ON THE INFLUENCE OF INDIVIDUAL FEATURES COEFFICIENTS OVER SPEAKER RECOGNITION

Pop G. Petre, Lupu Eugen

Technical University, Cluj-Napoca
Gh. Baritiu str. 25-27,fax: 0040-64-191689, Romania, E-mail:Petre.Pop@com.utcluj.ro

Abstract : **Features based on short-term spectral estimate have a strong dependece on individual speakers and consequently are used in speaker recognition. However, these features also contain information about the lexical content of the utterance. We study the influence of eliminating some coefficients from speech features (LPC cepstrum, MFCC) on speaker recognition using VQ.**

I. INTRODUCTION

Representations used in speaker recognition concentrate primarly on properties of the speech signal atributtable to the shape of the vocal tract. Speaker identity is correlated with the physiological and behavioral characteristics of the speaker, both encoded in the spectral envelope and in the supra-segmental features (voice source characteristics and dynamic feature spanning over several segments).

Representations are almost always derived from the short time power spectrum, ignoring the phase structure, primarly because human ears are very insensitive to phase effects. However, these spectral parameters also contain information about the lexical content of the utterance. That's why some coefficients need to be removed and some need to be emphasize while other need to be deemphasized.

We studied the influence of eliminating some of the the first coefficients (individually or in sequence) from features (LPC cepstrum, MFCC) on speech derived spectrum and on speaker recognition using VQ.

II. THE INFLUENCE ON THE DERIVED SPECTRUM

Cepstral analysis

The LPC cepstrum provides a stabler representation from one repetition to another of a speaker's utterances. Aditionally, LPC cepstrum coefficients are uncorrelated which is a premise for good results in speaker recognition.

The effect of elimating the k_0'th coefficient over derived spectrum is shown in fig. 1.

First coefficient (k_0=0) has the most important influence on the derived spectrum. The effect of eliminating this coefficient is accentuated in high frequency domain. Next coefficient (k_0=1) has smaller influence over the derived spectrum. For k_0>1, the new spectrum is very closed to the original, especially in first half of frequency domain.

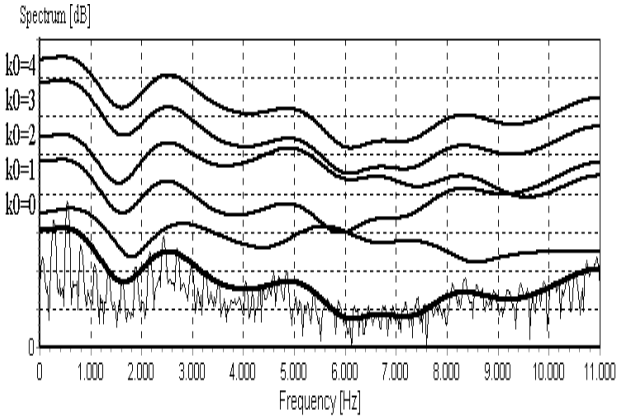

Figure 1. Power spectrum and LPC cepstrum derived spectrum for k_0=0,...,4.

The effect of elimating sequences of two coeffcients (seq=2) starting with the k_0'th coefficient over derived spectrum is shown in fig. 2.

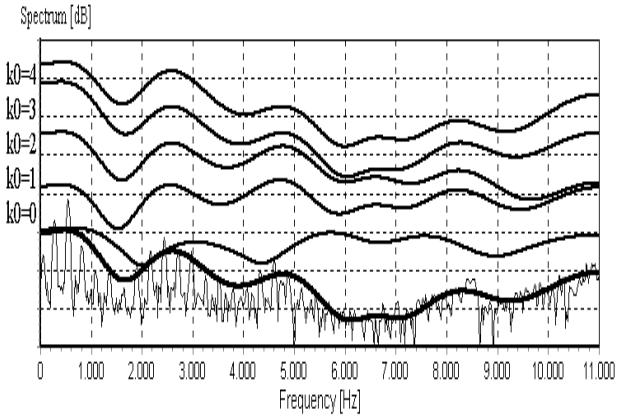

Figure 2. Power spectrum and LPC cepstrum derived spectrum for k_0=0,...,4, seq=2.

The first two coefficients, starting from k_0=0, alter the derived spectrum especially in high frequency domain. Next pairs of coefficients have smaller influence over the derived spectrum. For starting position k_0>1, the new spectrum is very closed to the original, especially in first half of frequency domain.
These experiments emphasize the idea that the first 4 coefficients are very important for the shape of the LPC cepstrum derived spectrum while high order coefficients seem to have very small influence over the derived spectrum.

Mel cepstral analysis

The MFCC derived spectrum is even more smoothed than LPC cepstrum derived spectrum.
The effect of eliminating the k_0'th coefficient over MFCC derived spectrum is shown in fig. 3.

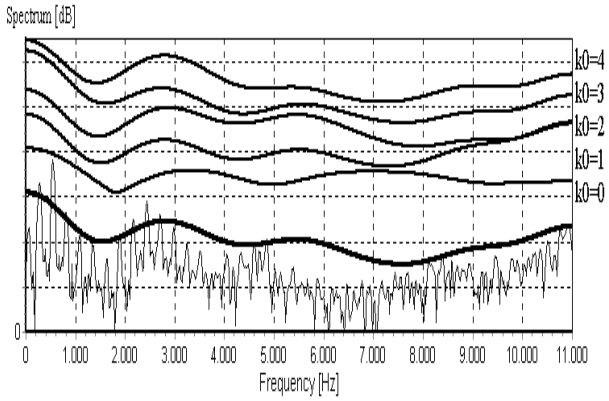

Figure 3. Power spectrum and MFCC derived spectrum for k_0=0,...,4.

As it can be seen, for k_0=0, the new spectrum is seriously damaged. The second coefficient alter the derived specrum at high frequencies while for k_0>3 the new spectra are very similar to the original spectrum.
Figure 4 presents the effect of elimating sequences of two coeffcients (seq=2) starting with the k_0'th coefficient over derived spectrum.

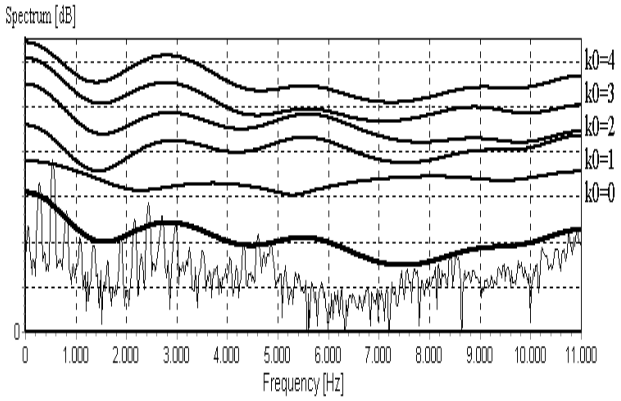

Figure 4. Power spectrum and MFCC derived spectrum for k_0=0,...,4, seq=2.

Again, the first two coefficients, starting from k_0=0, alter the derived spectrum much more than other pairs of coefficients from the features vector.
In MFCC case, the experiments shows that the first 2 coefficients are very important for the shape of the derived spectrum.

III. THE INFLUENCE ON THE SPEAKER RECOGNITION

We realized a set of experiments on speaker recognition using VQ. The data set consist of fixed utterances from 25 speakers, 15 males and 10 females. For each type of parameters (LPC cepstrum, MFCC), the position k_0 of eliminated coefficients was varied between 0 and 10 and the sequence length variation domain was between 1 and 4.
We used 16 LPC cepstral coefficients and 13 MFCC coefficients.

Figures 5 presents the variation of ERR (Error Recognition Rate) with starting position of eliminated coefficients (k_0) and with sequence length (seq), for LPC cepstral coefficients.

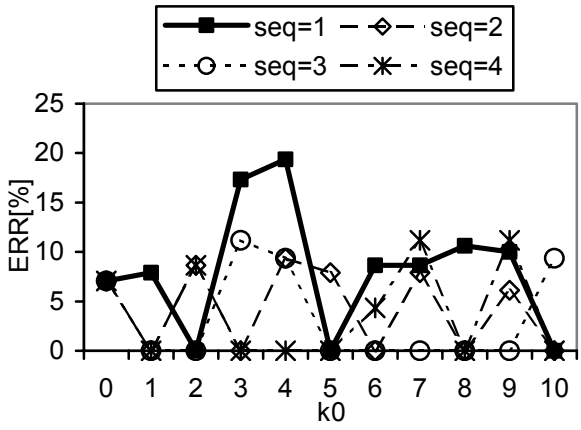

Figure 5. ERR(k_0) values using LPC cepstrum features

We found that coefficients c_2 and c_5, as individuals, have negative influence on recognition.
Sequences of two coefficients starting with k_0=1 or k_0=3 may be removed in order to improve the recognition performances.
If sequences of three coefficients starting with position between 5 and 9 are removed, the recognition performances are very good.

Figures 6 shows the variation of ERR (Error Recognition Rate) with starting position of eliminated coefficients (k_0) and with sequence length (seq), for MFCC coefficients.
We found that coefficients c_0 and c_6, c_7 as individuals, have negative influence on speaker recognition.
By eliminating sequences of two coefficients no improvements are obtain in the recognition

performances.Yet, c_1, c_2, c_3 and c_6, c_7 have smaller speaker specific information than other MFCC coefficients.

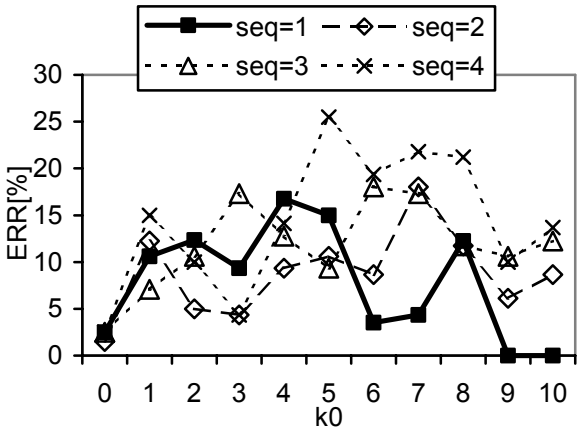

Figure 6. ERR(k_0) values using MFCC cepstrum features

IV. CONCLUSIONS

This paper try to figure out the influence of some coefficients from the parameter set (LPC cepstrum, MFCC) used in speaker recognition over the shape of the derived spectrum and finally over the process of recognition.

In both cases, LPC cepstrum and MFCC, the first coefficients have a significant influence over the spectrum shape. Removing the subsequent coefficients still preserve the spectrum shape.

A set of experiments on speaker recognition was realized using VQ in which we try to figure out the influence of the features coefficients (individually or in group) on the recognition results.

In LPC cepstrum case we found that coefficients c_2, c_5 have negative influence on recognition.

Sequences of two coefficients (starting with k_0=1 or k_0=3) or three coefficients (starting with position between 5 and 9) may be removed in order to improve the recognition performances.

In MFCC case, coefficients c_0, c_6, c_7, have negative influence on speaker recognition. We achieve no improvements by eliminating sequences. But, the experiments shows that c_1, c_2, c_3 and c_6, c_7 have smaller speaker specific information than other MFCC coefficients.

References

[1] Rabiner, L., Juang, B.H. [1993] “Funfamentals of Speech Recognition”, Prentice Hall.

[2] Huang, X., Acero, A., Hon, H. [2001] “Spoken Language Processing”, Prentice Hall .

[3] H. Hermansky [1999] "Mel cepstrum, deltas,double-deltas,.. - What else is new?", in *Proceedings of Robust Methods for Speech Recognition in Adverse Conditions, 1999, Tampere, Finland*.

[4] Picone, J.,W. [1993] "Signal modeling techniques in speech recognition", *Proc.IEEE* vol. 81 sept. 1993, pp.1215-1246.