

L3. Estimarea Frecvenței Fundamentale – Metoda Cepstrală

1. Analiza cepstrală.

O modalitate bună de a estima FF pentru semnale vocale staționare, fără zgomot și de durată suficientă este folosirea cepstrum-ului. Cepstrul se definește ca o prelucrare derivată din Transformata Fourier Rapida (FFT) fiind definit ca TFI a logaritmului amplitudinii spectrale a semnalului vocal.

Abscisa cepstrum-ului are ca unitate cufrența (s) și varfurile în cepstru (care sunt legate de periodicitatea în spectru) sunt numite ramonici.

Prelucrările cepstrale fac parte din clasa prelucrarilor homomorfe.

Metodele homomorfe de prelucrare a semnalului prezintă un interes deosebit în recunoașterea vorbirii. Sistemele homomorfe sunt o clasă de sisteme neliniare care respectă principiul superpoziției (sistemele liniare sunt un caz particular al sistemelor homomorfe). Prelucrarea vorbirii poate fi considerată ca o filtrare a excitației glotale de către tractul vocal.

În procesarea vorbirii un sistem homomorfic trebuie să aibă proprietatea :

$$D\left[\left[x_1(n)\right]^\alpha \cdot \left[x_2(n)\right]^\beta\right] = \alpha \cdot D\left[x_1(n)\right] + \beta \cdot D\left[x_2(n)\right] \quad (1)$$

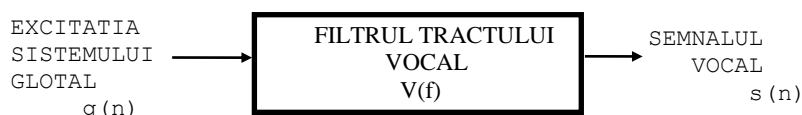


Fig. 1. Modelul acustic liniar de producere a vorbirii

Acesta este tipul de operație de superpoziție referitor la înmulțire, exponențială și adunare. O funcție logaritmă respectă proprietatea generalizată a superpoziției. Sistemele homomorfe sunt utile în prelucrarea semnalului deoarece permit separarea semnalului de excitație de forma tractului vocal. Procesul de separare a două componente se poate asimila cu o deconvoluție :

$$s(n) = g(n) \otimes v(n) \quad (2)$$

unde $g(n)$ - excitația

$v(n)$ - răspunsul tractului vocal la impuls

care în domeniul frecvență :

$$S(f) = G(f) \cdot V(f) \quad (3)$$

$$\log S(f) = \log [G(f) \cdot V(f)] = \log G(f) + \log V(f) \quad (4)$$

În domeniul logaritmă contribuția excitației și a tractului vocal sunt suprapuse și se pot separa, cel puțin în teorie, folosind metode convenționale de prelucrare a semnalului.

Pentru calculul cepstrumului se calculează mai întâi logaritmul amplitudinii spectrale, după care se calculează transformata Fourier inversă a logaritmului spectrului, figura 2.

$$c(n) = \frac{1}{N} \sum_{k=0}^{N-1} \log |S_{\text{med}}(k)| \cdot e^{(2\pi/N)kn}, \quad 0 \leq n \leq N-1 \quad (5)$$

unde $c(n)$ sunt coeficienții cepstrali.

Ne referim la coeficienții cepstrali calculați prin metoda anterior prezentată ca la coeficienții cepstrali derivați din transformata Fourier .

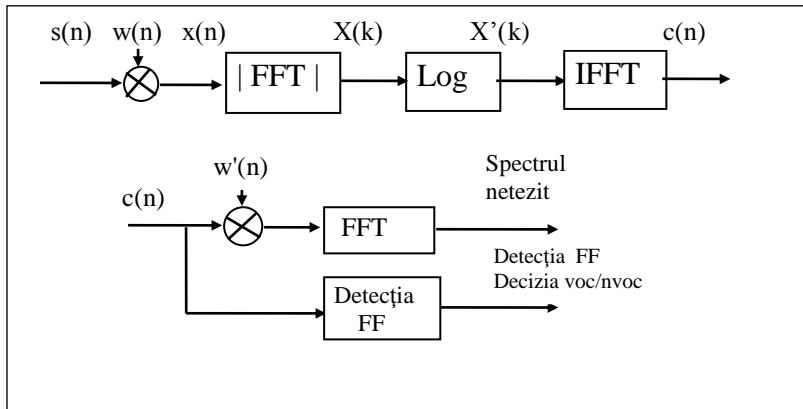


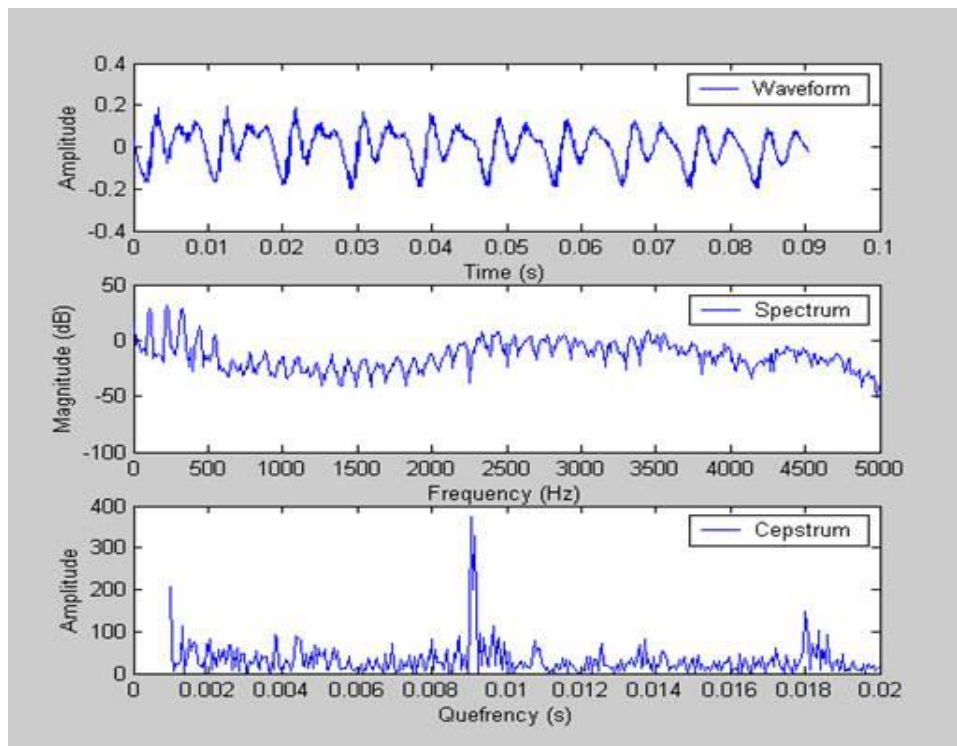
Fig. 2. Diagrama analizei cepstrale pentru extragerea anvelopei spectrale netezite si a frecvenței fundamentale.

Daca logarithmul amplitudinii spectrale contine mai multe armonici regulat distantate atunci analiza Fourier a acestui spectru va prezenta un varf corespunzator distantei dintre armonici adica FF. Spectrul semnalului este efectiv tratat ca un alt semnal cautand apoi periodicitate in el insusi.

2. Model pentru implementare

Pentru a obtine o estimare a FF din cepstrum se cauta un varf in domeniul cufrenta (s) corespunzator FF din s semnalului vocal. In continuare este prezentat un model.

```
%MODEL FF din CESPTRUM
% get a section of vowel
[x,fs]=wavread('six.wav',[24120 25930]);
ms1=fs/1000; % maximum speech Fx at 1000Hz
ms20=fs/50; % minimum speech Fx at 50Hz
%
% plot waveform
t=(0:length(x)-1)/fs; % times of sampling instants
subplot(3,1,1);
plot(t,x);
legend('Waveform');
xlabel('Time (s)');
ylabel('Amplitude');
%
% do fourier transform of windowed signal
Y=fft(x.*hamming(length(x)));
%
% plot spectrum of bottom 5000Hz
hz5000=5000*length(Y)/fs;
f=(0:hz5000)*fs/length(Y);
subplot(3,1,2);
plot(f,20*log10(abs(Y(1:length(f)))+eps));
legend('Spectrum');
xlabel('Frequency (Hz)');
ylabel('Magnitude (dB)');
%
% cepstrum is DFT of log spectrum
C=fft(log(abs(Y)+eps));
%
% plot between 1ms (=1000Hz) and 20ms (=50Hz)
q=(ms1:ms20)/fs;
subplot(3,1,3);
plot(q,abs(C(ms1:ms20)));
legend('Cepstrum');
xlabel('Quefrequency (s)');
ylabel('Amplitude');
```



Se cauta indexul varfului in domeniul cepstrului (cuefrenta) intre 1-20ms si apoi se converteste in Hertzi:

```
[c, fx]=max(abs(C(ms1:ms20)));
```

```
fprintf('Fx=%gHz\n', fs/(ms1+fx-1));
```

Metoda functioneaza cel mai bine cand FF nu se modifica prea mult, cand FF nu este prea mare si semnalul nu este zgomotos. Un dezavantaj important al metodei este faptul ca necesita resurse de prelucrare importante in domeniul frecventa.

Referinte: <http://mi.eng.cam.ac.uk/~ajr/SA95/node33.html>
<http://www.mathworks.com/>

3. DE IMPLEMENTAT !!!SE POATE TERMINA SI ACASA!!!

- Se decupeaza o sectiune suficient de lunga (>100ms) dintr-o vocala si se salveaza cu nume distinct (**pentru cei care nu au studiat PSV**)
- Se alege un fisier din cele anexate la L2 (**pentru cei care au studiat PSV**)
- Se afiseaza cu aplicatia COLEA valorile FF pentru sectiunea aleasa/fisierul ales prin metoda cepstrala si se salveaza rezultatul.
- Realizati implementarea proprie pe baza modelului
- Testati aplicatia pe acelasi fisier si comparati rezultatele . Comentarii.

SE PEDA RAPORT AL LUCRARI DE LABORATOR CU : IMPLEMENTARE (SURSA PERSONALA!!!), REZULTATE, CAPTURI, COMENTARII.

