THE RESEARCH FOR THE MELP VOCODER AND ITS REAL-TIME REPLEMENTATION

Jun Wang[1], Jiyin Zhao[1], Jingqiu Yang[2], Yusheng Yang[1]
[1]College of Communication Engineering, Jilin University, Changchun, P. R. China
[2]Changchun Techuni Testing Corporation, P.R.China

Abstract: An improved mixed excitation linear prediction speech coder is proposed based on the analysis of the standard MELP speech coding algorithm. Bit rate of LPC coefficient is reduced by the similarity of LPC coefficient. at the same time, the primary four-stage vector quantization is substituted by the three-stage vector quantizer., and solve the contradiction of rates of quantization and quantization quality well. depress bit-rate further, reduce the reserves of codebook and calculation complexity. By these two methods the bit rates of MELP coder can be reduced to as low as 1.8kbps without evident damage to the synthetic speech quality. On the base of the algorithm simulation, the realization of the algorithm with C and assembly language, and on DSK hardware develop platform is performed. Both simulation and experiment results show that the synthesis speech by this new algorithm is equal to the 4800bps Federal Standard FS 1016 (CELP).

Key words: mixed excitation linear prediction, vector quantization, synthetic speech

Introduction: In mobile, satellite and military communications systems, the technology of speech coding plays an important role in increasing the availability of the channel by compressing transmission bandwidth and reducing transmission bit-rate of the speech signal. In recent years, the technology of speech coding advances rapidly. With the development of the signal processing and communication technology, the focus of speech coding research is centralized on the study and realization of low and very low bite rate speech coding algorithms.

The MELP is a perfect algorithm in current low bit rate speech coding. MELP coder has been adopted as the new US Federal Standard at 2.4kbps. the algorithm combine the merit of LPC and MBE algorithm. Several careful research and many experiments on the aspects of speech analysis, parameter code/encode and speech synthesis have been carried out. some new methods and ameliorations are employed in pitch detection, vector quantization and transmittion of LPC parameters. An improved 1.8 kbps MELP coding algorithm is proposed in this paper.

MELP Low Bit Rate Speech Coding Algorithm: The MELP coder is based on the LPC model with additional features including mixed excitation, aperiodic pulses, adaptive spectral enhancement, pulse dispersion filtering, and Fourier magnitude modeling[1][2]. These additional parameters largely amend the excitation structure of the LPC model, at the same time eliminate the mechanical tone noise that come forth in LPC speech synthesize. These features allow the MELP vocoder to simulate accurately natural speech.Fig.1 is MELP decoder block Diagram.

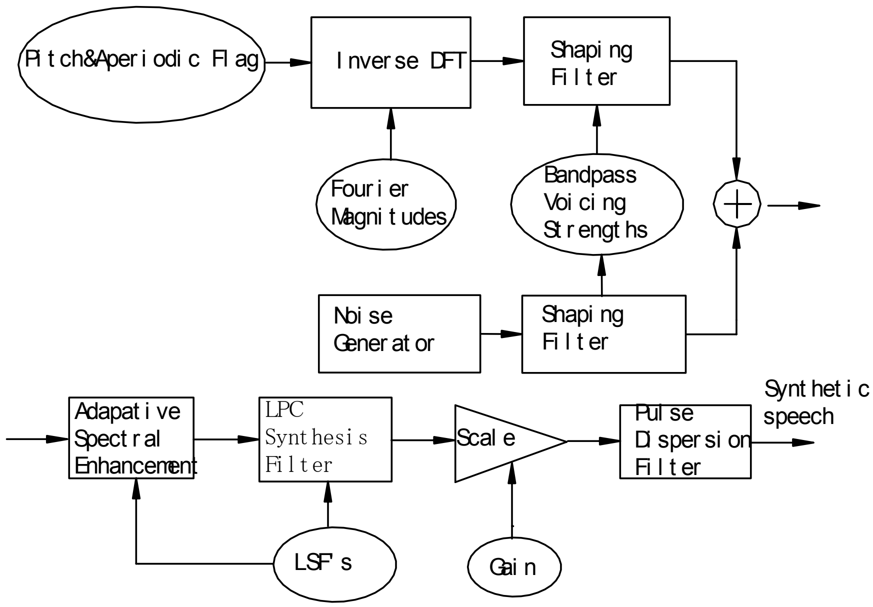

Figure 1. MELP Decoder Block Diagram

Experiments: Improved LSF Vector Quantization: The MELP coder is a parameter coder algorithm based on linear prediction analysis. It's parameters are transmitted and analysed frame-by-frame. However, it is weak that the characteristic of the track answer slow change isn't considered in the course of speech forming. So that, the similarity of the near speech frame is presented to reduce bit rate. If the similarity of the current frame and the previous frame exceed a value, the LSF parameters aren't transmitted and substituted by the previous frame LSF parameters.

Define a previous frame LSF parameters: $x(n), n = 1,2\Lambda\ 10$, Define the current frame LSF parameters $y(n), n = 1,2\Lambda\ 10$, Based on the cross-correlation theory, the cross-correlation function of $x(n)$ and $y(n)$ is:

$$R(x,y) = \sum_{i=1}^{10} c_i x(n) y(n) \tag{1}$$

c_i is the weighted coefficient, it is defined as:

$$c_i = \begin{cases} 1, & 1 \le i \le 7 \\ 0.8, & i = 8,9 \\ 0.4, & i = 10 \end{cases}$$

the normalized cross-correlation function is:

$$r(x,y) = \frac{R(x,y)}{\sqrt{R(x,x)R(y,y)}} \tag{2}$$

it is very obvious that $r(x,y)$ is larger also when the similarity of the two near LSF parameters become similar. Before $r(x,y)$ is used to judge, we need choosing a appropriate threshold r_0, when $r(x,y) \ge r_0$, the LSF parameters of the current frame aren't transmitted. the bit rate of speech coder will be reduced largely.

Improved LSF Vector Quantization:The standard MELP algorithm adopts the four-stage vector quantizer[3], The hunting route is 8. The method save effectively memory space of codebook vector, consider computer complexity and optimization of the quantiation vector. MSVQ[4] make satisfied result in standard MELP algorithm. But in the four-stage vector quantization of standard MELP algorithm, the capacity of the codebook is too large and the compensation of the fourth code-vector is still great. In this paper, the different vector quantization rank and bit assignment is obtained.

Considering the above analysis, the three-stage vector quantization is adopted and the hunting route is 16 in this paper, the bit assignment is 8,6,6 bits. firstly 16 vectors in the first codebook vectors that have minimum error with the input line spectrum are searched. Secondly, for this 16 vectors, 16 vectors in the second codebook vectors that have minimum error with the input line spectrum are searched respectively. Then 16*16=256 vectors are searched completely respectively in the three codebook vector, and calculate the distance between 256*64=16384 vectors and the input vector. The vector with minimum distance is the final quantization vector. The index of the three-stage codebook is the transmission codeword. Compare with four-stage vector quantization of the standard MELP algorithm, the three-stage vector quantization save effectively the space of the codebook vector and reduce the calculation complexity. The quantization bits of LSF parameters is descended to 20 bits from 25 bits[5][6]. It can be seen from subjective tests that the synthesis speech preserves the character of original speech.

Results: A great deal speech signals are analysed and find that it is very fit that r_0 is equal to 0.75. The LSF parameters of about 500 frame speech signals aren't transmitted and substituted by those of a previous frame speech signal in the 1000 frame speech signals. From Fig.2, The differentiation is very littler between the original speech signal and the synthetic speech signal.

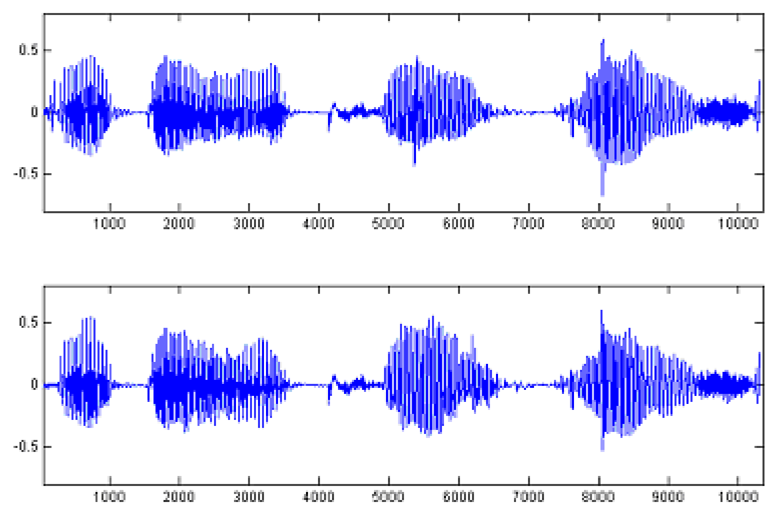

Figure 2. Comparison of two kinds of speech signals

An improved MELP Speech coding is proposed in this paper. Bit rate of LPC coefficient is reduced by the similarity of LPC coefficient. at the same time, the primary four-stage vector quantization is substituted by the three-stage quantization., and solve the contradiction of rates of quantization and quantization quality well. depress bit-rate further, reduce the reserves of codebook and calculation complexity. The improved MELP speech coding bit assignment is obtained in table 1. If $r_0 > 0.75$, LSF parameters of the current frame speech signal aren't transmitted and substituted by those of a previous frame speech signal. The table 1 is the bit allocation of LSF parameters.

The aims of a low bit rate speech coding algorithm is to attain high efficiency and concise algorithm on the basis of high quality of synthesize speech on the one hand, and on other hand is implemented in real time by DSP. In this paper, using TMS320C5416 DSK as hardware development platform and C and assembly language as software tools, the realization of the algorithm has been achieved with c language. It can seen from the Fig.3 that the synthesis speech with matlab and c language preserves the character of original speech.

Table 1. Bit allocation

Parameter		unvoice	Improved MELP($r_0 = 0.75$)			
			$r \geq r_0$		$r < r_0$	
			voice	unvoice	voice	unvoice
LSF's	25	25	20	20	2	2
fs_mag	8		8		8	
Gain	8	8	8	8	8	8
pitch	7	7	7	7	7	7
bpvc	4		4		4	
jitter	1		1		1	
Error Protection		13		13		13
syn	1	1	1	1	1	1
sum	54	54	49	49	31	31

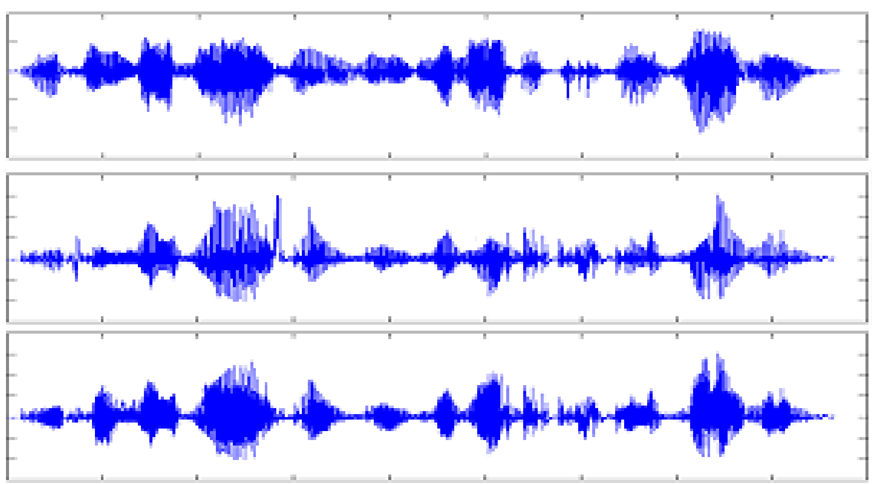

Figure 3. Simulation results

Discussion: An improved mixed excitation linear prediction speech coder is proposed based on the analysis of the standard MELP speech coding algorithm. Bit rate of LPC coefficient is reduced by the similarity of LPC coefficient. at the same time, the primary four-stage vector quantzation is substituted by the three-stage vector quantization, and solve the contradiction of rates of quantization and quantization quality well. depress bit-rate further, reduce the reserves of codebook and calculation complexity. By these two methods the bit rates of MELP coder can be reduced to as low as 1.8kbps without evident damage to the synthetic speech quality. It can be seen from the simulation result that the synthesis speech preserves the character of original speech. Furthermore, subjective tests show that the score of MOS can reach 3.5. On the base of the algorithm simulation, the realization of the algorithm with C and assembly language, and on DSK hardware develop platform is performed.

Conclusions: The bit rates of MELP coder can be reduced to as low as 1.8kbps without evident damage to the synthetic speech quality by the similarity of LPC coefficient and the three-stage vector quantization. LSF coefficient. Both simulation and experiment results show that the synthesis speech by this new algorithm is equal to the 4800bps Federal Standard FS 1016 (CELP).

References:

1. A. McCree, K.Truong, E.B.Gerorge,T.P.Barnwell, and V.Viswanathan. A 2.4kb/s MELP Coder candidate for the NEW U.S Federl Standard, Proceeding of IEEE ICASSP 1996:200-203
2. Takahiro Unno, Thomas P. Barnwell III, and Kwan Truong, An Inproved Mixed Excitation Linear Prediction (MELP) coder, Proceeding of IEEE ICASSP 1999:245-248
3. W. P. I. EBlanc, B. Bhattacharya, S.A.Mahmoud, and V.Cuperman. Efficient Search and Design Procedures for Robust Multi-Stage VQ of LPC Parameters for 4kbps Speech Coding, IEEE Transaction on Speech and Audio Processing, Vol.1,No.4,1993:373-385
4. K. K. Paliwal and B.S. Atal, "Efficient vector quantization of LPC parameters at 24bits/frame", IEEE Trans.Speech Audio Proceeding, Vol 1,No 1,Jan.1993:3-14
5. Y. Linde, A. Buzo, and R.M. Gray. An algorithm for vector quantization design", IEEE Trans. Commun, Vol. COM-28, Jan, 1980, 84-95
6. R. M. Gray. Vector Quantization. IEEE Transzction on ASSP,1984,1,4-29