European Conference on
Speech Technology
Edinburgh, Scotland, UK
September 1987

NEW QUANTIZATION METHOD FOR THE TRANSMISSION OF THE SPEECH LPC PARAMETERS

O. Abdel Alim*, M. Nagi* and H. El-Asaal**

ABSTRACT

In this paper we apply a new method for quantizing the LPC parameters obtained from the analysis of different speech data. This new method uses the redundancy between the LPC coefficients of successive frames to get additional sub-quantization levels suitable for each coefficient group. This method is applied for english as well as for arabic sounds. The data under test comprises mono- and disyllabic sounds. The results show that the bit rate could be minimized in the case of maintaining the same error as that of the traditional method. If the bit rate is maintained the same, the error will be reduced.

INTRODUCTION

In LPC analysis the speech signal is divided into frames each is represented by a vector of estimated vocal tract parameters, representing the spectral information of speech, assumed to be constant throughout the frame interval.

For many sounds these parameters do not change significantly from one frame to the next /1/. As the analysis frame is made short enough to represent quickly varying speech, such as stop consonants, then the slowly varying sounds are several frames long, and the information representing this frames show very small variation.

Moreover, for voiced speech it has been observed that k1 (the first reflection coefficient of each frame) has a skewed distribution near (-1), while the second coefficient k2 is skewed near (+1). The other k-parameters are usually bounded by (0.7) in magnitude and more evenly distributed near zero /2/.

For continuous speech, the k-parameters do not change very fast from one frame to the next and some of them can often be represented enough by the parameters of the previous frame /3/.

The redundancy which exist between successive frames is reflected in the nature of the parameters' vectors representing successive frames.

A new method is introduced to use the redundancy which exists between successive frames to reduce the number of bits used to code reflection coefficients, or to improve the speech quality if the number of bits is kept constant.

* Faculty of Engineering, Alex. University, Alexandria,Egypt
** Siemens Scientific Office, Alexandria,Egypt

10.21437/ECST.1987-173

THEORY

Let us assume that we have N successive frames, at first they are analyzed to obtain the k-parameters representing the spectral information of each frame using the auto-correlation method.
If the analysis filter has M-delay sections, then we have M*N reflection coefficients.
If a matrix is composed out of M*N coefficients with its rows are the successive frames, thus the element k(m,n) is the mth coefficient of the nth frame in that block of frames.
For each column we have elements with values close to each other and concentrated near a local mean value (for most speech sounds).
The max. and min. of each column are quantized and used to generate sub-quantization levels which cover the column element range.
To transmit the elements of each column, the quantized max. and min. is sent and then the quantized element values. At the receiver, the max. and min. are used to generate back the sub-quantization levels for that column.

SAMPLES

In /4/ the suggested method has been applied on english sounds such as the word "cattle" and the utterances "əna" and "əzha ".
Samples obtained from 7 different arabic words (mono- and disyllabic) have been analyzed. Waveforms of some of these words are shown in figure (1). The number of samples per frame is 200, sample frequency is 10 kHz and the order of analysis filter (M) is 8.
The analysis is performed using Markel's method and the newly suggested method under the following two conditions:
1- The total square error (TSE) is held constant, as the error obtained in Makrel's method, and the number of bits is calculated.
2- The same number of bits is held constant, as obtained in Makrel's method, and the mean square error (EMS) is calculated.

RESULTS

The results are shown in table (1) and (2). In Table (1) the error is held approximately constant for one group and for more than one group per word and the bit reduction is given for each case. In table (2) the number of bits is held approximately constant and the EMS reduction is given.

DISCUSSION AND CONCLUSION

Table (1) shows that the new method reduces at general the number of bits for the same error. It shows also that for words with fricatives, it is more useful to divide the word into more than one group to gather each part of the word for

Table (1)

arabic latin	N	Makrel's method no. bits	TSE *	sugg. method one group no. bits	TSE *	bit red%	suggested method with group division group	no. bits	TSE *	bit red%
bab باب	38	1520	51	1301	48	16.8	1-19 20-38	1276	56	19.0
shab شاب	44	1760	57	1671	45	5.3	1-22 23-44	1567	55	12.3
khab خاب	43	1720	68	1550	64	10.9	1-22 23-43	1514	55	13.6
sab ساب	42	1680	56	1515	58	10.8	1-21 22-42	1480	53	13.5
ssab صاب	44	1760	73	1542	75	14.1	1-22 23-44	1504	60	17.0
rehab رحاب	48	1920	62	1725	68	11.3	1-24 25-48	1655	64	16.0
asbab أسباب	55	2200	174	1808	159	21.6	1-20 21-40 41-55	1739	156	26.0

Table (2)

arabic latin	N	Makrel's method no. bits	EMS *	New method no. bits	EMS *	EMS red%
bab باب	38	1520	51	1486	15	230
shab شاب	44	1760	57	1714	38	50
khab خاب	43	1720	68	1718	31	123
sab ساب	42	1680	56	1638	30	88
ssab صاب	44	1760	73	1671	48	53
rehab رحاب	48	1920	62	1940	27	123
asbab أسباب	55	2200	174	2186	43	301

* These values are multiplied by 1000.

obtaining similar coefficients in each group.
From table (2) we can see that if the number of bits is held constant, the speech quality will be improved considerably. However, the optimal case could be obtained when we reduce the number of bits under the constraint of a suitable error.

REFERENCES

/1/ Markel, J.D. and Gray, A.H., 1976, " Linear prediction of speech ", Springer Verlag, New York.

/2/ Markel, J.D. and Gray, A.H., 1974, IEEE ASSP, 22 124-134.

/3/ Panagiotis, E. et al, 1983, IEEE ASSP, 31, 706-713

/4/ O. Abdel Alim and H. El-Asaal, "Minimization of coded speech information in analysis/synthesis methods", IEE International conference on speech input/output, conference Publication No. 258, p 266-269

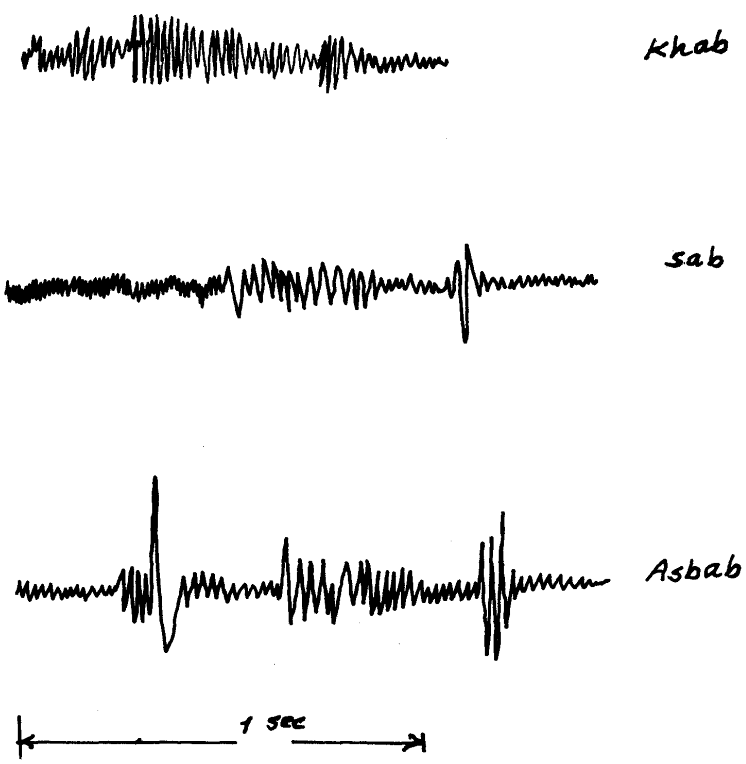

Fig.(1) : Waveforms of some arabic sounds