



Contribuții la recunoașterea vorbirii și a vorbitorului

Margit Antal

Facultatea de Electronică, Telecomunicații și Tehnologia Informației

Cond.șt. Prof. Dr. Ing. Gavril Todorean



**Contribuții la
recunoașterea
vorbirii și a
vorbitorului**

Margit Antal

Bazele teoretice

**Contribuții la
recunoașterea
vorbitorului**

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

**Contribuții la
recunoașterea
vorbirii**

Clasificarea fonemelor

Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

**Direcții de
cercetare**

- 1 Bazele teoretice**
- 2 Contribuții la recunoașterea vorbitorului**
- 3 Contribuții la recunoașterea vorbirii**
- 4 Direcții de cercetare**

Contribuții la recunoașterea vorbirii și a vorbitorului

Margit Antal

Bazele teoretice

Contribuții la recunoașterea vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la recunoașterea vorbirii

Clasificarea fonemelor

Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

Dirjecții de cercetare

- Studiul modelelor de vorbitori din punct de vedere fonetic
- Ordonarea categoriilor fonetice pe baza puterii lor de discriminare a vorbitorilor
- Studiul modelelor de vorbitori structurate fonetic

Contribuții la recunoașterea vorbirii și a vorbitorului

Margit Antal

Bazele teoretice

Contribuții la recunoașterea vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la recunoașterea vorbirii

Clasificarea fonemelor

Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

Dirjecții de cercetare

- Studiul modelelor de vorbitori din punct de vedere fonetic
- Ordonarea categoriilor fonetice pe baza puterii lor de discriminare a vorbitorilor
- Studiul modelelor de vorbitori structurate fonetic

Contribuții la recunoașterea vorbirii și a vorbitorului

Margit Antal

Bazele teoretice

Contribuții la recunoașterea vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la recunoașterea vorbirii

Clasificarea fonemelor

Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

Dirjecții de cercetare

- Studiul modelelor de vorbitori din punct de vedere fonetic
- Ordonarea categoriilor fonetice pe baza puterii lor de discriminare a vorbitorilor
- Studiul modelelor de vorbitori structurate fonetic

Contribuții la recunoașterea vorbirii și a vorbitorului

Margit Antal

Bazele teoretice

Contribuții la recunoașterea vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la recunoașterea vorbirii

Clasificarea fonemelor

Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

Dirjecții de cercetare

- Studiul comportamentelor modelelor Markov ascunse fără probabilitățile de tranziție între stări
- Clasificarea și recunoașterea fonemelor cu modelul propus
- Studiul duratelor fonemelor și posibilitatea integrării acestora în algoritmul Viterbi

Contribuții la recunoașterea vorbirii și a vorbitorului

Margit Antal

Bazele teoretice

Contribuții la recunoașterea vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la recunoașterea vorbirii

Clasificarea fonemelor

Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

Dirjecții de cercetare

- Studiul comportamentelor modelelor Markov ascunse fără probabilitățile de tranziție între stări
- Clasificarea și recunoașterea fonemelor cu modelul propus
- Studiul duratelor fonemelor și posibilitatea integrării acestora în algoritmul Viterbi

Contribuții la recunoașterea vorbirii și a vorbitorului

Margit Antal

Bazele teoretice

Contribuții la recunoașterea vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la recunoașterea vorbirii

Clasificarea fonemelor

Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

Direcții de cercetare

- Studiul comportamentelor modelelor Markov ascunse fără probabilitățile de tranziție între stări
- Clasificarea și recunoașterea fonemelor cu modelul propus
- Studiul duratelor fonemelor și posibilitatea integrării acestora în algoritmul Viterbi

Contribuții la
recunoașterea
vorbirii și a
vorbitorului

Margit Antal

Bazele teoretice

Contribuții la
recunoașterea
vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la
recunoașterea
vorbirii

Clasificarea fonemelor

Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

**Dirjecții de
cercetare**

1. Elemente de fonetică
2. Extragerea caracteristicilor
3. Tehnici de grupare, cuantizare vectorială
4. Modele Markov ascunse

Contribuții la
recunoașterea
vorbirii și a
vorbitorului

Margit Antal

Bazele teoretice

Contribuții la
recunoașterea
vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la
recunoașterea
vorbirii

Clasificarea fonemelor

Recunoașterea
fonemelor

Modelarea duratei
fonemelor

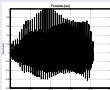
Recunoașterea
cuvintelor

Concluzii

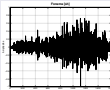
Dirjecții de
cercetare

Clase fonetice

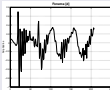
Vocale



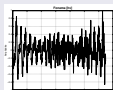
Fricative



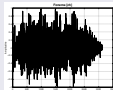
Plozive



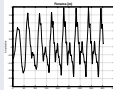
Semivocale



Africate



Nazale



- Preprocesare

$$s_p(n) = s(n) - a \cdot s(n-1) \quad (1)$$

- Analiză cepstrală

- FFT - Spectru de amplitudine

- Spectru de putere

- Spectru de energie

- Spectru de densitate de energie

- Postprocesare

$$\Delta C_m(t) = \frac{\sum_{k=-K}^K k C_m(t+k)}{\sum_{k=-K}^K k^2} \quad (2)$$

- Preprocesare

$$s_p(n) = s(n) - a \cdot s(n-1) \quad (1)$$

- Analiză cepstrală

- FFT - Spectru de amplitudine
- Filtre așezate conform cu scara Mel
- Logaritmare
- Transformată cosinus

- Postprocesare

$$\Delta C_m(t) = \frac{\sum_{k=-K}^K k C_m(t+k)}{\sum_{k=-K}^K k^2} \quad (2)$$

- Preprocesare

$$s_p(n) = s(n) - a \cdot s(n-1) \quad (1)$$

- Analiză cepstrală

- FFT - Spectru de amplitudine

- Filtre așezate conform cu scara Mel

- Logaritmare

- Transformată cosinus

- Postprocesare

$$\Delta C_m(t) = \frac{\sum_{k=-K}^K k C_m(t+k)}{\sum_{k=-K}^K k^2} \quad (2)$$

- Preprocesare

$$s_p(n) = s(n) - a \cdot s(n-1) \quad (1)$$

- Analiză cepstrală

- FFT - Spectru de amplitudine
- Filtre așezate conform cu scara Mel
- Logaritmare
- Transformată cosinus

- Postprocesare

$$\Delta C_m(t) = \frac{\sum_{k=-K}^K k C_m(t+k)}{\sum_{k=-K}^K k^2} \quad (2)$$

- Preprocesare

$$s_p(n) = s(n) - a \cdot s(n-1) \quad (1)$$

- Analiză cepstrală
 - FFT - Spectru de amplitudine
 - Filtre așezate conform cu scara Mel
 - Logaritmare
 - Transformată cosinus

- Postprocesare

$$\Delta C_m(t) = \frac{\sum_{k=-K}^K k C_m(t+k)}{\sum_{k=-K}^K k^2} \quad (2)$$

- Preprocesare

$$s_p(n) = s(n) - a \cdot s(n-1) \quad (1)$$

- Analiză cepstrală

- FFT - Spectru de amplitudine
- Filtre așezate conform cu scara Mel
- Logaritmare
- Transformată cosinus

- Postprocesare

$$\Delta C_m(t) = \frac{\sum_{k=-K}^K k C_m(t+k)}{\sum_{k=-K}^K k^2} \quad (2)$$

- Preprocesare

$$s_p(n) = s(n) - a \cdot s(n-1) \quad (1)$$

- Analiză cepstrală
 - FFT - Spectru de amplitudine
 - Filtre așezate conform cu scara Mel
 - Logaritmare
 - Transformată cosinus
- Postprocesare

$$\Delta C_m(t) = \frac{\sum_{k=-K}^K k C_m(t+k)}{\sum_{k=-K}^K k^2} \quad (2)$$

Contribuții la recunoașterea vorbirii și a vorbitorului

Margit Antal

Bazele teoretice

Contribuții la recunoașterea vorbitorului

Analiza fonetică a modelelor de vorbitori

Modele de vorbitori fonetic pure

Modele de vorbitori structurate fonetic

Concluzii

Contribuții la recunoașterea vorbirii

Clasificarea fonemelor

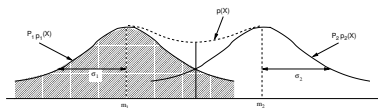
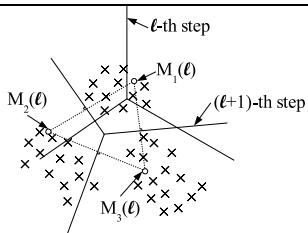
Recunoașterea fonemelor

Modelarea duratei fonemelor

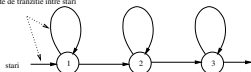
Recunoașterea cuvintelor

Concluzii

Dirjecții de cercetare



probabilitate de tranziție între stări



probabilitati de emisie

secvența de observații



Contribuții la
recunoașterea
vorbirii și a
vorbitorului

Margit Antal

Bazele teoretice

Contribuții la
recunoașterea
vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la
recunoașterea
vorbirii

Clasificarea fonemelor

Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

Dirjecții de
cercetare

	TIMIT	OASIS-NUMBERS
Creator	TI si MIT	Szeged University
Conditii	birou	birou
Parametri	16KHz, 16 biti	22.05KHz, 16 biti
Scop	Rec.vorbirii continue	Rec.numerelor
Nr. vorbitori	630	26
Continut	10 propozitii(30s)	26 de cuvinte
Nr. foneme	61	31

Contribuții la
recunoașterea
vorbirii și a
vorbitorului

Margit Antal

Bazele teoretice

Contribuții la
recunoașterea
vorbitorului

Analiza fonetică a
modelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la
recunoașterea
vorbirii

Clasificarea fonemelor

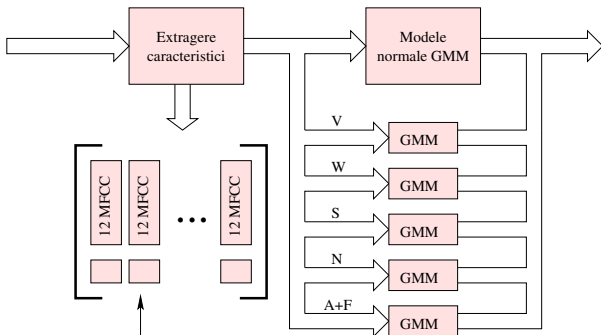
Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

Dirjecții de
cercetare



Grupa fonetica {V, W, S, N, A, F, O}

Contribuții la
recunoașterea
vorbirii și a
vorbitorului

Margit Antal

Bazele teoretice

Contribuții la
recunoașterea
vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la
recunoașterea
vorbirii

Clasificarea fonemelor

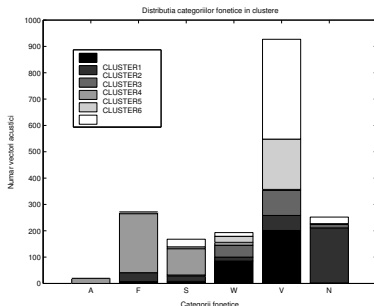
Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

Dirjecții de
cercetare



Contribuții la
recunoașterea
vorbirii și a
vorbitorului

Margit Antal

Bazele teoretice

Contribuții la
recunoașterea
vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la
recunoașterea
vorbirii

Clasificarea fonemelor

Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

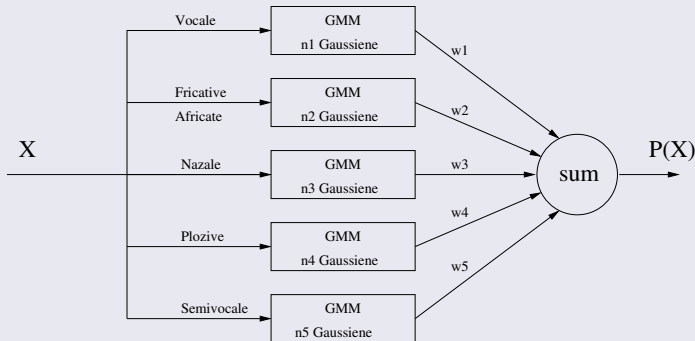
Concluzii

Dirjecții de
cercetare

TIMIT 630 de vorbitori

Categorie	Antrenare	Testare	Rata id.
Vocale	9.65s	2.27s	95.39%
Nazale	1.34s	0.38s	70.31%
Fricative	3.27s	0.95s	44.60%
Semivocale	2.38s	0.47s	41.74%
Plozive	1.43s	0.35s	10.47%

PSGMM - Calculul similarității



Contribuții la
recunoașterea
vorbirii și a
vorbitorului

Margit Antal

Bazele teoretice

Contribuții la
recunoașterea
vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la
recunoașterea
vorbirii

Clasificarea fonemelor

Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

Dirjecții de
cercetare

TIMIT 630 de vorbitori

Tip model	Nr.mixt.	Rata id.(3s)	Rata id.(6s)
GMM	6	93.02%	98.58%
PSGMM1	6	92.14%	98.74%
PSGMM2	6	92.06%	98.16%

Contribuții la
recunoașterea
vorbirii și a
vorbitorului

Margit Antal

Bazele teoretice

Contribuții la
recunoașterea
vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la
recunoașterea
vorbirii

Clasificarea fonemelor

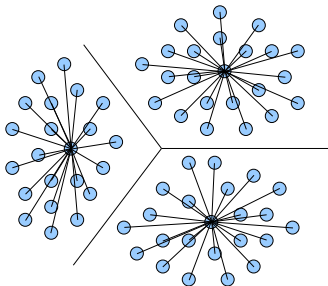
Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

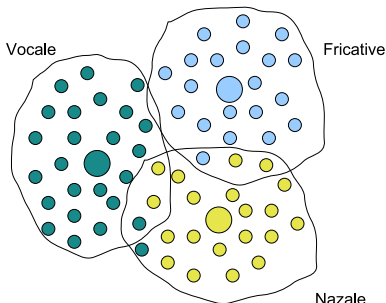
Concluzii

Dirjecții de
cercetare



a. Inițializare clasică

Eroare: 2,53%



b. Inițializare cu clasele fonetice

Eroare: 1,11%

Contribuții la recunoașterea vorbirii și a vorbitorului

Margit Antal

Bazele teoretice

Contribuții la recunoașterea vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la recunoașterea vorbirii

Clasificarea fonemelor

Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

Dirjecții de cercetare

- Componentele unui model de vorbitor nu sunt pure d.p.d.v. fonetic
- Nazalele și vocalele s-au dovedit a fi sunetele cele mai specifice vorbitorilor
- Modelele structurate fonetic (PSGMM), care utilizează doar informații relative la categoria fonemelor, au performanțe similare cu cele obișnuite (GMM)

Contribuții la recunoașterea vorbirii și a vorbitorului

Margit Antal

Bazele teoretice

Contribuții la recunoașterea vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la recunoașterea vorbirii

Clasificarea fonemelor

Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

Direcții de cercetare

- Componentele unui model de vorbitor nu sunt pure d.p.d.v. fonetic
- Nazalele și vocalele s-au dovedit a fi sunetele cele mai specifice vorbitorilor
- Modelele structurate fonetic (PSGMM), care utilizează doar informații relative la categoria fonemelor, au performanțe similare cu cele obișnuite (GMM)

Contribuții la recunoașterea vorbirii și a vorbitorului

Margit Antal

Bazele teoretice

Contribuții la recunoașterea vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la recunoașterea vorbirii

Clasificarea fonemelor

Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

Dirjecții de cercetare

- Componentele unui model de vorbitor nu sunt pure d.p.d.v. fonetic
- Nazalele și vocalele s-au dovedit a fi sunetele cele mai specifice vorbitorilor
- Modelele structurate fonetic (PSGMM), care utilizează doar informații relative la categoria fonemelor, au performanțe similare cu cele obișnuite (GMM)

Contribuții la recunoașterea vorbirii și a vorbitorului

Margit Antal

Bazele teoretice

Contribuții la recunoașterea vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la recunoașterea vorbirii

Clasificarea fonemelor

Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

Dirjecții de cercetare

- Clasificarea fonemelor
- Recunoașterea fonemelor
- Modelarea duratei fonemelor
- Recunoașterea cuvintelor

Contribuții la recunoașterea vorbirii și a vorbitorului

Margit Antal

Bazele teoretice

Contribuții la recunoașterea vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la recunoașterea vorbirii

Clasificarea fonemelor

Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

Direcții de cercetare

- Clasificarea fonemelor
- Recunoașterea fonemelor
- Modelarea duratei fonemelor
- Recunoașterea cuvintelor

Contribuții la recunoașterea vorbirii și a vorbitorului

Margit Antal

Bazele teoretice

Contribuții la recunoașterea vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la recunoașterea vorbirii

Clasificarea fonemelor

Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

Direcții de cercetare

- Clasificarea fonemelor
- Recunoașterea fonemelor
- Modelarea duratei fonemelor
- Recunoașterea cuvintelor

Contribuții la recunoașterea vorbirii și a vorbitorului

Margit Antal

Bazele teoretice

Contribuții la recunoașterea vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la recunoașterea vorbirii

Clasificarea fonemelor

Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

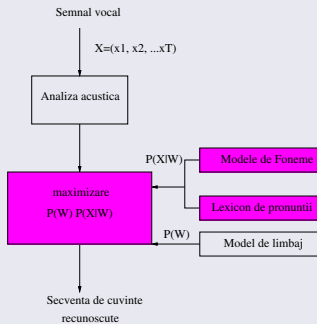
Concluzii

Direcții de cercetare

- Clasificarea fonemelor
- Recunoașterea fonemelor
- Modelarea duratei fonemelor
- Recunoașterea cuvintelor

Problema recunoașterii vorbirii

$$\hat{W} = \max_W P(W|X) = \max_W P(X|W) \cdot P(W) \quad (3)$$



Contribuții la
recunoașterea
vorbirii și a
vorbitorului

Margit Antal

Bazele teoretice

Contribuții la
recunoașterea
vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la
recunoașterea
vorbirii

Clasificarea fonemelor

Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

Dirjecții de
cercetare

Antrenare-Testare-Evaluare

Vectori acustici: 39MFCC, 20ms/10ms

$$Rata_{recunoastere} = \frac{\text{numar_foneme_corect_clasificate}}{\text{numar_total_foneme}} \quad (4)$$

TIMIT

Antrenare: 176.929 foneme, 462 vorbitori

Testare: 63.644 foneme, 168 vorbitori

OASIS Numbers

Antrenare: 7.461 foneme, 20 vorbitori

Testare: 2.244 foneme, 6 vorbitori

Contribuții la
recunoașterea
vorbirii și a
vorbitorului

Margit Antal

Bazele teoretice

Contribuții la
recunoașterea
vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la
recunoașterea
vorbirii

Clasificarea fonemelor

Recunoașterea
fonemelor

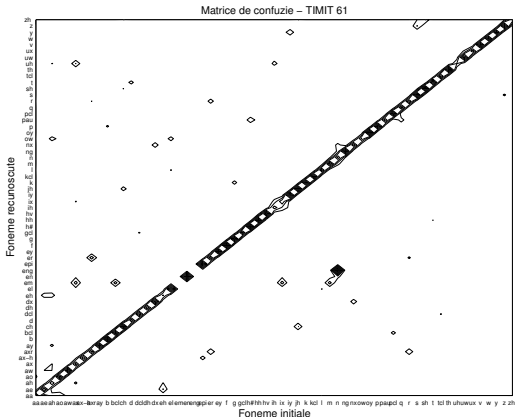
Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

Dirjecții de
cercetare

Rata clasificării: 63.34%



Contribuții la
recunoașterea
vorbirii și a
vorbitorului

Margit Antal

Bazele teoretice

Contribuții la
recunoașterea
vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la
recunoașterea
vorbirii

Clasificarea fonemelor

Recunoașterea
fonemelor

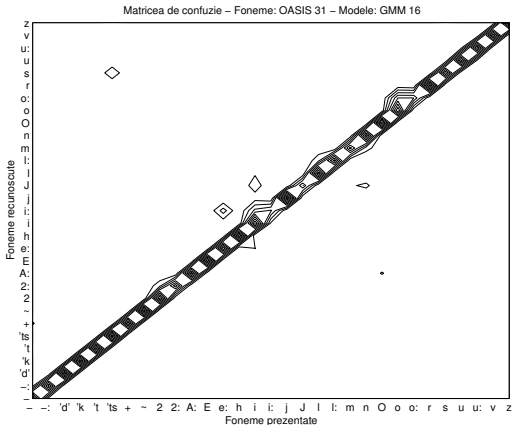
Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

Dirjecții de
cercetare

Rata clasificării: 92.74%



Contribuții la
recunoașterea
vorbirii și a
vorbitorului

Margit Antal

Bazele teoretice

Contribuții la
recunoașterea
vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la
recunoașterea
vorbirii

Clasificarea fonemelor

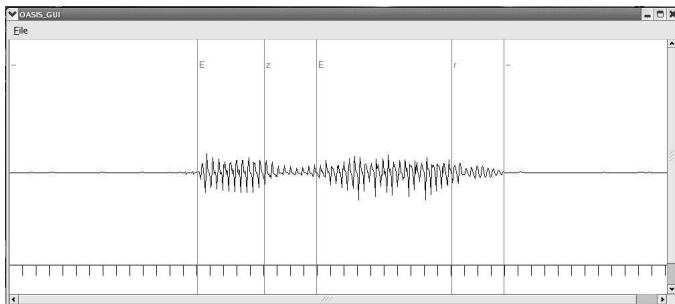
**Recunoașterea
fonemelor**

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

Dirjecții de
cercetare



Modificarea algoritmului Viterbi

$$\alpha_t(j) = \max_{1 \leq i \leq n} \alpha_{t-1}(i) \cdot a_{ij} \cdot b_j(o_t) \quad (5)$$

$$\log \alpha_t(j) = \max_{1 \leq i \leq n} \{ \log \alpha_{t-1}(i) + \log a_{ij} \} + \log b_j(o_t) \quad (6)$$

$$\log \alpha_t(j) = \max_{1 \leq i \leq n} \{ \log \alpha_{t-1}(i) + l_{ij} \} + \log b_j(o_t) \quad (7)$$

$$l_{ij} = \begin{cases} \beta, & \text{daca } i = j \\ 0, & \text{caz contrar} \end{cases} \quad (8)$$

Contribuții la
recunoașterea
vorbirii și a
vorbitorului

Margit Antal

Bazele teoretice

Contribuții la
recunoașterea
vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la
recunoașterea
vorbirii

Clasificarea fonemelor

**Recunoașterea
fonemelor**

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

Dirjecții de
cercetare

Metrica Levenshtein

$$PER = 100 \cdot \frac{INS + SUBST + DEL}{N} \quad (9)$$

$$ACCURACY = 100 - PER \quad (10)$$

Contribuții la
recunoașterea
vorbirii și a
vorbitorului

Margit Antal

Bazele teoretice

Contribuții la
recunoașterea
vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la
recunoașterea
vorbirii

Clasificarea fonemelor

**Recunoașterea
fonemelor**

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

Dirjecții de
cercetare

TIMIT-61

β	DEL	INS	SUBST	ACCURACY
11	6.79%	7.39%	34.54%	51.27%
13	9.13%	4.76%	33.12%	52.98%
15	11.42%	3.24%	31.66%	53.67%
20	17.16%	1.31%	28.12%	53.40%

Contribuții la
recunoașterea
vorbirii și a
vorbitorului

Margit Antal

Bazele teoretice

Contribuții la
recunoașterea
vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la
recunoașterea
vorbirii

Clasificarea fonemelor

**Recunoașterea
fonemelor**

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

Dirjecții de
cercetare

OASIS 31

β	DEL	INS	SUBST	ACCURACY
11	0.52%	27.93%	5.18%	66.36%
13	0.82%	21.80%	5.53%	71.85%
20	2.37%	11.27%	6.00%	80.35%
25	4.06%	7.77%	5.78%	82.38%
27	4.83%	7.08%	5.83%	82.25%
30	6.43%	5.82%	5.70%	82.03%

Contribuții la
recunoașterea
vorbirii și a
vorbitorului

Margit Antal

Bazele teoretice

Contribuții la
recunoașterea
vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la
recunoașterea
vorbirii

Clasificarea fonemelor

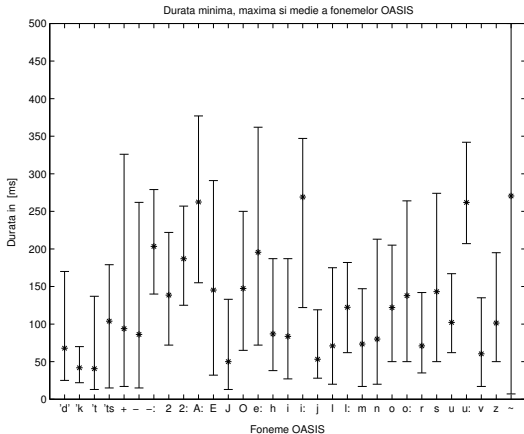
Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

Dirjecții de
cercetare



$$\alpha_t(j) = \max_{1 \leq i \leq n} \left\{ \max_{\tau_{min} \leq \tau \leq \tau_{max}} \left\{ \alpha_{t-\tau}(i) \cdot a_{ij} \cdot \prod_{\Theta=0}^{\tau-1} b_j(o_{t-\Theta}) \right\} \right\} \quad (11)$$

$$\alpha_t(j) = \max_{1 \leq i \leq n} \left\{ \max_{\tau_{min}(j) \leq \tau \leq \tau_{max}(j)} \left\{ \alpha_{t-\tau}(i) \cdot a_{ij} \cdot \prod_{\Theta=0}^{\tau-1} b_j(o_{t-\Theta}) \right\} \right\} \quad (12)$$

Contribuții la
recunoașterea
vorbirii și a
vorbitorului

Margit Antal

Bazele teoretice

Contribuții la
recunoașterea
vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la
recunoașterea
vorbirii

Clasificarea fonemelor

Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

Dirjecții de
cercetare

OASIS 31

Durate	ACC.	INS	SUBST	DEL
1..50	11.70%	85.53%	2.63%	0.12%
2..50	60.70%	36.22%	2.84%	0.21%
3..50	76.29%	18.39%	4.23%	1.07%
$\tau_{min} \dots \tau_{max}$	76.64%	16.32%	5.95%	1.07%

Contribuții la
recunoașterea
vorbirii și a
vorbitorului

Margit Antal

Bazele teoretice

Contribuții la
recunoașterea
vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la
recunoașterea
vorbirii

Clasificarea fonemelor

Recunoașterea
fonemelor

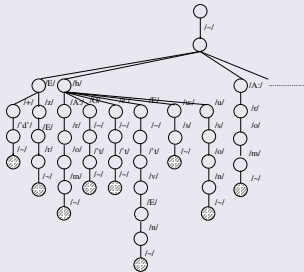
Modelarea duratei
fonemelor

**Recunoașterea
cuvintelor**

Concluzii

Dirjecții de
cercetare

Rezultate experimentale, OASIS 26 de cuvinte



REC. FONEME

REC. CUVINTE

84.36%

86.54%

Contribuții la recunoașterea vorbirii și a vorbitorului

Margit Antal

Bazele teoretice

Contribuții la recunoașterea vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la recunoașterea vorbirii

Clasificarea fonemelor

Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

Dirjecții de cercetare

- Rezultatele experimentale demonstrează că probabilitățile de tranziție în modelele Markov ascunse sunt ne semnificative comparativ cu probabilitățile de emisie ale stărilor
- Am adaptat algoritmul Viterbi pentru modele de foneme GMM
- Am propus o serie de modificări ai algoritmului Viterbi pentru integrarea cunoștințelor legate de durata fonemelor.

Contribuții la recunoașterea vorbirii și a vorbitorului

Margit Antal

Bazele teoretice

Contribuții la recunoașterea vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la recunoașterea vorbirii

Clasificarea fonemelor

Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

Dirjecții de cercetare

- Rezultatele experimentale demonstrează că probabilitățile de tranziție în modelele Markov ascunse sunt ne semnificative comparativ cu probabilitățile de emisie ale stărilor
- Am adaptat algoritmul Viterbi pentru modele de foneme GMM
- Am propus o serie de modificări ai algoritmului Viterbi pentru integrarea cunoștințelor legate de durata fonemelor.

Contribuții la recunoașterea vorbirii și a vorbitorului

Margit Antal

Bazele teoretice

Contribuții la recunoașterea vorbitorului

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

Contribuții la recunoașterea vorbirii

Clasificarea fonemelor

Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

Direcții de cercetare

- Rezultatele experimentale demonstrează că probabilitățile de tranziție în modelele Markov ascunse sunt ne semnificative comparativ cu probabilitățile de emisie ale stărilor
- Am adaptat algoritmul Viterbi pentru modele de foneme GMM
- Am propus o serie de modificări ai algoritmului Viterbi pentru integrarea cunoștințelor legate de durata fonemelor.

**Contribuții la
recunoașterea
vorbirii și a
vorbitorului**

Margit Antal

Bazele teoretice

**Contribuții la
recunoașterea
vorbitorului**

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

**Contribuții la
recunoașterea
vorbirii**

Clasificarea fonemelor

Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

**Direcții de
cercetare**

- **Recunoașterea vorbitorului**
 - Construirea modelelor de vorbitori structurate fonetic prin adaptarea modelelor generale de foneme
 - Verificarea identității pe baza unor informații lingvistice ca: durata unor foneme sau cuvinte specifice
- **Recunoașterea vorbirii**
 - Modelarea mai adecvată a duratei fără mărirea complexității calculelor



**Contribuții la
recunoașterea
vorbirii și a
vorbitorului**

Margit Antal

Bazele teoretice

**Contribuții la
recunoașterea
vorbitorului**

Analiza fonetică a
modelelor de vorbitori

Modele de vorbitori
fonetic pure

Modele de vorbitori
structurate fonetic

Concluzii

**Contribuții la
recunoașterea
vorbirii**

Clasificarea fonemelor

Recunoașterea
fonemelor

Modelarea duratei
fonemelor

Recunoașterea
cuvintelor

Concluzii

**Direcții de
cercetare**

Mulțumesc pentru atenție! Întrebări?