procedează astfel : se adaugă valoarea C_L la cele pozitive, respectiv se scade valoarea C_L la cele negative sau se atribuie valoarea maximă pozitivă respectiv valoarea maximă negativă (limitare infinită). În continuare se calculează funcția de autocorelație. Aceasta prezintă maxime la intervale de timp direct corelate cu perioada asociată frecvenței fundamentale .

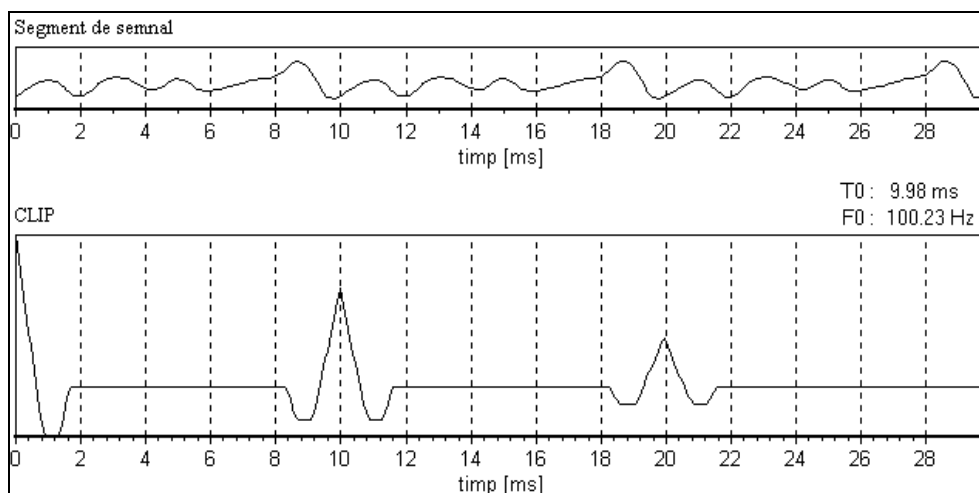


Determinarea F_0 cu metoda CLIP

Metoda este sensibilă la valoarea de prag folosită. Se poate folosi pentru C_L valoarea medie sau valoarea maximă a semnalului pe segmentul considerat, eventual înmulțită cu un coeficient subunitar. Scăderea pragului apropie forma funcției de cea de autocorelație obișnuită în timp ce creșterea acestuia simplifică forma finală a funcției .

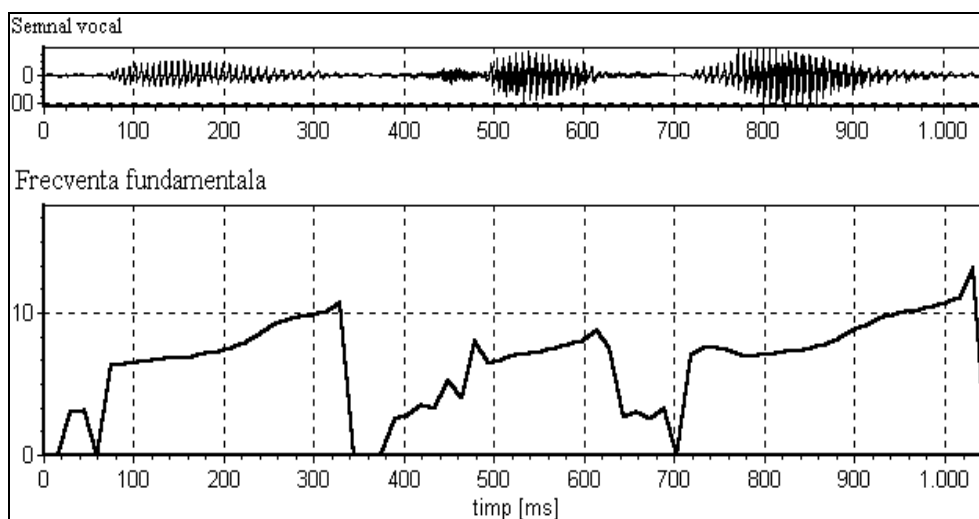


Determinarea F_0 cu metoda CLIP, prag 20%.

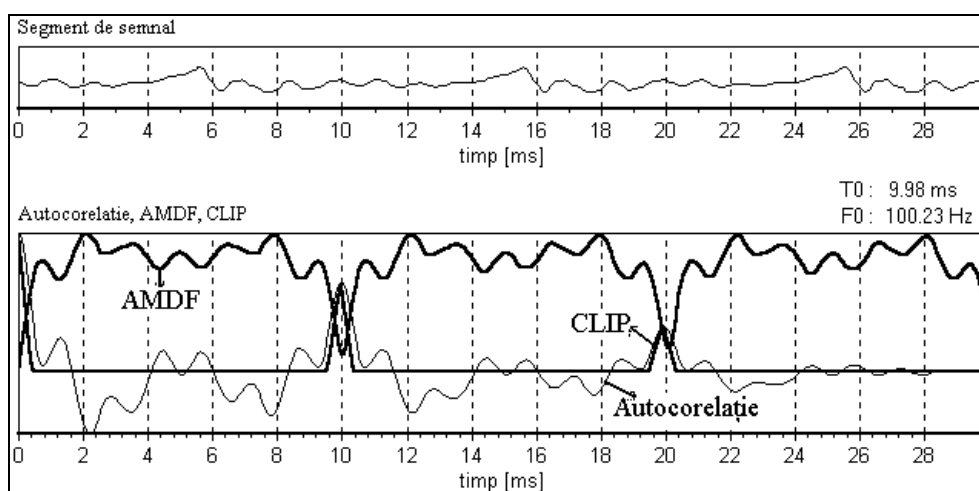


Determinarea F_0 cu metoda CLIP, prag 60%.

Conturul frecvenței fundamentale este prezentat în fig.



Semnal vocal și frecvența fundamentală asociată obținută prin metoda CLIP, prag 30%



Convergența metodelor de determinare a frecvenței fundamentale bazate pe autocorelație

2. TEMA DE IMPLEMENTAT

- Se vizualizeaza fisierele .wav atasate, cu aplicatia COLEA.
- Se alege un fisier de lucru si se decupeaza o sectiune suficient de lunga dintr-o zona sonora (in jurul segmentului de energie maxima)si se salveaza cu nume distinct.
- Se alege un fisier din cele anexate si se reprezinta conturul FF pentru intreg fisierul ales cu aplicatia COLEA.
- Se afiseaza cu aplicatia COLEA valorile FF pentru sectiunea aleasa prin metodele disponibile (autocorelatie) si se noteaza rezultatul+captura.
- Realizati implementarea proprie pe baza teoretica din lucrare, folosind una din metodele prezentate: AMDF, CLIP, ACC sau altele.
- Testati aplicatia proprie pe acelasi fisier/segment si comparati rezultatele obtinute cu cele din COLEA
- Faceti detectia sonor/nesonor si determinati conturul FF pentru fisierul ales (pitch tracking)

Raport SE PREDA RAPORTUL LUCRARII DE LABORATOR CU : IMPLEMENTARE (SCRIPT PERSONAL), REZULTATE COMPARATIVE, CAPTURI, COMENTARII.