T2R2 東京科学大学 リサーチリポジトリ
Science Tokyo Research Repository

論文 / 著書情報
Article / Book Information

Title	Dithered Subband Coding With Spectral Subtraction
Authors	Chatree Budsabathon, Akinori Nishihara
Citation	IEICE Trans. Fundamentals., Vol. E89-A, No. 6, pp. 1788-1793
Pub. date	2006, 6
URL	http://search.ieice.org/
Copyright	(c) 2006 Institute of Electronics, Information and Communication Engineers

Powered by T2R2 (Science Tokyo Research Repository)

IEICE TRANS. FUNDAMENTALS, VOL.E89–A, NO.6 JUNE 2006

PAPER

Dithered Subband Coding with Spectral Subtraction

Chatree BUDSABATHON[†a)], *Student Member and* **Akinori NISHIHARA**[†b)], *Fellow*

SUMMARY In this paper, we propose a combination-based novel technique of dithered subband coding with spectral subtraction for improving the perceptual quality of coded audio at low bit rates. It is well known that signal-correlated distortion is audible when the audio signal is quantized at bit rates lower than the lower bound of perceptual coding. We show that this problem can be overcome by applying the dithering quantization process in each subband. Consequently, the quantization noise is rendered into a signal-independent white noise; this noise is then estimated and removed by spectral subtraction at the decoder. Experimental results show an effective improvement by the proposed method over the conventional one in terms of better SNR and human listening test results. The proposed method can be combined with other existing or future coding methods such as perceptual coding to improve their performance at low bit rates.

key words: *dither, subband coding, spectral subtraction, perceptual coding*

1. Introduction

Subband coding is a well-known technique for lossy compression of wideband speech, audio, image and video information [1]–[3]. In subband coding, the input signal is decomposed into a set of band limited components, called subbands, which can be reassembled to reconstruct the original signal without error (perfect reconstruction). The signal in each subband can be quantized by different quantization levels and different coding methods to achieve better coding performace than direct coding of the original signal. The quantizer in subband coding systems always introduces a quantization error, especially when the coding bit rate is low and the quantization error correlates with the input signal. So far, in the case of audio, most successful compression methods apply perceptual audio coding, where the masking properties of human auditory system are employed for bit rate reduction [3]. The quantization noise is masked if its power spectrum is below a masking threshold at frequencies determined by the input masking signal. Obviously, this limits the lower bound of the bit rates to be controlled for transparent quality [4]. Therefore, when the quantization bit is not enough or the bit rate is lower than this bound, the quantization error will correlate with the input signal and inevitably introduce a noticeable variety of undesirable artifacts, including harmonic distortion and sudden silences at the end of sounds with decaying envelopes, that are generally very annoying [5].

In this paper, we propose a method that combines a dithered subband coding with spectral subtraction at the decoder. There are two methods of dithering subband decoding based on subtractive dithering and non-subtractive dithering [6]. In subtractive dithering, the quantization distortion is eliminated by adding a dither to the signal before the quantization process and subtracting the same dither at the decoder [7]–[9]. In a non-subtractive dithering system, there is no need to subtract the dither at the decoder. The resulting error in the dithered subband coding system becomes a signal-independent white noise which is less perceptible and easier to be removed by a noise reduction process.

Although dither can remove a distortion and change it to a signal-independent white noise, this independent noise in each subband is still high and decreases the performance in some applications. A spectral subtraction method is applied to remove this noise. The spectrum of the noise is estimated from the information of dither signal and then subtracted from the coded audio signal at the decoder. Moreover, the error after subtraction process is reduced by residual noise reduction technique. Therefore, the output noise level is minimized and the quality of the audio signal is better than the conventional undithered subband system.

This paper is organized as follows. In Sect. 2, dithered subband coding is introduced. System design and the effects of dithering are explained. In Sect. 3, spectral subtraction with residual noise reduction technique is presented. The experimental results are shown in Sect. 4. Finally, the conclusion is given in Sect. 5.

2. Dithered Subband Coding

Dithering was first introduced by Gray [7], which is widely used to avoid harmonic distortion from the quantization process. An encoder of dithered subband coding is implemented by adding dither signal before the quantization operation. The model of dithered subband encoder is shown in Fig. 1. The input signal $x(n)$ is divided into N subbands by analysis filters $H_0, H_1, \ldots, H_{N-1}$, and downsampled by the number equal to the number of channels. The dither is added to the subband signal before being coded, by quantization functions $Q_0, Q_1, \ldots, Q_{N-1}$, respectively, according to criteria specifications of each band. In particular, the number of bits per sample in each band is different. Typically, the low frequency bands that contain more information will be

Manuscript received August 29, 2005.
Manuscript revised January 17, 2006.
Final manuscript received March 13, 2006.

[†]The authors are with the Department Communications and Integrated Systems, Tokyo Institute of Technology, Tokyo, 152-8552 Japan.
a) E-mail: chatree@nh.cradle.titech.ac.jp
b) E-mail: aki@cradle.titech.ac.jp

DOI: 10.1093/ietfec/e89–a.6.1788

Copyright © 2006 The Institute of Electronics, Information and Communication Engineers

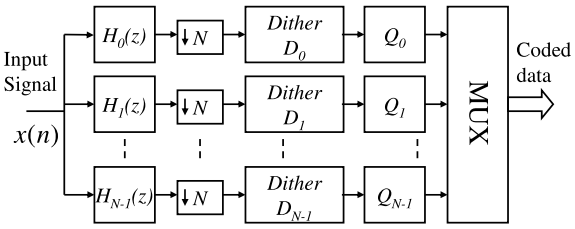

Fig. 1 Dithered subband encoder.

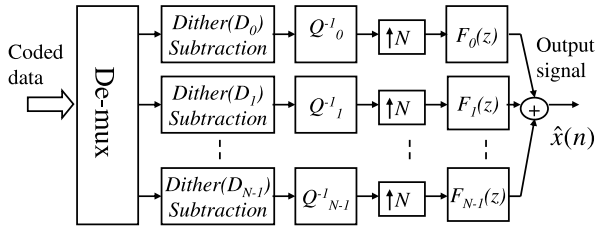

Fig. 2 Subtractive dithered subband decoder.

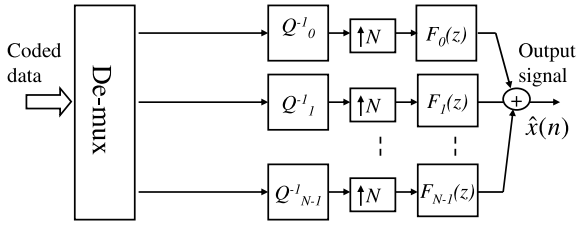

Fig. 3 Non-subtractive dithered subband decoder.

quantized with relatively more bits per sample than the high frequency bands.

The subtractive dithered subband decoder is shown in Fig. 2, where the dither signal is subtracted from the coded data before the de-quantization process, then upsampled and filtered by the synthesis filters $F_0, F_1, \ldots, F_{N-1}$. The subband signals are recombined to form an output signal $\hat{x}(n)$ that is an approximate of the original signal. Dithering effectively removes signal-dependent quantization noise in the case of coarse quantizer.

The other method is non-subtractive dithered encoder as shown in Fig. 3, where the dither signal is not subtracted at the decoder. This method will reduce signal dependent noise at the expense of random noise. The effects of this random noise are larger than in the case of subtractive dithering but better than the signal dependent noise. The noise removal can be applied on the individual subband components. Since different amount of noise is introduced in the different subbands, the noise removal operation in the individual subbands can be adapted to match the amount of noise that was introduced in each subband. This would translate into computational saving and optimum noise removal.

The dither signal itself is a random noise with certain probability density function (pdf) as shown in Fig. 4. The rectangular probability density function dither or uniform

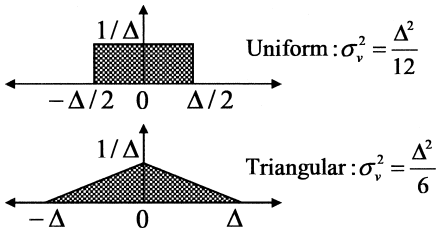

Fig. 4 Model of dither signal.

Table 1 Comparison between subtractive and non-subtractive dither.

Subtractive	Non-subtractive
Render the total error statistically independent of the system input	Render the MSE independent of the system input by using triangular dither
Not increase the noise	Increase the noise
Need dither at decoder	Implement easily

Table 2 Mean square error of dithering quantization.

MSE	Uniform pdf	Triangular pdf
Subtractive	$\frac{\Delta^2}{12}$	$\frac{\Delta^2}{12}$
Non-subtractive	$> \frac{\Delta^2}{12}$	$\frac{\Delta^2}{12} + \frac{\Delta^2}{6}$

dither is a white noise sequence having uniform distribution in the interval $[-\frac{\Delta}{2}, \frac{\Delta}{2}]$ and the triangular pdf dither can be a white noise sequence having triangular distribution in the interval $[-\Delta, \Delta]$, where Δ is the step size of the quantizer. In practice, the dither signal can be a pseudo-random sequence so that it can be generated and synchronized at both the encoder and the decoder. The mean square error (MSE) depends on the type of dithering model and the probability density function (pdf) of the dither signal. The comparison of two types of dither is shown in Table 1 and the summary of MSE between original signal and reconstructed signal when applied both types of dither is shown in Table 2.

We can see that in subtractive dithered system, both uniform dither and triangular dither can render the total error statistically independent of the input signal and achieve the same MSE ($\frac{\Delta^2}{12}$), however, in non-subtractive dithered system, we need the triangular pdf dither to render MSE of quantization noise to be independent of the system input and have a constant MSE ($\frac{\Delta^2}{12} + \frac{\Delta^2}{6}$) that is larger than in subtractive dithered system. Even subtractive dither usually show better performance in making the reconstruction error independent of the input signal [5], [7] with lower MSE but non-subtractive dither is preferred in many applications since it does not need a storage of the dither signal and is easy to implement.

In our system, the uniform dithered subtraction subband coding is chosen. The dither is a pseudo-random sequence with uniform distribution in the interval of $[-\frac{\Delta_i}{2}, \frac{\Delta_i}{2}]$, where Δ_i is the step-size of the uniform quantizer Q_i be-

ing used at subband i. The pseudo-random sequence can be generated by linear congruential algorithm or more complex algorithms. In linear congruential algorithm, the next number r_{n+1} is calculated from the current number r_n by

$$r_{n+1} = (\mathrm{A}\, r_n + \mathrm{B}) \mod \mathrm{M}, \tag{1}$$

where A and M are relatively prime numbers. In order to synchronize the dither at both the encoder and the decoder, the r_0, A, B and M are sent with the header of the first frame. The damaged or corrupted header may cause the error when synchronizing the dither signal at the decoder. To avoid this issue we can apply cyclic redundancy check (CRC) or use the same standard dither at both encoder and decoder.

The different amount of the dither are added to the different subbands so that an optimal tradeoff is achieved between bit rate and signal quality. When the subtractive uniform dither is applied to N-band filterbank with quantization step-size Δ_i at i^{th} subband, the resulting MSE of the reconstructed signal in each subband is signal-independent and exactly equal to $\frac{\Delta_i^2}{12}$. Since $\frac{\Delta_i^2}{12}$ is signal-independent so the overall quantization noise (σ_v^2) in the coded signal is also signal-independent and given by

$$\sigma_v^2 = \frac{1}{N}\sum_{i=0}^{N-1}\frac{\Delta_i^2}{12}\left[\int_{-\pi}^{\pi}|F_i(e^{j\omega})|^2\frac{d\omega}{2\pi}\right], \tag{2}$$

where $F_i(e^{j\omega})$ is the frequency response of the synthesis filter of i-th subband. Note that this is equal to the noise variance in the case of no dithering. The reconstructed signal will have none of the artifacts that occur due to correlations between the quantization noise and the input signal but the entire signal may appear the sound with the uniform noise. We noted that high-frequency bands generally have very small energy. Without dithering, these bands are usually quantized to zero, in which the quantization noise variance is equal to the signal variance itself. Dithering always introduces a noise of variance $\frac{\Delta_i^2}{12}$ so the dither should not be added in those bands.

3. Spectral Subtraction

The Spectral subtraction was first introduced by Boll [10] for a speech enhancement process. In our system, spectral subtraction is utilized to remove the signal-independent white noise after dithering process at the decoder. The block diagram of spectral subtraction is shown in Fig. 5. In the time domain, the noisy signal or reconstructed signal at the decoder $y(n)$ is composed of original signal $x(n)$ and the uncorrelated additive noise signal $n(n)$ as

$$y(n) = x(n) + n(n), \tag{3}$$

This noisy signal model can be expressed in the frequency domain as

$$Y(e^{j\omega}) = X(e^{j\omega}) + N(e^{j\omega}), \tag{4}$$

where $Y(e^{j\omega})$, $X(e^{j\omega})$, and $N(e^{j\omega})$ are the Fourier transforms

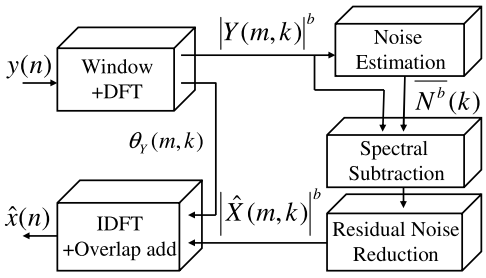

Fig. 5 Spectral subtraction method.

of $y(n)$, $x(n)$, and $n(n)$, respectively. Spectrum density function of the signal mixed with uncorrelated white noise equals to the spectrum density of the signal plus the spectrum density of the noise, hence the noise can be subtracted from the mixed signal as

$$|\widehat{X}(e^{j\omega})|^b = |Y(e^{j\omega})|^b - \overline{|N(e^{j\omega})|^b}, \tag{5}$$

where $|\widehat{X}(e^{j\omega})|^b$ is an estimate of the original signal spectrum $|X(e^{j\omega})|^b$, $\overline{|N(e^{j\omega})|^b}$ is the time-averaged noise spectra. The exponent b is set to 1 when magnitude spectral subtraction is performed and set to 2 in the case of power spectral subtraction. Spectral subtraction may be implemented in the power or the magnitude spectral domains, however, in our experimental simulations, the power spectral subtraction gave a little better results. Therefore, short time power spectral subtraction is applied in our implementation. $Y(m,k)$ is the short-time spectrum of noisy input signal $y(n)$ for each frequency bin k and each frame m, where it can be calculated by windowing the signal $y(n)$ by using a Hann or a Hamming window, and then transforming via discrete fourier transform (DFT). The $\overline{N^2}(k)$ is the time-average of estimated spectrum of noise signal at bin k. The power spectrum of estimated signal $|\widehat{X}(m,k)|$ is calculated as

$$|\widehat{X}(m,k)|^2 = \begin{cases} 0, \text{if } |Y(m,k)|^2 - \overline{N^2}(k) < 0 \\ |Y(m,k)|^2 - \overline{N^2}(k), \text{otherwise.} \end{cases} \tag{6}$$

Since human cannot detect the phase distortion of audio signal, the phase function of noisy signal $\theta_y(m,k)$ is combined with the estimated spectrum $|\widehat{X}(m,k)|$ and inversely transformed from the frequency domain into time domain by IDFT and overlap added method.

$$\widehat{x}(n) = \mathrm{IDFT}\left\{|\widehat{X}(m,k)|e^{j\theta_y(m,k)}\right\}. \tag{7}$$

3.1 Noise Estimation

In general, it is difficult to estimate the noise spectrum due to the coarse quantization step-size in every frame because this quantization noise is neither white nor uncorrelated with the input signal [5]. However, in the dithered quantizer, the quantization noise is made white and input-independent. The quantization noise spectrum can be estimated frame by

frame using information at the decoder such as the quantization step-size and the type of dither signal. We used pseudo-random sequence to generate dither, thus we can send only the initial state instead of the entire dither signal. The estimated quantization noise $n(n)$ at the output is the summation of the dither signal $d_i(n)$ convoluting with the synthesis filter bank $f_i(n)$ as

$$n(n) = \sum_{i=1}^{N} \Big[f_i(n) * d_i(n) \Big]. \tag{8}$$

The amplitude spectrum of this noise at bin k of each frame m can be calculated by windowing the output noise with Hann window then transforming it into frequency domain. The time-average power spectrum of the estimated noise at each bin k is

$$\overline{N^2}(k) = E\Big[N(m,k)^2 \Big], \tag{9}$$

where $N(m,k)$ is the estimated amplitude noise spectrum at frame m, bin k and the operation $E[x]$ means the expectation of x. The maximum residual noise is the maximum of the error between the noise power $N^2(m,k)$ and the time-average of noise power $\overline{N^2}(k)$ at frequency bin k:

$$\max \left| N_R^2(k) \right| = \max \left| N^2(m,k) - \overline{N^2}(k) \right|. \tag{10}$$

3.2 Residual Noise Reduction

It is obvious that the effectiveness of this method is dependent on the accuracy of spectral noise estimate. The better the noise estimate, the less residual noise appears in the spectrum. If the estimated noise is lower than the actual noise then the noise remains in the output signal and if the estimated noise is larger than the actual noise then the audio signal is distorted. The causes of residual noise are the variations of the instantaneous noise power spectrum around the mean, the cross product terms, and the non-linear mapping of the estimated spectra that fall below a threshold. This residual noise can be compensated by over-subtraction technique [11] and by averaging the residual noise with the adjacent frames. The general expressions of a spectral subtraction with over-subtraction are given by

$$|\widehat{T}(m,k)|^2 = |Y(m,k)|^2 - \alpha \overline{N^2}(k) \tag{11}$$

$$|\widehat{X}(m,k)|^2 = \begin{cases} |\widehat{T}(m,k)|^2 \text{ if } |\widehat{T}(m,k)|^2 > \beta \overline{N^2}(k) \\ \beta \overline{N^2}(k), \quad \text{otherwise,} \end{cases} \tag{12}$$

where $\alpha > 1$ minimizes the appearance of negative values that generate spectral spikes, and $0 < \beta \ll 1$ sets a spectral floor which reduces the perception of musical noise. Selection of over-subtraction factor α is also an important issue. The higher the amount of α is, the stronger components are attenuated, resulting in better noise suppression. However, too strong over-subtraction will result in over-suppression of components in the original signal and therefore it will introduce more distortions. The optimal value of α can be set as a function of the segmental noisy signal to noise ratio (NSNR) so that high SNR frames need less compensation than low SNR frames. The NSNR is calculated for every frame as

$$NSNR\,(\text{dB}) = 10 \log_{10} \frac{\sum_{k=1}^{K} |Y(k)|^2}{\sum_{k=1}^{K} \overline{N^2}(k)}. \tag{13}$$

The over-subtraction factor α can be calculated [12] as

$$\alpha = \alpha_0 - \frac{3}{20} NSNR, \tag{14}$$

for $-5\,\text{dB} \le NSNR \le 20\,\text{dB}$. α_0 is the α at 0 dB which is determined to be 4. K is the number of frequency bins. Residual noise can be suppressed additionally by averaging the power spectrum or replacing its current value, at a given frequency bin, with its minimum value chosen from the adjacent analysis frames.

$$|\widehat{X}(i,k)|^2 = \min \left\{ |\widehat{X}(j,k)|^2 \right\}, j = i-1, i, i+1,$$
$$\text{when} \quad |\widehat{X}(i,k)|^2 < \delta \max |N_R^2(k)|. \tag{15}$$

This process is done only if $|\widehat{X}(i,k)|^2$ is less than the maximum residual noise multiplied with a correction factor δ. The δ varies from 0.3 to 1.0 to prevent distortions of the audio signal.

4. Experiments and Results

The proposed technique is combined with the conventional subband coder whose specifications closely resemble those for MPEG-1 layer I (MP-1) [13]. Figure 6 shows the implemented system that includes a 32-channel QMF filterbank, a constant bit allocator, subtractive dithering quantizer and spectral subtraction. The audio signal is divided into 32 equal width subband streams in the frequency domain by 512-coefficient analysis window $C(n)$. The 32 new samples are shifted into the buffer of 512 PCM (Pulse Code Modulation) samples $x(n)$ in every computation cycle. The output $s(i)$ of each filter i for $i = 0$ to 31 can be written as

$$s(i) = \sum_{k=0}^{63} M(i,k) \sum_{j=0}^{7} C(i+64j) x(i+64j), \tag{16}$$

where analysis matrix $M(i,k)$ is defined by

$$M(i,k) = \cos \left[\frac{(2i+1)(k-16)\pi}{64} \right]. \tag{17}$$

A uniform dither is added to every subband except the bands which are allocated zero bits. The same dither is subtracted and quantization noise spectrum is estimated at the decoder using the information of bit allocation and type of dither signal from the encoder. The estimated noise spectrum is set to zero for subbands with zero quantization bits allocated. The analysis window length of the spectral subtraction must be chosen to ensure a good frequency resolution and to prevent smearing of signal transients. In our simulation, the analysis window is a Hann window of 2048 samples (about 50 ms

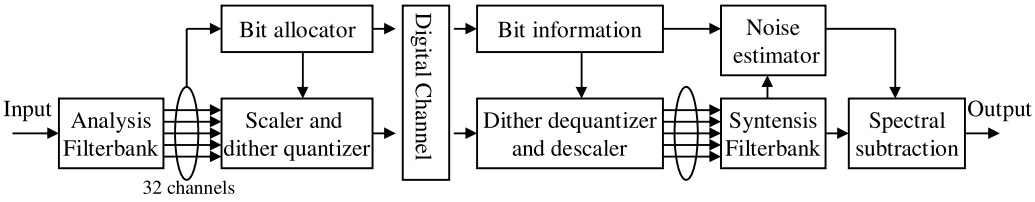

Fig. 6 Dithered subband coding system implementation.

at 44.1 kHz sampling rate) with 1024 samples overlapping. β is set to 0.03 and δ is set to 0.6 by listening test. The test signals are male speech, female speech, Japanese music, and English music from audio Compact Disk (CD) at 0.7056 Mbps/channel (16 bits/sample, 44.1 kHz). The conventional MP-1 system (without dither and spectral subtraction) and the proposed system are compared by encoding the test signals at low bit rates (less than 64 kbps/channel). The objective measure in terms of the segmental SNR is used to evaluate the signal quality enhancement. The segmental SNR is given by the average of the SNR in each frame of the signal as

$$SNR_{seg} = \frac{1}{M}\sum_{m=0}^{M-1} 10\log_{10}\frac{\sum_{k=1}^{K} x^2(m,k)}{\sum_{k=1}^{K}\left(x(m,k)-\widehat{x}(m,k)\right)^2}, \quad (18)$$

where $x(m,k)$ and $\widehat{x}(m,k)$ are respectively the original and the decoded audio signals at the m-th frame.

From the simulation, our proposed method can improve the segmental SNR of encoded signal about 0 dB to 3 dB compared to the conventional one. The spectrograms of the original music signal, the coded signal by the conventional method, and by the proposed method are shown in Fig. 7, Fig. 8, and Fig. 9, respectively. The audio is coded at low bit-rate so that the signal in the high frequency band (low information) is quantized into zero. Therefore, there is no signal spectrum in high frequency band. This is also similar with the resulting signal coded by the conventional MP-1. We can see that in Fig. 8, the coding bit rate is very low until the available bit is not enough for those bands then the signal spectrum at the middle frequency range in the coded signal is scattering and sounds like a musical noise. This is clearly perceptible and very annoying. Figure 9 clearly shows that the proposed system can reduce the artifact from quantization noise and enhance the signal quality. The signal spectrum becomes smoother and the quantization noise is less perceptible compared with the conventional MP-1.

The subject tests were also carried out to confirm the effective improvement by the proposed method. Seven university students were trained to get familiar with quality degradation in coded audio signal. In the first set of tests, after the listeners listened to the original test signal, the coded signals by the conventional and the proposed method were presented randomly. Therefore, listeners did not know which one is conventional or proposed method. Then they were asked to select the signal that they thought better between both of them. They could answer that they have the same quality. The results are averaged and given in Table 3,

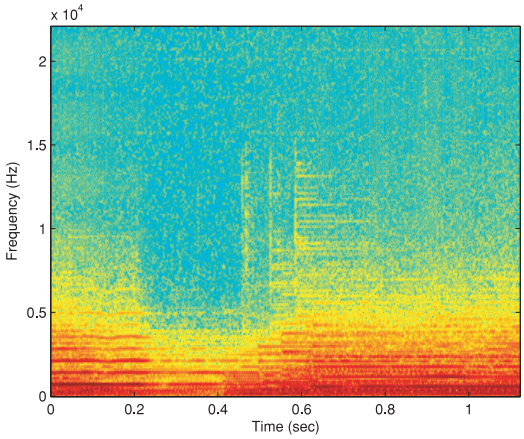

Fig. 7 Spectrogram of original music signal.

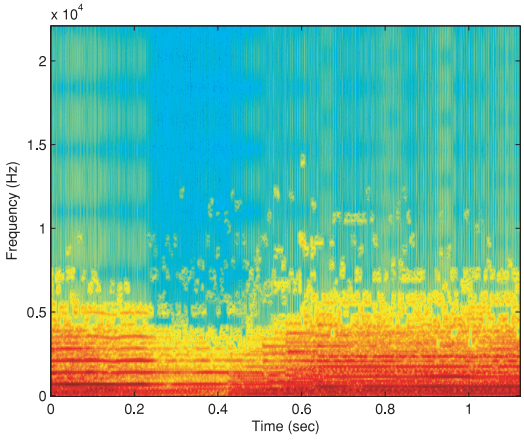

Fig. 8 Spectrogram of coded signal by MP-1.

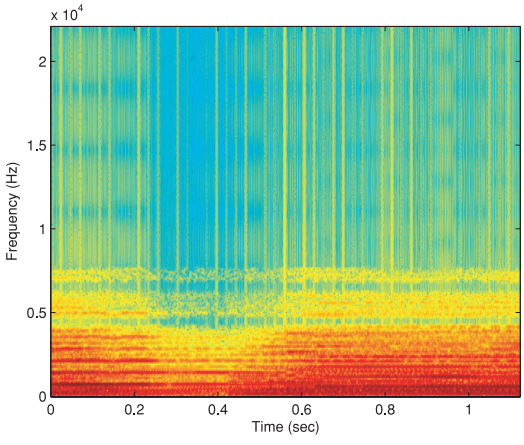

Fig. 9 Spectrogram of coded signal by proposed method.

Table 3 Results of human listening tests.

Audio Signal	Average probability of listener thought that proposed encoded signal is better than conventional encoded signal
Pop music	1.0
Rock music	0.86
Male speech	0.71
Female speech	0.71

Table 4 Average score of MP-1 and the proposed method for 2 audio signal at 32 kbps and 64 kbps.

Audio Signal	Average score	
	MP-1	Proposed
Pop music (32 kbps)	1.57	2.86
New Age music (32 kbps)	1.43	3.00
Pop music (64 kbps)	2.43	3.57
New Age music (64 kbps)	2.14	3.57

Note: All the differences with 95% confidence level.

which shows the superiority of the proposed method. The other tests were performed to support the results by the same group. The test signals were played by high quality headphones. The listeners were asked to rate the quality of each signal by scoring between one (very low quality) and five (very high quality) similar to mean opinion score (MOS) scale [14]. The sample signals are music coded at 32 kbps and 64 kbps. The averages of these scores are shown in Table 4. The music coded by the proposed method has higher average score than the same music coded by the conventional method at the same bit rate with 95% confidence level. These results clearly show an improvement of the proposed method over the conventional method.

5. Conclusion

In this paper, a dithered subband coding with spectral subtraction for a low bit rate audio coding is proposed. In the proposed method, the quantization noise is firstly rendered into signal-independent white noise by subtractive dithering and then suppressed by spectral subtraction with residual noise reduction. The results show output signal of the proposed method has higher SNR and better perceptual audio quality compared to the conventional one. The proposed system can be combined with other perceptual coding to improve their performance at low bit rates.

Acknowledgment

The first author would like to thank the Japanese government scholarship for the financial support.

References

[1] M. Vetterli and J. Kovacevic, Wavelets and Subband Coding, Prentice Hall, NJ, 1995.
[2] J.W. Woods, Subband Image Coding, Kluwer Academic Publishers, 1991.
[3] T. Painter and A. Spanias, "A review of algorithms for perceptual coding of digital audio signals," Proc. Digital Signal Processing International Conference, vol.1, pp.179–208, July 1997.
[4] G. Theile, M. Link, and G. Stoll, "Low bit rate coding of high quality audio signals," AES Conventional, Preprint, p.2431, March 1987.
[5] N.S. Jayant and P. Noll, Digital Coding of Waveform, Principle and Application in Speech and Video, Prentice Hall, NJ, 1984.
[6] R.A. Wannamaker, The theory of dithered quantization, Ph.D. thesis, University of Waterloo, Ont., Canada, 1997.
[7] R.M. Gray and T.G. Stockham, "Dithered quantizers," IEEE Trans. Inf. Theory, vol.39, no.3, pp.805–812, May 1993.
[8] B. Denckla, "Subtractive dither for internet audio," J. Audio Eng. Soc., vol.46, no.7/8, pp.654–655, July/Aug. 1998.
[9] T. Chen, "Elimination of subband coding artifacts using the dithering technique," Proceeding ICIP-04 IEEE International Conference, vol.2, pp.874–877, Nov. 1994.
[10] S.F. Boll, "Suppression of acoustic noise in speech using spectral subtraction," IEEE Trans. Acoust. Speech Signal Process., vol.27, no.2, pp.113–120, April 1979.
[11] M. Berouti, R. Schwartz, and J. Makhoul, "Enhancement of speech corrupted by acoustic noise," Proc. ICASSP, vol.4, pp.208–211, April 1979.
[12] R.M. Udrea and S. Ciochina, "Speech enhancement using spectral over-subtraction and residual noise reduction," Proc. SCS 2003, vol.1, pp.165–168, July 2003.
[13] D. Pan, "A tutorial on MPEG/audio compression," IEEE Trans. Multimed., vol.2, no.2, pp.60–74, summer 1995.
[14] S.R. Quackenbush, T.P. Barnwell III, and M.A. Clements, Objective Measures of Speech Quality, Prentice-Hall, Englewood Cliffs, NJ, 1988.

Chatree Budsabathon was born in Bangkok, Thailand, on December 1, 1979. He received the B.Eng. degree (with first class honors) from Chulalongkorn University, Bangkok, Thailand, in 2000. Since 2000, he has received the Japanese Government Scholarship and has continued his education in Japan as the research student. He received the M.Eng. degree in Communications and Integrated Systems Engineering from Tokyo Instituted of Technology, Tokyo, Japan, in 2003. He is currently pursuing the Ph.D. course at Tokyo Instituted of technology, engaging in the research on digital signal processing techniques for multimedia and communication.

Akinori Nishihara received the B.E., M.E. and Dr. Eng. degrees in electronics from Tokyo Institute of Technology in 1973, 1975 and 1978, respectively. Since 1978 he has been with Tokyo Institute of Technology, where he is Professor of the Center for Research and Development of Educational Technology. His main research interests are in signal processing, and its application to educational technology. He served as an Associate Editor of the IEICE Trans. Fundamentals from 1990 to 1994, an Associate Editor of the IEEE Transactions on Circuits and Systems II from 1995 to 1997, and Editor-in-Chief of Trans. IEICE Part A from 1998 to 2000. He served as an ExCom member of IEEE Region 10 (Asia Pacific Region), as Student Activities Committee Chair, Treasurer, and Educational Activities Committee Chair. He now serves as a member of IEICE Strategic Planning Committee, a member of IEEE EAB Committee of Global Accreditation Activities, Chair of IEEE Circuits and Systems Society Japan Chapter, and adviser of IEEE Japan Council Women in Engineering Affinity Group. He received the IEICE Best Paper Award in 1999 and the IEEE Third Millennium Medal in 2000. Dr. Nishihara is a Fellow of IEEE, a member of EURASIP, ECS, JSET and IEEK.