Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Adaptive Compression of Respiratory and Swallowing Sounds

A. Yadollahi, *Student Member,* Z. Moussavi, *Senior Member,* P. Yahampath, *Member*

Abstract— **Compression of biomedical signals is of great interest in telemedicine applications for a fast and reliable online data transfer. In this paper, an adaptive method for compression of respiratory and swallowing sounds is proposed. The method is based on the transform coding of the signal and adaptively assigns two different bit allocations for encoding the stationary and non-stationary parts of the signal. The method was applied to data of 12 subjects and its performance was compared with other methods of bit allocation such as constant bit allocation. The results show that the proposed adaptive bit allocation method improves the SNR of the signal by 3 dB compared to that of constant bit allocation method. Furthermore, the possibility of removing the requirement of updating the bit allocation method for each subject was investigated. The results confirm that with a few training data, the proposed method can be used in a fully automated mode without the need to adjust the adaptive bit allocation for every signal separately; hence, it is faster for online coding applications.**

I. Introduction

With advances in digital signal processing, compression of biomedical signals has received great attentions for use in telemedicine applications [1-4]. These studies were mainly focused on the applications of wavelet transforms or predictive coding for compression of ECG signals, EMG signals or biomedical images. Our research group is involved in the development of an integrated respiratory and swallowing sounds remote assessment tool, in which the sounds with a bandwidth of over 5 kHz have to be transferred online. Hence, compression of such data is of interest for a fast and reliable transmission. Since the characteristics of respiratory and swallowing sound signals are more similar to speech signals than to ECG or EMG signals, speech compression methods were investigated more thoroughly in this study.

Frequency domain coding is one of the most important classes of techniques for speech compression. In these methods, the frequency components of the speech signal are extracted using either a filter bank or a suitable transform, and then encoded with scalar quantizers [5, 6]. Comparing different transforms, discrete cosine transform (DCT) has been found to have near optimum performance for coding of speech signals [7]. In addition to finding an optimum transform, allocating bits to different transform coefficients in an efficient manner is a challenge in transform coding applications. Many researchers have attempted to develop adaptive bit allocation methods based on the non-stationary characteristics of the speech signals in which the transform coefficients change with time [6-9].

In this study a compression method based on DCT coding was investigated for respiratory and swallowing sounds signals. The adaptive bit allocation methods developed in [6-9] are not well suited for our application as they are dependent on the special characteristics of the speech signals. Those methods also require time consuming computations. Furthermore, the swallowing sounds are non-stationary signals by nature, while the respiratory sounds can be considered as wide-sense stationary signals, especially in parts where the airflow plateaus. Therefore, a new adaptive bit allocation method was sought in this study, which uses the difference between the characteristics of the DCT coefficients of the stationary and non-stationary parts of the respiratory and swallowing sounds signal.

II. Method

A. Data

In this study 14 respiratory and swallowing sound recordings of 12 healthy children (ages 3-16 years) were adopted from a previous study data set [10]. During the test, all participants were fed 180 ml thin liquid (juice) and thick liquid (pudding diluted with milk) in single boluses of 5 ml each. Respiratory and swallowing sounds were recorded using Siemens accelerometer (EMT25C) placed over suprasternal notch using double adhesive tapes. The respiratory and swallowing sounds were amplified, band-pass filtered (30-2500 Hz) and digitized at 10240 Hz sampling rate and 12 bits/sample (National Instruments™ DAQCard).

B. Transform Coding

In transform coding the signal is coded in two steps. In the first step a linear transformation is applied to the signal and then the transform coefficients are quantized individually (Fig. 1). Choices of the optimum transform function and the optimum bit allocation method are the major problems in transform coding.

The total quantization error is defined as:

$$\overline{D}=E\left(\left\|X-\hat{X}\right\|^2\right), \tag{1}$$

where X and $\hat{X}$ are the actual and the compressed signals, respectively and $E(.)$ is the mean value. It was shown that the minimum error is achieved using Karhunen-Loeve

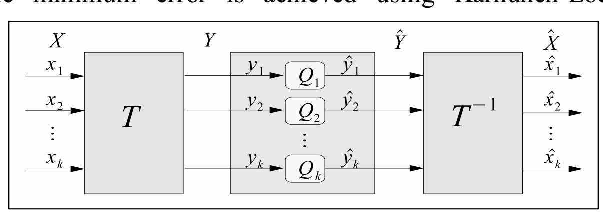

Fig. 1. Schematic of the transform coding system.

transform (KLT) [6, 11]. The main disadvantages of KLT are due to its signal-dependent characteristics and the computational difficulties in calculating its parameters [7]. Some other transforms such as the Walsh-Hadamard transform (WHT), discrete slant transform (DST), discrete Fourier transform (DFT) and DCT have been investigated in comparison with KLT for speech processing purposes in [7]. They found that among these transforms the performance of DCT was better than the other transforms and its performance was close to that of the KLT [7]. KLT is based on the eigenvectors of the correlation matrix of X which is a Toeplitz matrix. The near optimum performance of DCT arises from the fact that the basis vectors of the DCT closely approximate the eigenvectors of a class of Toeplitz matrices [6]. The main advantage of the DCT to KLT is its signal-independent definition. Also, comparing with DFT, the next best transform, DCT produces less boundary effects in the transform coding applications.

The DCT of a real M point signal $v(n)$ is defined as [6]:

$$V_c(k)=\sum_{n=0}^{M-1} v(n)c(k)\cos\left[\frac{(2n+1)\pi k}{2M}\right], k=0,1,\cdots,M-1, \quad (2)$$

where

$$c(k)=\begin{cases} 1 & k=0 \\ \sqrt{2} & k=1,2,\cdots,M-1 \end{cases} \quad (3)$$

The inverse of DCT is defined as [6]:

$$v(n)=\frac{1}{M}\sum_{k=0}^{M-1} V_c(k)c(k)\cos\left[\frac{(2n+1)\pi k}{2M}\right], n=0,1,\cdots,M-1. \quad (4)$$

The choice of bit allocation determines the accuracy of encoding the transform coefficients and it is highly dependent on the significance of the different coefficients. Let B and M represent the total number of bits and transform coefficients, respectively. The simplest method is the uniform bit allocation defined as [11]:

$$b_i=\bar{b}=\frac{B}{M}, \quad (5)$$

where b_i shows the number of bits associated to the i^{th} coefficient. Here, it is assumed that all coefficients have equal significance. Another bit allocation method takes into account the variance of the coefficients assuming all coefficients have similar probability density function (pdf) [11]:

$$b_i=\bar{b}+\frac{1}{2}\log_2\left[\frac{\sigma_i^2}{\rho^2}\right], \quad (6)$$

where ρ^2 is the geometric mean of the variance of the transform coefficients. Another bit allocation method considers different pdf for the coefficients and it is defined as [11]:

$$b_i=\bar{b}+\frac{1}{2}\log_2\left[\frac{\sigma_i^2}{\rho^2}\right]+\frac{1}{2}\log_2\left[\frac{h_i^2}{H^2}\right], \quad (7)$$

where

$$h_i=\frac{1}{12}\left\{\int_{-\infty}^{\infty}[f_i(x)]^{1/3}dx\right\}^3, \quad (8)$$

$f_i(x)$ is the pdf of the normalized signal and H is the geometric mean of h_i. These three bit allocation methods were investigated in this study.

C. Respiratory and Swallowing Sounds Compression

In order to compress the respiratory and swallowing sounds, the recorded signal was segmented into windows (frames) of 128 samples (12.5 ms). In each window the signal was normalized by the estimated standard deviation of that window. This normalization factor is sent to the receiver as side information [7]. Then, each normalized window was encoded using a 128-point DCT.

In order to evaluate the effects of bit allocation, the three bit allocation methods described by Eq. 5, 6 and 7 were investigated. For each subject 20 s of the signal was used as training data to find the bit allocation method for DCT coefficients and design the optimum quantizer. In the case of allocating equal bits to all DCT coefficients (Eq. 5), there are not enough bits for encoding all of the coefficients (about 2 bits at 16 kbit/sample). Therefore, only the first 30 DCT coefficients that were found to be more important than the rest were encoded and the rest of the coefficients were set to zero.

When the DCT coefficients are assumed to be identically distributed as in Eq. 6, the only unknown parameter is the standard deviation which was estimated from the training data. The use of Eq. 7, on the other hand, requires estimating a pdf for each DCT coefficient. In this case, the histogram was used to approximate the pdf $f_i(x)$ and to find the value h_i in Eq. 8 of the i^{th} DCT coefficient.

The generalized Lloyd algorithm [11] was used to find the optimum encoder and decoder for quantizing the DCT coefficients. Having found the optimum quantizer, the rest of the signal was used as the test signal to evaluate the performance of different bit allocation methods. In this study, the measure of SNR:

$$SNR=10*\log_{10}\left[\frac{\mathrm{var}(X)}{\bar{D}}\right], \quad (9)$$

where D is mean value of the error as defined in Eq 1, and X is the original signal, was used to evaluate the quantizer performance.. The mean and standard deviation of the SNR values were calculated among the subjects. After finding the superior bit allocation method, it was used for further investigation. Also, the performance of the methods for different bit rates of 32, 16, 8, 4, 2 and 1 kbit/sample were considered to find the effects of bit rate and examine the smallest possible bit rate for transmitting the data.

In previous studies, the objective of adaptive transform coding was to account for the non-stationary properties of speech signals in different windows [6-9]. Instead of using a constant bit allocation for all windows, the researchers tried to estimate the spectral characteristics of the DCT coefficients in each window. Then, this information was used to update the bit allocation and it was sent to the receiver as side information. They applied averaging and

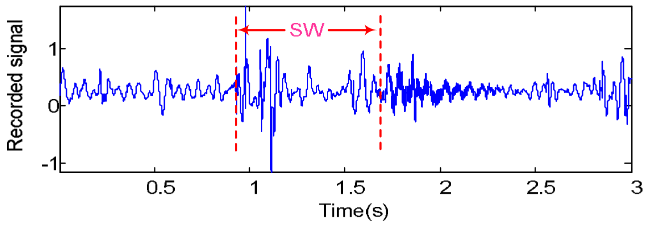

Fig. 2.A typical respiratory and swallowing sounds record.

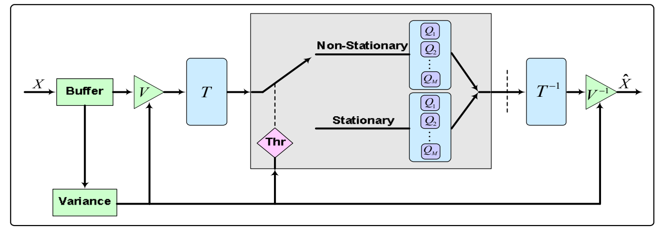

Fig. 3. Schematic of adaptive compression system.

interpolation of the coefficients [7], linear predictive coding (LPC) model of speech signal [6] and pseudo cepstrum [9] to find the spectral characteristics of the signal. However, these methods are time-consuming and also specialized for speech signals. In addition, they require about 2 kbit/sample for transmitting the side information. Therefore, a new adaptive bit allocation method for improving the performance of the coding system was investigated in this study.

Figure 2 shows a typical record of normal respiratory and swallowing sounds. As mentioned earlier, the swallowing sounds are mainly non-stationary while the respiratory sounds can be mostly considered stationary. Hence, their characteristics are significantly different. These changes in the characteristics of different parts of the signal were considered to develop an adaptive method for coding the recorded respiratory and swallowing sounds signal.

Using the training data for each subject, the mean value of the variance estimated within in each window was calculated. This value was used as a threshold to classify each window as stationary or non-stationary. Then, two different bit allocations were used for stationary and non-stationary windows. Consequently, different optimum quantizers were also designed for stationary and non-stationary windows. Figure 3 shows the schematic of the proposed adaptive compression system.

In the last experiment the possibility of using some general fixed bit allocations for the stationary and non-stationary parts of any respiratory and swallowing sounds signal was investigated. The signals of 5 subjects were used to determine the bit allocations for DCT coefficients of stationary and non-stationary parts of the signals. Then, instead of finding the optimum bit allocation for each subject individually, the averages of these bit allocations for each stationary and non-stationary part was used as a "*predefined adaptive bit allocation*" to design the quantizers and compressing the signals of each subject. The goal of this experiment was to investigate the robustness of the method to the changes in the bit allocations of the DCT coefficients due to variations of the signals among the subjects.

III. Results and Discussion

In the first experiment the effects of different bit allocation methods of DCT coefficients were investigated. Using uniform bit allocation for all DCT coefficients (Eq. 5), the performance of the method was found very poor with average SNR of 12.47 ± 6.02 at 16 kbit/sample compared to those of other methods (20.48 ± 5.53 and 20.07 ± 5.78 for similar and different pdf based methods at 16 kbit/sample, respectively). Therefore, this method was removed from further investigations. Figure 4 shows the results of the two other bit allocation methods (similar pdf for all of the DCT coefficients (Eq. 6) or different pdf for each DCT coefficient (Eq. 7)) for different bit rates. Comparing the results, it is clear that the performances of both bit allocation methods are similar. Investigating the pdf of different DCT coefficients, it was found that almost all of the coefficients (except a few initial coefficients) have similar pdfs. Therefore, incorporating the parameters h_i and H which indicate the difference in the pdf of DCT coefficients does not change the results. On the other hand, the bit allocation method presented in Eq. (6) is superior in terms of computational cost. Therefore, this method was chosen for further investigation.

Next, the difference in the characteristics of DCT coefficients among the windows including stationary or non-stationary parts of the signal was considered. The proposed adaptive bit allocation method in this study, switches between two different bit allocations for stationary and non-stationary windows based on the variance of the signal in each window. The results of this adaptive bit allocation transform coding method compared to the other bit allocation methods (Adaptive-Predefined and Constant with similar pdf) at different bit rates are shown in Fig. 5.

As it can be observed in Fig. 5, the adaptive bit allocation

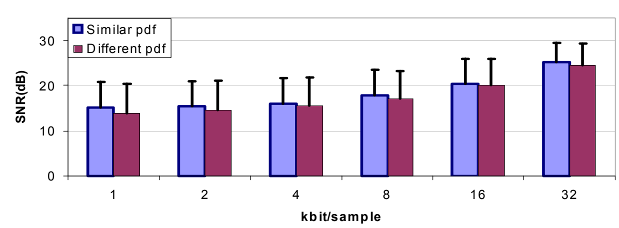

Fig. 4. Comparison of different bit allocation method.

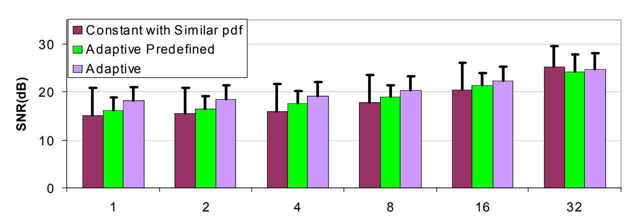

Fig. 5. Results of transform coding of respiratory and swallowing sounds with different bit allocation methods.

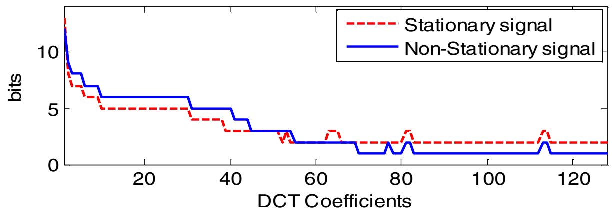

Fig. 6. Predefined adaptive bit allocations of stationary and non-stationary signals.

method improves the SNR of the constant bit allocation method by about 3dB at bit rates below 32 kbit/sample but their performances were similar at 32 kbit/sample. However, the standard deviation of the adaptive method was consistently smaller than that of the constant bit allocation method, indicating the adaptive method is more robust to the changes among the subjects.

In [7] it was shown that using adaptive bit allocation method improves the performance by about 4 dB. However, the adaptive bit allocation method proposed in [7] requires updating the bit allocation in each window based on the changes in the DCT coefficients of the signal; hence it is time consuming. On the other hand, the proposed bit allocation method in this study switches between two different bit allocations based on the variance of the signal in each window. This makes our method to be much faster, as it requires much less computations compared to that proposed in [7] and yet achieves a similar improvement (3 dB improvement in SNR compared to that of the constant bit allocation).

In the last experiment, the robustness of the proposed adaptive DCT coding system with respect to the changes in the bit allocations among different subjects were investigated by using the average bit allocation of 5 subjects' signals and apply it to the rest of the signals. The goal of this experiment was to investigate the possibility of using two predefined bit allocations that is adaptively being assigned to stationary and non-stationary parts of the signal, regardless of the subject. The results, shown in Fig. 6, reveal that for the non-stationary parts most of the bits were assigned to the first DCT coefficients, while for the stationary parts the bits were allocated more uniformly.

The SNR results of the predefined adaptive bit allocation method for different values of bit rate, shown in Fig. 5, indicate that the SNR values of this method are only slightly smaller than that of the adaptive method. Note that the proposed adaptive method adapts to the given subject and does not require off-line training. On the other hand, the predefined adaptive method requires off-line training but it is faster for online coding and transmission. Also, the performance of the predefined adaptive bit allocation method is better than that of the constant bit allocation method in spite of using similar bit allocations for all subjects. These results assert the necessity of using different bit allocations for stationary and non-stationary portions of the signal.

IV. Conclusion

In this study the compression of respiratory and swallowing sounds was investigated. The proposed compression method is based on the DCT coding of the signal. Since the significance of the coefficients is not equal, different constant bit allocations were examined. Then, considering the changes in the stationary characteristics of the signal, an adaptive bit allocation method was sought. The proposed method switches between two bit allocations based on the variance of the signal in each window. The results indicate that the SNR of the adaptive bit allocation method is about 3dB higher than that of the constant bit allocation method. Also, comparing with the other adaptive bit allocation methods used for speech signals, this method is faster and requires less computation. The average SNR of the adaptive method for compressing the respiratory and swallowing sounds was found to be 24.80 ± 3.38 dB.

Finally, the robustness of the proposed method to the changes in the bit allocations among different subjects was investigated. It was shown that using two constant predefined bit allocations for all of the subjects does not affect the performance of the adaptive method greatly. Therefore, the proposed adaptive method has the ability to work in a fully automated mode without the need to be adjusted for every signal separately; hence, it is faster for online coding applications.

References

[1] A. Cohen, and Y. Zigel, "Compression of multichannel ECG through multichannel long-term prediction ", IEEE Engineering in Medicine and Biology Magazine, Vol. 17, No. 1, pp109–115, 1998.
[2] R.S.H. Istepanian, L.J. Hadjileontiadis, and S.M. Panas, "ECG data compression using wavelets and higher order statistics methods", IEEE Trans. on Information Technology in Biomedicine, Vol. 5, No. 2, pp108-115, 2001.
[3] H. Wen-Jyi, C. Ching-Fung, and L. Kuo-Jung, "Scalable medical data compression and transmission using wavelet transform for telemedicine applications", IEEE Trans. on Information Technology in Biomedicine, Vol. 7, No. 1, pp54-63, 2003.
[4] E.S.G. Carotti, J.C. De Martin, D. Farina, and R. Merletti," Linear predictive coding of myoelectric signals", Proceedings of ICASSP'05, Vol. 5, pp629-632, 2005.
[5] A.S. Spanias, "Speech Coding: A Tutorial Review", Proceedings of the IEEE, Vol. 82, No. 10, pp 1539-1582, 1994.
[6] J.M. Tribolet, and R.E. Crochiere, "Frequency Domain Coding of Speech", IEEE Trans. on ASSP, Vol. ASSP-27, No. 5, pp. 512-530, 1979.
[7] R. Zelinski, and P. Noll, "Adaptive Transform Coding of Speech Signals", IEEE Trans. on ASSP, Vol. ASSP-25, No. 4, pp. 299-309, 1977.
[8] R. Zelinski, and P. Noll, "Approaches to Adaptive Transform Speech Coding at Low Bit Rates", IEEE Trans. on ASSP, Vol. ASSP-27, No. 1, pp. 89-95, 1979.
[9] R.V. Cox, and R.E. Crochiere, "Real-Time Simulation of Adaptive Transform Coding", IEEE Trans. on ASSP, Vol. ASSP-29, No. 2, pp. 147-154, 1981.
[10] G. Rempel, and Z. Moussavi, "The effect of altering the consistency of food on the breath and swallow pattern of young people with cerebral palsy", *Journal of Dysphagia,* Vol. 20, No. 2. pp. 108-112, 2005.
[11] A. Gersho, and R.M. Gray, "Vector Quantization and Signal Compression", Norwell, MA:Kluwer, 1992.