論文 94-19-12-17

Dynamic Excitation Modeling Scheme Applied for Variable Low Bit-Rate Homomorphic Vocoder

Jae Ho Chung* *Regular Member*

가변 저 전송율 호모몰픽 보코더에 응용된 동적 음원 모델링 기법

正會員 鄭 在 皓*

ABSTRACT

In this paper, a new dynamic excitation modeling scheme is proposed. Based upon the proposed excitation modeling scheme, two variable bit rate homomorphic vocoders are designed, whose average bit rates are 3.8 Kbps and 4.4 Kbps. The performance of the proposed excitation modeling scheme is thên evaluated through the subjective listening tests. In the tests, the performances of two speech coders designed in this paper are compared with the one of 4.8 Kbps homomorphic vocoder designed by Chung and Schafer, in which conventional static excitation modeling scheme applied. The subjective listening tests show that proposed dynamic excitation modeling scheme improves synthesized speech quality while lowering the average bit rate of speech coders.

要 約

본 논문에서는 음원신호의 새로운 동적 모델링 기법을 제안하였다. 제안된 동적 모델링 기법을 적용하여, 평균 3.8 Kbps와 4.4 Kbps에서 작동하는 두개의 가변 저 전송율 호모몰픽 보코더를 구현하였다. 본 논문에서 제시한 모델링 기법의 성능은 청음 테스트를 사용하여 분석하였으며, 청음 테스트에서는 본 논문에서 구현한 두개의 코더가 기존의 정적 음원 모델링 방법을 사용하여 Chung과 Schafer에 의하여 구현된 4.8 Kbps 호모몰픽 보코더가 비교되었다. 제안된 동적 음원 모델링 기법이 음성통신의 평균 전송율을 낮추면서, 반면에 합성된 음의 음질을 향상시킴을, 청음 테스트를 통하여 보였다.

I. Introduction

The use of analysis-by-synthesis in determining the excitation signal has significantly advanced the state of the art in LPC vocoders [1, 2]. The same principle has been adopted to obtain the excitation signal in the homomorphic vocoder, i.e.,

*인하대학교 전자공학과 디지탈 신호처리 연구실
Department of Electronics Engineering, Digital Signal Processing Laboratory, Inha University
論文番號 : 94247
接受日字 : 1994年 9月 14日

an exhaustive search procedure is used to determine the excitation by minimizing a perceptually weighted difference between the original speech and the output of the vocoder synthesizer [3, 4]. The systematic evaluation has shown that the homomorphic vocoder model with analysis-by-synthesis excitation can produce high quality synthetic speech. In this paper, the homomorphic vocoder is further developed by applying a very effective excitation modeling scheme, called the dynamic excitation modeling scheme.

After a brief review of previous work [3, 4] in Section Ⅱ, a new improved excitation model is described in Section Ⅲ. This new model emphasizes either the pitch dependent periodic nature or the more random nature of the sound sources depending on the state of each speech frame. Since different voicing states result in different excitation modeling, the classification of a given speech block into a correct voicing state is important. The cepstrum(which is the basis for the homomorphic vocoder) is used as a primary tool for classifying a given speech segment into three different classes each of which is modeled differently. The voicing classification algorithm is discussed in detail in Section Ⅲ. In Section Ⅳ, the proposed new excitation modeling scheme is applied for the design of two variable bit rate homomorphic vocoders whose average bit rates are 3.8 and 4.4 Kbps. In Section Ⅴ, the performance of the proposed excitation modeling scheme is analyzed by performing the subjective tests on the designed coders. In the subjective listening tests, the performances of two speech coders designed in this paper were compared with the one of 4.8 Kbps homomorphic vocoder designed by Chung and Schafer [3, 4], in which conventional static excitation modeling scheme applied. The evaluation results show that the new excitation modeling method improves the performance of the homomorphic vocoder in low bit rate coding environments. Finally, the conclusions of the paper are given in Section Ⅵ.

Ⅱ. The Fundamental Vocoder Framework [3, 4]

The homomorphic filtering procedure which provides the vocal tract information in the homomorphic vocoder is depicted in Figure 1. Each block of a pre-emphasized speech signal $s(n)$ is first weighted by a Hamming window $v(n)$, and then the corresponding cepstrum $c(n)$ is computed by applying the sequential operations diagrammed in the upper part of Figure 1. The low-time cepstrum $\hat{h}(n)$ representing the vocal tract information

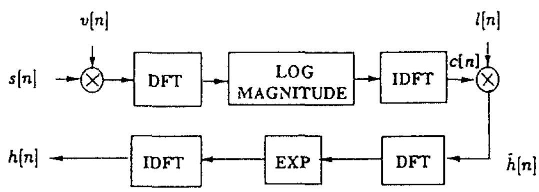
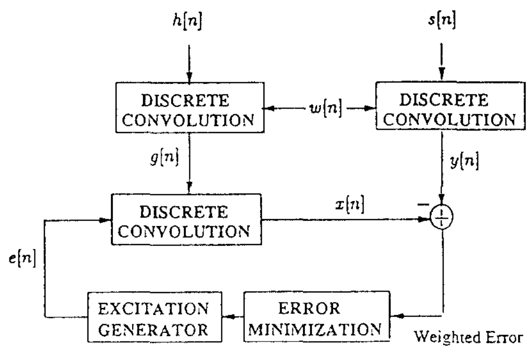


Figure. 1. Homomorphic filtering for estimating the vocal tract impulse response.



$s[n]$: pre-emphasized speech
$w[n]$: perceptual weighting filter impulse response
$y[n]$: "weighted" speech
$h[n]$: vocal tract impulse response
$g[n]$: "weighted" vocal tract impulse response
$x[n]$: "weighted" synthetic speech
$e[n]$: excitation signal

Figure. 2. Analysis-by-synthesis method for obtaining the excitation sequence $e(n)$.

is then extracted using a lifter $l(n)$: i.e., $\hat{h}(n) = c(n)l(n)$ with

$$l(n) = \begin{cases} 2, & 1 \le n \le n_0 \\ 0, & otherwise, \end{cases} \tag{1}$$

where n_0 is less than the pitch period. Finally, applying the operations depicted at the lower part of Figure 1, the impulse response $h(n)$ is derived from the low-time cepstrum $\hat{h}(n)$. Observe that the impulse response $h(n)$ is normalized automatically because zeroth $\hat{h}(n)$ has the value of zero. This impulse response represents the vocal tract response during the time interval corresponding to the input block of speech samples.

In the homomorphic vocoder developed by Chung and Schafer [3, 4], the excitation signal $e(n)$ is determined based on the analysis-by-synthesis excitation algorithm shown in Figure 2. The excitation signal $e(n)$ is composed of the following two parts : $\beta_1 f_{\gamma_1}(n)$, where $f_{\gamma_1}(n)$ is a zero-mean Gaussian codebook sequence corresponding to index γ_1 in the codebook, and $\beta_2 e(n-\gamma_2)$, which represents a short segment of the past (previously computed) excitation beginning γ_2 samples before the present excitation frame, i.e.,

$$e(n) = \beta_1 f_{\gamma_1}(n) + \beta_2 e(n-\gamma_2). \tag{2}$$

The perceptually weighted synthetic speech $x(n)$ corresponding to $e(n)$ has the form of

$$x(n) = \beta_1 x_1(n) + \beta_2 x_2(n), \tag{3}$$

where $x_1(n) = g(n) * f_{\gamma_1}(n)$, $x_2(n) = g(n) * e(n-\gamma_2)$, and $g(n) = w(n) * h(n)$ is the perceptually weighted vocal tract impulse response. The weighting of the speech signal and the impulse response with the weighting filter $w(n)$ before synthesis is to improve subjective speech quality by concentrating the coding noise in the formant regions of the spectral envelope [5, 6].

The optimum parameters, β_1, γ_1, β_2, and γ_2 are determined pairwise in the following manner. First, the parameters γ_2 and β_2 are chosen by minimizing the mean-squared error

$$E_0 = \sum_n |y(n) - \beta_2 x_2(n)|^2. \tag{4}$$

For a given γ_2, the value of β_2 that minimizes the mean-squared error in Equation (4) is given by

$$\beta_2 = \frac{\sum_n y(n)\, x_2(n)}{\sum_n x_2^2(n)}. \tag{5}$$

The optimum values for γ_2 and β_2 are found by an exhaustive search with values of γ_2 restricted to a finite range. Then the residual signal $y_1(n) = y(n) - \beta_2 x_2(n)$ is formed and the parameters γ_1 and β_1 are chosen by an exhaustive search of the Gaussian codebook to minimze

$$E_1 = \sum_n |y_1(n) - \beta_1 x_1(n)|^2. \tag{6}$$

As before, the value of β_1 that minimizes the mean-squared error for a given codebook sequence $f_{\gamma_1}(n)$ is

$$\beta_1 = \frac{\sum_n y_1(n)\, x_1(n)}{\sum_n x_1^2(n)}. \tag{7}$$

At the synthesizer, each block of an excitation sequence $e(n)$ is convolved with the corresponding vocal tract impulse response $h(n)$ to produce the synthetic speech output $\tilde{s}(n)$ using the overlap-add method [7].

Ⅲ. New Modeling Scheme of the Excitation Signal

The excitation generator introduced in Section Ⅱ routinely searches through a fixed interval of the past excitation signal, and then searches through the Gaussian codebook without considering the state of the speech segment, i.e., voiced

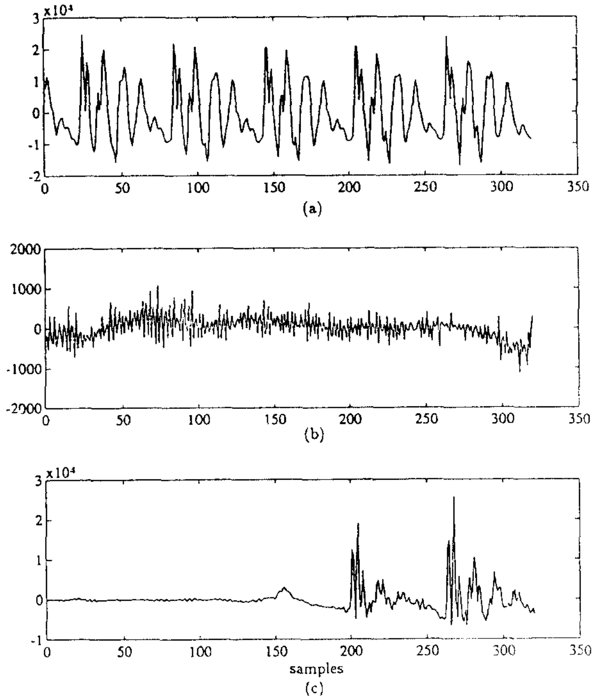


Figure. 3. A speech segment of (a) voiced, (b) unvoiced, and (c) mixed.

or unvoiced. It is likely that this static excitation modeling scheme is not the best way to describe the nonstationary behavior of the sound sources. Figure 3 shows three distinctly different types of speech segments, namely, voiced, unvoiced, (including silence), and mixed (including transient). In the case of a voiced segment, the speech waveform is dominated by the quasi-periodic nature. For an unvoiced segment, the speech waveform appears to vary randomly. Finally, a mixed speech segment has the quasi-periodic nature as well as the random nature. Since three waveform segments are quite different from one another, in the excitation modeling, it would be more effective to emphasize either the pitch dependent periodic nature or the more random nature of the sound sources depending on the voicing state of a given speech segment. This suggests the dynamic excitation modeling scheme.

3.1 The dynamic Excitation modeling

In dynamic excitation modeling, different excitation modeling strategies are applied for the different calsses of speech sounds, i.e., voiced, unvoiced, and mixed. For voiced segments, the excitation signal $e(n)$ is comprised of two sequences selected from a time-varying queue of the past excitation history, i.e.,

$$e(n)=\beta_0 e(n-\gamma_0)+\beta_1 e(n-\gamma_1). \tag{8}$$

This emphasizes the pitch dependent periodic nature. In the unvoiced case, the excitation signal $e(n)$ is modeled by a Gaussian codebook sequence, i.e.,

$$e(n)=\beta f_\gamma(n), \tag{9}$$

emphasizing the random nature. Finally, in the case of segments classified as mixed excitation, the excitation $e(n)$ is modeled as the sum of a Gaussian codebook sequence and a sequence selected from the fixed interval of the past excitation history, i.e.,

$$e(n)=\beta_0 e(n-\gamma_0)+\beta_1 f_{\gamma_1}(n). \tag{10}$$

The memory of the past excitation history is updated by whatever model is chosen during a given excitation frame.

Since different voicing states result in different excitation modeling, the classification of a given speech block into a correct voicing state is important. The cepstrum has been successfully applied as a tool for voicing decision as well as detection [8, 9, 10]. Figure 4 shows the corresponding cepstrums of the speech segments shown in Figure 3. The periodic nature of the voiced speech segment allows strong cepstral peaks, whereas the random nature of the unvoiced speech segment does not reflect any strong cepstral peaks. In the homomorphic vocoder, the vocal tract information is represented by the low-time cepstrum and therefore the cepstrum is available

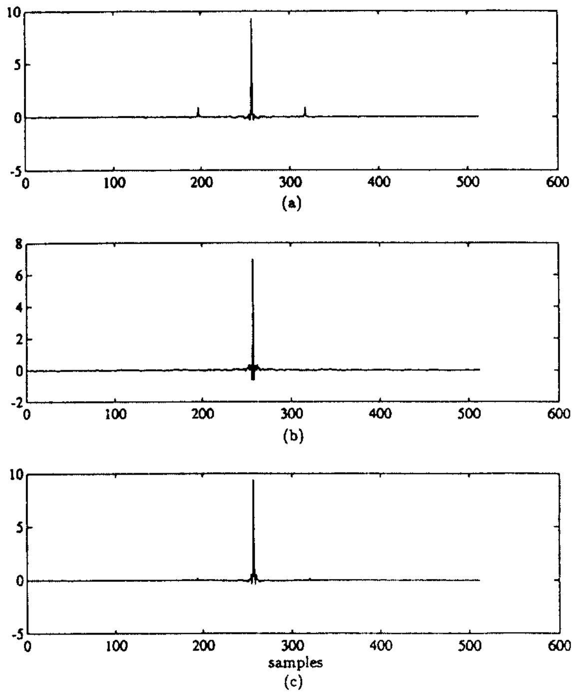


Figure. 4. Cepstrums of speech segment of (a) voiced, (b) unvoiced, and (c) mixed.

without any additional computation. Consequently, a voicing classification method primarily based on the cepstrum is adopted in our dynamic excitation modeling system. The voicing classification algorithm is discussed in the following.

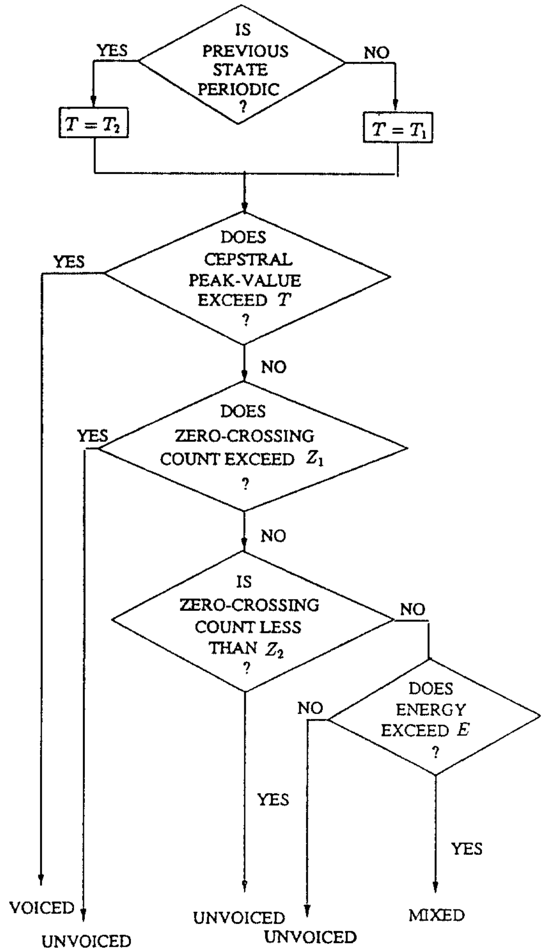


Figure. 5. The voicing classification algorithm.

3.2 The Voicing Classification Algorithm

Figure 5 shows a flow diagram of the voicing classification algorithm suggested in this paper. The state of the speech segment is first classified into either periodic (i.e., voiced) or non-periodic (mixed and unvoiced) based on the strength of the cepstral peak value in the range $25 \leq n \leq 100$. The cepstral peak position coincides with the pitch mark. Thus, the cepstral peak search is performed in the range $25 \leq n \leq 100$. This range is adequate to include both female and male speakers. If the value of this peak exceeds a threshold T, the segment is classified as periodic, otherwise as non-periodic. The threshold T is switched between two values T_1 and $T_2 (T_1 > T_2)$ depending on the state of the previous speech segment. More precisely, the value of T_1 is used as a threshold if the previous state is non-periodic, whereas the value of T_2 is used as a threshold if the previous state is periodic. The size of FFT used is 512. Values of $T_1 = 0.3$ and $T_2 = 0.2$ yields good performance.

Non-periodic segments are further classified into either mixed or unvoiced (also include silence) based on the zero-crossing count and the energy of the original speech segment. If the zero-crossing count either exceeds a threshold Z_1 or does not exceed a threshold Z_2 ($Z_1 > Z_2$), the segment is classified as unvoiced. If the zero-crossing count lies in between Z_1 and Z_2, the energy of the block is computed. If the energy of the block exceeds a threshold E, the state of the speech block is classified as mixed. Otherwise it is classified as unvoiced.

When the misclassification of voicing happens, the corresponding excitation blocks will be modeled less effectively. However, less effectiveness does not mean no modeling, since each excitation block will be modeled anyway using one of the three types. The worst misclassification would be the classification of a voiced block as unvoiced one. In this case, the excitation block will be modeled very poorly, but the chance of this misclassification occurring is very slim. The most probable misclassification would be the confusion between the voiced and mixed blocks. In these cases, however, the excitation modeling will be switched from one to the other between the two excitation types given in Equations (8) and (10). Thus, in these cases, the misclassifications would not be serious.

Table 1 shows the ratio among the speech segments of voiced, unvoiced, and mixed for a total of 4,092 speech segments (160 samples long each) collected from four different speakers (three male, 1 female, 11 sentences/speaker). The ratio shows that almost half of the total speech segments are unvoiced (including silence). Since a large portion of speech signal is unvoiced, and since unvoiced segments are represented by a single Gaussian codebook sequence in the dynamic excitation modeling, a significant savings in bits for representing unvoiced segments can be derived, while maintaining the synthetic speech quality. Indeed,

Table 1. Ratio among the speech segments of voiced, unvoiced, and mixed.

	number of segments	ratio
voiced	1732/4092	0.423
unvoiced	1977/4092	0.483
mixed	383/4092	0.094

in the next section, two variable bit rate homomorphic vocoders whose average bit rates are 3.8 and 4.4 Kbps are designed, using the proposed dynamic excitation modeling scheme.

Ⅳ. Design of Variable Bit Rate Homomorphic Vocoders

To verify the performance of the dynamic excitation modeling scheme, 3.8 Kbps and 4.4 Kbps homomorphic vocoders are designed. The performances of these two variable bit rate coders will be compared with the one of 4.8 Kbps fixed bit rate homomorphic vocoder, designed by Chung and Schafer [3, 4], in which the excitation was modeled using the scheme introduced in Section Ⅱ.

First, the cepstrums are coded using the vector quantization scheme. Since the voicing state is classified into three different classes, and since the spectral characteristics of each state are very dissimilar, the codebook used to code the low-time cepstrum is composed of three different groups for three different voicing states, i.e., voiced, unvoiced, and mixed.

To design the codebook, first, 4092 cepstral training vectors were obtained from 44 utterances (3 male, 1 female, 11 utterances/speaker) by performing the vocal tract analysis for every 160 samples long speech block. Then, the cepstral training vectors were classified into three different subsets using the voicing classification algorithm discussed earlier, and then each of the three different subsets was used to design a sub-codebook. The Euclidean cepstral distance measure was used as a distance measure, and the maximin-distance algorithm was used to provide the initial sub-codebooks [11]. The number of training vectors in the subsets are summarized in Table 1.

From the training data, a cepstrum codebook represented by 12-bits was designed. The first region of the codebook having 1,732 entries is used for the voiced speech segments. The second region is used for the unvoiced speech segments, and has 512 entries. The last region has 256 entries, and used to code the cepstrums of mixed speech segments. Since the codebook is composed of three sub-codebooks, by arranging the sub-codebooks sequentially the search for the best codeword can be restricted to the particular sub-codebook depending on the voicing state.

Two different systems have been designed. Table 2 shows the bit allocations used for the first system. The maximum bit rate is 5.6 Kbps when the voicing state is mixed, and the minimum bit rate is 3.2 Kbps when the voicing state is unvoiced. The average bit rate is about 4.4 Kbps.

Table 3 shows the bit allocations used for the second system. The maximum bit rate is 4.8 Kbps when the voicing state is either voiced or mixed, and the minimum bit rate is 2.667 Kbps when the voicing state is unvoiced. The average bit rate is about 3.8 Kbps.

In the case of a voiced segment, the scale factors β_0 and β_1 were quantized separately using non-uniform scalar quantizers. The locations γ_0 and γ_1 were coded separately. For the unvoiced case, the gain parameter β was quantized using an APCM coder introduced by Jayant [12]. In the case of a mixed segment, the scale factor β_0 was coded using a non-uniform quantizer, and the gain parameter β_1 was quantized using an APCM coder.

Computer simulations were conducted to evaluate the performance of the designed homomorphic vocoders. The simulations showed that the quality of the synthetic speech synthesized using the dynamic excitation modeling was superior to that of the homomorphic vocoder which uses fixed form of excitation pattern (see Section Ⅱ). The variable bit rate vocoders designed in this section will be tested subjectively in the next section. The designed two 3.8 and 4.4 Kbps homomorphic vocoders will be denoted as VBHV_3.8 and VBHV_4.4, respectively.

Ⅴ. Performance Evaluation

For the subjective evaluation of the variable bit rate homomorphic vocoders designed in Section Ⅳ, a variation of the Paired Acceptability Rating Method (PARM) was used as a subjective quality measure [13, 14]. The PARM measure is an isometric speech quality measure in which, the listeners rate the perceived speech quality of a set of speech communication systems on a scale of 0 to 100. The PARM test involves the comparison of all possible pairs of systems in that PARM module. The test begins with the presentation of the high and low anchors of the PARM module and their corresponding scores. The synthetic speech from a pair of systems is then presented to the listeners, and the listeners are

Table 2. Bit allocations used for 4.4 Kbps homomorphic vocoder.

	Bits/Frame					Samples/Frame		Bit Rate Kbps
	cepstrum	β_0	β_1	γ_0	γ_1	excitation	cepstrum	
voiced	12	5	5	7	7	40	160	5.4
unvoiced	12	5	0	8	0	40	160	3.2
mixed	12	5	5	7	8	40	160	5.6

Table 3. Bit allocations used for 3.8 Kbps homomorphic vocoder.

	Bits/Frame					Samples/Frame		Bit Rate Kbps
	cepstrum	β_0	β_1	γ_0	γ_1	excitation	cepstrum	
voiced	12	5	5	7	7	45	180	4.8
unvoiced	12	5	0	7	0	45	180	2.667
mixed	12	5	5	7	7	45	180	4.8

asked to rate each sentence on a score of 0 to 100. The high and low anchors have fixed scores of 80 and 20 respectively, and the anchors are presented to the listeners at the beginning and periodically during the test.

In our subjective tests, an undistorted original signal was used as the high anchor. The low anchor was a heavily distorted synthetic speech signal produced by a homomorphic vocoder using a Gaussian random ensemble as the excitation. For the tests, 16 untrained listeners participated in the subjective tests.

The test consisted of the three fully quantized homomorphic vocoders besides the high and low anchors. Among the three vocoder systems involved in the test, two of them were the variable bit rate homomorphic vocoders designed in Section Ⅳ, namely, VBHV_3.8 and VBHV_4.4. The last coder included was a homomorphic vocoder operating at the fixed bit rate of 4.8 Kbps. Let's denote the coder as FBHV_4.8. The coder was designed according to the description in Section Ⅱ. The coder FBHV_4.8 was used as a reference for measuring the performance of VBHV_3.8 and VBHV_4.4 vocoders. The bit allocation of FBHV_4.8 is shown in Table 4. The parameters of FBHV_4.8 were quantized according to the schemes introduced in Chung and Schafer [3, 4]. A sequence of 12 cepstrums was vector quantized using a 256 size full search codebook designed using Euclidean cepstral distance measure. Each of the two excitation gain parameters β_0 and β_1 was coded using a 4-bit APCM coder.

The mean score and corresponding standard deviation for each of systems are summarized in Table 5. The level of significances (or confidences) among the systems are tabulated in Table 6. Each entry indicates the level of significance of the system at the top row of its column with respect to the system at the leftmost column of its row. Several observations can be made from the test results. First, the coder VBHV_4.4 received the highest score 54.5 which was higher than the one received by FBHV_4.8. The level of significance of the difference between the means of the systems VBHV_4.4 and FBHV_4.8 is 0.55. Second, the coder VBHV_3.8 scored 51.4 which was slightly lower than the score of FBHV_4.8. The level of the significance between the systems FBHV_4.8 and VBHV_3.8 is 0.60. Consequently, we can conclude from the test results that the dynamic excitation modeling is a very effective scheme to lower the average bit rate while maintaining the level of speech quality.

Table 5. Subjective test results.

	Average Score	Standard Deviation
High Anchor	74.8	6.9
VBHV_4.4	54.5	6.2
FBHV_4.8	53.7	5.6
VBHV_3.8	51.4	6.8
Low Anchor	25.0	4.6

Table 6. The level of significances of the systems for the subjective tests.

	HIGH	VBHV_4.4	FBHV_4.8	VBHV_3.8
VBHV_4.4	1.0			
FBHV_4.8	1.0	0.55		
VBHV_3.8	1.0	0.60	0.60	
LOW	1.0	1.00	1.00	1.00

Ⅵ. Conclusions

In this paper, a new excitation modeling scheme was proposed. This new model emphasizes either the pitch dependent periodic nature or the more

Table 4. Bit allocations for FBHV_4.8 kbps homomorphic vocoder.

Bits/Frame					Samples/Frame		Bit Rate (Kbps)
cepstrum	β_0	β_1	γ_0	γ_1	excitation	cepstrum	
8	4	4	7	7	40	160	4.8

random nature of the sound sources depending on the state of each speech frame. Based upon the proposed excitation modeling scheme, two variable bit rate homomorphic vocoders were designed whose average bit rates are 3.8 Kbps and 4.4 Kbps, named VBHV_3.8 and VBHV_4.4, respectively. To evaluate the performance of the proposed excitation modeling scheme, the subjective tests were performed on the designed vocoders. The subjective tests included three coders. Two of them were VBHV_3.8 and VBHV_4.4 designed in this paper. The last coder included in the test was the 4.8 Kbps fixed bit rate homomorphic vocoder which was designed by Chung and Schafer [3, 4], named FBHV_4.8. The coder FBHV_4.8 was used as a reference coder to measure the performances of VBHV_3.8 and VBHV_4.4 coders. The test results showed that VBHV_4.4 coder performs better than FBHV_4.8. Also, the test results indicated the mean score achieved by VBHV_3.8 was slightly lower than that of FBHV_4.8. Consequently, we can conclude from the test results that the proposed dynamic excitation modeling scheme is very effective to lower the average bit rate while improving the quality of synthesized speech signal.

References

1. B. S. Atal and J. R. Remde, "A new model of LPC excitation for producing natural-sounding speech at low bit rates," *Proc. Int. Conf. Acoust, Speech, Signal Processing,* pp.614-617, 1982.
2. B. S. Atal and M. R. Schroeder, "Predictive coding of speech and subjective error criteria," *IEEE Trans. Acoust., Speech, and Signal Proc.,* vol.ASSP-27, pp.247-254, June 1979.
3. J. H. Chung and R. W. Schafer, "A 4.8 Kbps homomorphic vocoder using analysis-by-synthesis excitation analysis," *Proc. Int. Conf. Acoust, Speech, Signal Processing,* pp.144-147, 1989.
4. J. R. Deller, J. G. Proakis, and J. H. Hansen, *Discrete-time processing of speech signals,* Macmillan Publishing Company, pp.467-469, 1993.
5. B. S. Atal, "Predictive coding of speech at low bit rates," *IEEE Trans. on Communications,* vol. 30, pp.600-614, April 1982.
6. A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing,* Prentice-Hall, Englewood Cliffs, N.J., 1988.
7. A. Noll, "Cepstrum pitch determination," *Journal of Acoust. Soc. Amer.,* pp.293-309, vol.41, Feb. 1967.
8. R. W. Schafer and L. R. Rabiner, "System for automatic formant analysis of voiced speech," *Journal of Acoust. Soc. Amer..* pp.634-648, vol.47, Feb. 1970.
9. L. Rabiner, M. Cheng, A. Rosenberg, and C. McGonegal, "A comparative performance study of several pitch detection algorithm," *IEEE Trans. on Acoust., Speech, and Signal Proc.,* vol. ASSP-24, no.5, Oct. 1976.
10. J. T. Tou and R. C. Gonzalez, *Pattern Recognition Principles,* Addison-Wesley, 1974.
11. N. S. Jayant, "Adaptive quantization with a one word memory," *Bell System Technical Journal,* pp.1119-1144, September 1973.
12. W. D. Voiers, "Methods of predicting user acceptance of voice communications systems," *Final Report 100-74-C-0056, DCA,* April 1976.
13. R. C. Rose and T. P. Barnwell, "Design and performance of an analysis-by-synthesis class of predictive speech coders," *IEEE Trans. on Acoust., Speech, and Signal Proc.,* vol.ASSP-38, no.9, pp.1489-1503, Sep. 1990.

※본 연구는 93년도 인하대학교 연구비지원에 의하여 수행되었음.

鄭 在 皓(Jae Ho Chung) 정회원
1982년 : 美國 University of Maryland(학사)
1984년 : 美國 University of Maryland(석사)
1990년 : 美國 Georgia Institute of Technology(박사)
1984년～1985년 : 美國 Naval Surface Warfare Center, Electronic Engr.
1991년～1992년 : 美國 AT&T Bell Laboratories, 연구원
1992년～현재 : 인하대학교 공과대학 전자공학과, 조교수