# A 128 kb/s Hi-Fi Audio CODEC Based on Adaptive Transform Coding with Adaptive Block Size MDCT

Masahiro Iwadare, *Member, IEEE*, Akihiko Sugiyama, *Member, IEEE*, Fumie Hazu, Akihiro Hirano, *Member, IEEE*, and Takao Nishitani, *Member, IEEE*

*Abstract*—**A hi-fi audio CODEC with an improved adaptive transform coding (ATC) algorithm is developed using digital signal processors (DSP's). An audio signal with a 20 kHz bandwidth sampled at 48 kHz is coded at a rate of 128 kb/s. The algorithm utilizes adaptive block size selection, which is effective for pre-echo suppression. A modified discrete cosine transform (MDCT) with a simple window set is employed to reduce block boundary noise without decreasing the performance of transform coding. In addition, a fast MDCT calculation algorithm, based on a fast Fourier transform, is adopted. Weighted bit allocation is employed to quantize the transformed coefficients. The CODEC is realized by a multiprocessor system composed of newly developed DSP boards. Subjective tests with the CODEC show that the coding quality is comparable to that of compact disc signals.**

## I. Introduction

RECENTLY, audio signal coding has been widely studied for integrated services digital network (ISDN) and digital recording media applications [1]–[4]. One of the current standardization activities for hi-fi audio coding is being carried out by the Moving Picture Expert Group of the International Standards Organization (ISO/MPEG) [5] for next-generation digital recording media. The main target bit rate is 128 kb/s for a monophonic audio signal with a 20 kHz bandwidth.

Adaptive transform coding (ATC) [6]–[8] is one of the most promising algorithms for hi-fi audio coding. The transformed coefficients are quantized according to their bit assignments, which are determined adaptively by using the coefficient power. Therefore, redundancy in signals is effectively removed and a high compression ratio is achieved. However, conventional ATC with discrete cosine transform (DCT) involves quality problems due to the following two coding errors: a pre-echo caused in the silent period before an abrupt increase of the signal magnitude, or an attack, and noise at block boundaries degrading the subjective quality. In addition to these problems, the decrease in the performance of transform coding [6] is serious for signals with a widely spread spectrum.

Manuscript received September 1990; revised August 10, 1991.

The authors are with the C&C Systems Research Laboratories, NEC Corporation, Kanagawa, 216 Japan.

IEEE Log Number 9104245.

This paper presents a new ATC CODEC that does not suffer from these problems. A modified DCT (MDCT) with adaptive block size has been introduced to reduce pre-echo and block boundary noise. A shorter block size is selected during nonstationary periods in order to make use of the backward masking effects of human ears [9]. As the decrease in transform coding performance cannot be avoided with flat-spectrum signals, weighted bit allocation, based on the simultaneous masking effects [10], is employed only for quantizing the flat spectrum signal coefficients. In Section II, the problems of conventional ATC are demonstrated with examples. The features of the new coding algorithms are presented in Section III. Hardware implementation is described in Section IV. Finally, the coding quality for various kinds of signals is evaluated in Section V.

## II. ATC with a Fixed-Block-Size DCT

A blockdiagram for a conventional ATC with a fixed-block-size DCT [8] is shown in Fig. 1. Input samples are buffered and a block of samples is transformed to frequency domain coefficients by a DCT. Based on these coefficients, the smoothed spectrum is calculated and transmitted as side information. Bit assignment is accomplished in proportion to the spectrum power. Based on the assignment, the transformed coefficients are quantized and transmitted as main information.

Fig. 2 shows an example of source and reconstructed signals by ATC, when a fixed-block-size DCT was used. Fig. 2(a) shows the waveform for a source generated from a drum. Segments A, B, C, and D in this figure are blocks of 256 samples. In Fig. 2(b), segments A through D are coded and decoded as a single block of 1024 samples. In Fig. 2(c), segments are processed as four independent blocks. Comparing Fig. 2(a) and (b), it is obvious that noise components have been added to the decoded signal in segments A, B and the first half of C. These noise components are due to quantization errors in the frequency domain and result in pre-echoes. The pre-echo disappears in segments A and B when a block size of 256 is applied, as in Fig. 2(c). As individual segments A through D are processed independently, segment C, where

0733-8716/92$03.00 © 1992 IEEE

Roku, Inc. EX1024
U.S. Patent No. 6,339,757
Authorized licensed use limited to: George Mason University. Downloaded on February 23,2023 at 00:54:33 UTC from IEEE Xplore. Restrictions apply.

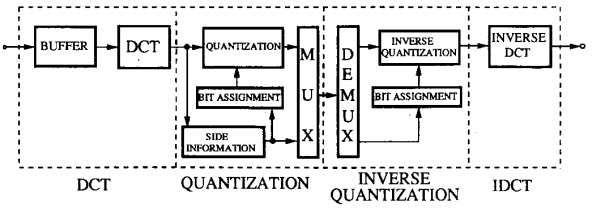



Fig. 1. Conventional ATC block diagram.

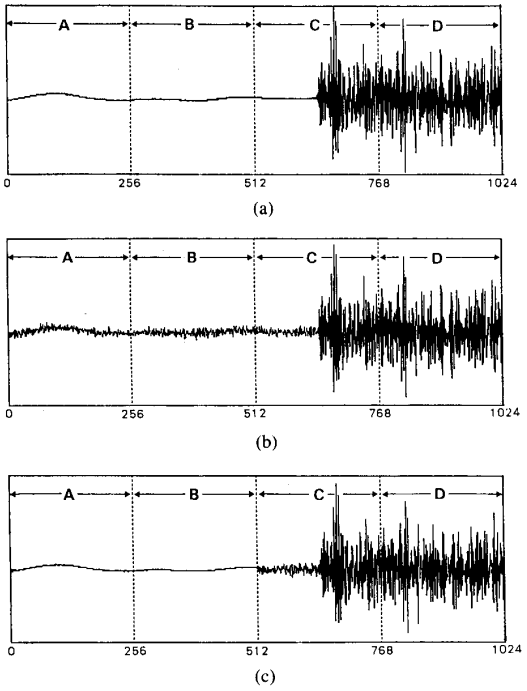



Fig. 2. Source and reconstructed signals by conventional ATC. (a) Source signal. (b) Reconstructed signal with a block size of 1024. (c) Reconstructed signal with a block size of 512.

the attack is located, is processed separately, thereby affecting no other blocks.

If the block size is set small enough, the loss of quality due to pre-echoes can be reduced. This is beacuse pre-echoes are masked by attacks owing to the backward masking effects of human ears when the periods for pre-echoes are short [9]. Experiments have shown that the time period corresponding to the block size should be below 5 ms. On the other hand, when the block size is small, the performance of transform coding decreases and the coding quality is degraded [6] for most signals. The reason is that the bit rate supplied for encoding the main information is reduced for a small block size due to the overhead required for side information. The tradeoff between the block size and coding quality is a design problem in the conventional ATC.

The noise at block boundaries, which is inherent in block operations, also degrades the coding quality. As each block is encoded independently, the quantization error affects successive blocks differently. As a result, a large noise is sometimes caused at the block boundaries. To reduce the noise, the overlapped window technique [11] has been commonly used. However, the performance of transform coding decreases because the overlapped portion is transformed twice. Therefore, a modified DCT (MDCT) [12] has been introduced for block boundary noise reduction without the performance decrease of transform coding. This MDCT, however, cannot cope with the tradeoff problem between the block size and coding quality.

The coding quality degradation with signals having a wide spectrum is a serious problem for high-quality audio coding. As assigned bits are distributed to a wide frequency range depending on the spectrum, the performance of transform coding decreases for those signals [6]. The degradation is clearly perceived in coding cabasa sources.

## II. ATC with Adaptive Block Size MDCT

### A. System Configuration

A new ATC algorithm features adaptive block size MDCT. This adaptive block size selection [13] is effective for reducing pre-echoes. As the original MDCT proposed by Princen *et al.* [12] assume a fixed block size, a new set of windows is equipped to apply adaptive block size selection.

A block diagram of the proposed algorithm is presented in Fig. 3. Input samples are buffered and an appropriate block size is selected after the signal characteristics have been evaluated in the time domain. The selected number of samples is then transformed into coefficients in the frequency domain using the MDCT. The coefficients are averaged and decimated for transmission, as side information. The spectrum distribution is calculated using this information. Simultaneous psychoacoustic weighting is applied in bit assignment only for flat-spectrum signals in quantizing transformed coefficients in order to retain as much signal fidelity as possible.

### B. Adaptive Block-Size Selection Algorithm

Adaptive block-size selection improves the coding quality by selecting a larger block size for pseudostationary parts of an input signal to achieve high transform coding gain. It chooses a smaller block for nonstationary parts to reduce pre-echoes due to backward masking effects. If the block size were determined by segmental signal-to-noise ratios (SNR's) for all possible combinations of block sizes, high coding quality would be achieved. However, this method increases the computation amount required for the local decoding process in the coder to enable evaluating the SNR's. Therefore, a simplified but effective block-size selection algorithm has been developed [13]. A restriction on block size combinations is first settled. The largest candidate block size is defined as a unit block, which is divided into subblocks equal in size. Each unit block is composed of a single block of the largest size, or

Authorized licensed use limited to: George Mason University. Downloaded on February 23,2023 at 00:54:33 UTC from IEEE Xplore. Restrictions apply.

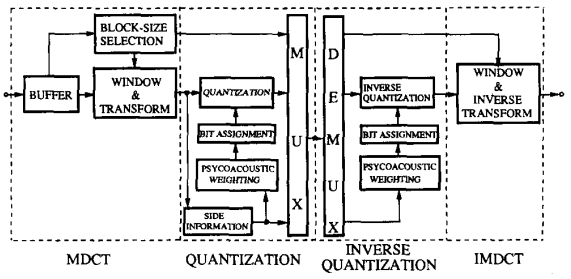



Fig. 3. A block diagram of a newly developed ATC.

two subblocks of the next largest size, and so on, assuming that the subblock size is decreased by a factor of 2. An example of a block size combination is shown in Fig. 4.

The block-size selection evaluates changes in the spectrum. The sum of interblock difference power in every frequency has been shown [14] to be identical to that in the time domain by using Parseval's theorem. Therefore, the interblock difference $x_{\mathrm{DIF}}(N, m)$ in the time domain is calculated for the $m$th subblock as:

$$x_{\mathrm{DIF}}(N, m) = \frac{1}{N} \sum_{j=1}^{N} \{x_N(j, m) - x_N(j, m-1)\}^2 \quad (1)$$

where $x_N(j, m)$ is the normalized $j$th time-domain sample in the $m$th subblock of block size $N$. The normalization is achieved independently according to the variance in each subblock. The parameter $m$ ranges from 1 to the number of subblocks $M_{\max}$ in a unit block.

From among multiple interblock difference values over an entire unit block, the maximum value described by:

$$x_{\max}(N) = \max \{x_{\mathrm{DIF}}(N, m) | m = 1, 2, \cdots, M_{\max}\} \quad (2)$$

is detected. (1) and (2) are repeatedly evaluated for all the candidate block sizes from the smallest subblock size $N_{\min}$ to the largest subblock size $N_{\max}$, increasing $N$ by a factor of 2 for example. The block size $N$ giving the smallest $x_{\max}(N)$ is chosen as the block size. Thus, this algorithm selects the block size for which the changes in the spectrum are the smallest.

To detect an attack with a higher permeability, the interblock difference ratio $C_R(m)$ of the $m$th and $(m+1)$th subblock of the smallest block size $N_{\min}$ is used for this purpose. $C_R(m)$ is defined by:

$$C_R(m) = x_{\mathrm{DIF}}(N_{\min}, m+1) / x_{\mathrm{DIF}}(N_{\min}, m). \quad (3)$$

If there is any $m$ whose $C_R(m)$ or $1/C_R(m)$ exceeds a predetermined threshold $C_{\mathrm{Th}}$, the smallest block size $N_{\min}$ is selected independently from the evaluation results of the spectrum change. The threshold $C_{\mathrm{th}}$ is determined as $2^6$ by experiments.

### C. Window Condition for Adaptive Block Size MDCT

To reduce block boundary noise without decreasing the performance of transform coding, MDCT based on time-domain aliasing cancellation [12] is employed. The definition of MDCT for the block size $N$ is given by the following equations.

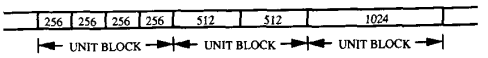



Fig. 4. Block size combination.

$$X(k) = \sum_{n=0}^{N-1} x(n) \cdot h(n) \cdot \cos\left\{\frac{2\pi}{N} \cdot \left(k + \frac{1}{2}\right) \cdot (n + n_0)\right\} \quad (4a)$$

$$y(n) = \frac{4}{N} \cdot h(n) \cdot \sum_{k=0}^{N/2-1} X(k) \cdot \cos\left\{\frac{2\pi}{N} \cdot \left(k + \frac{1}{2}\right) \cdot (n + n_0)\right\} \quad (4b)$$

$$k = 0, 1, \cdots, \frac{N}{2} - 1, \quad n = 0, 1, \cdots, N-1,$$

$$n_0 = \frac{N}{4} + \frac{1}{2}$$

$$\hat{x}(n) = y'\left(n + \frac{N}{2}\right) + y(n), \quad n = 0, 1, \cdots, \frac{N}{2} - 1 \quad (4c)$$

where $X(k)$ is the transformed coefficient, $\hat{x}(n)$ is the reconstructed signal, $y(n)$ is the inverse transformed time domain sample, $y'(n)$ is the inverse transformed sample in the previous block, and $h(n)$ is a window function.

When the symmetry for window function $h(n)$ is assumed, the condition for a window function $h(n)$ is:

$$h(i)h(i) + h(i + N/2)h(i + N/2) = 1. \quad (5)$$

This condition indicates that the analysis/synthesis window for MDCT is 50% overlapped. In the analysis filter, every input sample is transformed twice using the 50% overlap window. However, the following symmetry at the coefficients $X(k)$ is derived from (4a), when $N$ is quadruple:

$$X(k) = -X(N - 1 - k). \quad (6)$$

Therefore, only half of the coefficients should be transmitted. Thus, there is no reduction in the transform coding performance due to the fact that the overlapped portion is transformed twice.

In order to realize perfect reconstruction using block-size adaptation with MDCT, special attention must be paid to the design of the overlapping window shape. The conditions for MDCT windows of different block sizes have been studied in detail in [15]. As the window shape should be decided depending only on the current block size for simple implementation, a possible window function $h_N(n)$

Authorized licensed use limited to: George Mason University. Downloaded on February 23,2023 at 00:54:33 UTC from IEEE Xplore. Restrictions apply.

for the block size $N$ is given by the following equations:

$$h_N(n) = \begin{cases} 0 & \begin{cases} n = 0, \cdots, \frac{1}{4}N - L - 1, \\ n = \frac{3}{4}N + L, \cdots, N - 1, \end{cases} \\ \cos\left(\frac{\pi \cdot (n - N/4 - L)}{4L}\right) & n = \frac{1}{4}N - L, \cdots, \frac{1}{4}N + L - 1, \\ \cos\left(\frac{\pi \cdot (n - 3N/4 + L)}{4L}\right) & n = \frac{3}{4}N - L, \cdots, \frac{3}{4}N + L - 1, \\ 1 & n = \frac{1}{4}N + L, \cdots, \frac{3}{4}N - L - 1, \end{cases} \tag{7}$$

where $L$ stands for the half of the substantial overlapped length for the shoulder part defined by cosine function. The shoulder shape is the same for all the window sizes. The other parts are 0 or 1. The window functions defined by (7) are shown in Fig. 5 for $N = 256, 512, 1024$, $L = 64$.

As the shape for each block size is unique independent of the adjacent block size, the transform can be accomplished without waiting for the decision to be made regarding the next block size. This leads to a short coding delay and easy hardware implementation.

### D. Fast MDCT Calculation

The forward MDCT definition is shown in (4a). The cosine function in (4a) can be broken down into data shift, preprocessing, fast Fourier transform (FFT), and postprocessing, using the cosine function symmetry. A fast MDCT calculation algorithm is summarized as follows [16].

*1) Window for Transform:* Multiply the input signals by the window function.

$$xh(n) = x(n)h(n), \qquad n = 0 \cdots N - 1 \tag{8}$$

where $N$ stands for the current block size.

*2) Data Shift:*

$$x2(n) = -xh(n + 3N/4), \qquad n = 0 \cdots N/4 - 1, \tag{9a}$$

$$x2(n) = xh(n - N/4), \qquad n = N/4 \cdots N - 1. \tag{9b}$$

*3) Differential Calculation:* Calculate differences for $N/2$ samples.

$$x3(n) = x2(2n) - x2(N - 1 - 2n), \qquad n = 0 \cdots N/2 - 1. \tag{10}$$

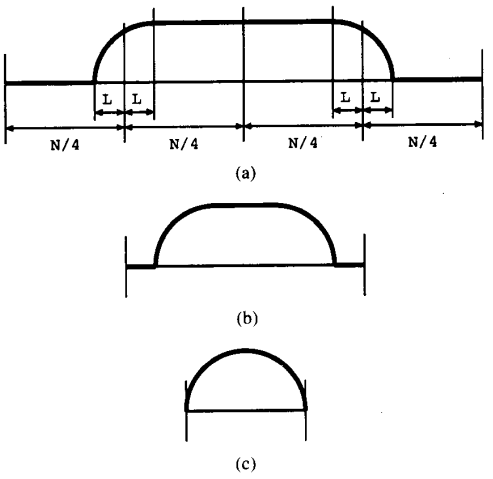



Fig. 5. Window functions for adaptive block size MDCT. (a) $N = 1024$, $L = 64$. (b) $N = 512$, $L = 64$. (c) $N = 256$, $L = 64$.

*4) Preprocessing:* Multiply coefficients for FFT.

$$x4(n) = x3(n) \exp(-2\pi jn/N), \qquad n = 0 \cdots N/2 - 1. \tag{11}$$

*5) FFT: $N/2$ Point FFT:*

$$xf(k) = \sum_{n=0}^{N/2-1} x4(n) \exp\{-2\pi jkn/(N/2)\}, \qquad k = 0, \cdots N/2 - 1. \tag{12}$$

*6) Postprocessing:* Multiply coefficients; take real parts.

$$X(k) = \text{real}\{\exp(-\pi j(k + 1/2)/N)\, xf(k)\}, \qquad k = 0 \cdots N/2 - 1. \tag{13}$$

Using this fast calculation algorithm, the number of multiplications can be reduced from $N^2$ to $(N/2)\log_2(N/2)$. The number of additions can be reduced from $N(N - 1)$ to $N \log_2(N/2)$.

The fast calculation algorithm for the inverse MDCT is described in [16].

### E. Weighted Bit Allocation

Conventional ATC obtains good coding quality with source signals having most of their power concentrated at lower frequencies, such as a piano or an orchestra. However, when the input signals have a widely spread spectrum the bits are assigned to the wide spectrum range and the transform coding performance is decreased [6]. To prevent this quality degradation, weighted bit allocation is used.

Fig. 6 shows the effects of the weighted bit assignment with cabasa as a source. Fig. 6(a), (b), (c), and (d) illustrate MDCT coefficients, the assigned number of bits without weighting, a weighting function, and the assigned number of bits with weighting, respectively. The weights

Authorized licensed use limited to: George Mason University. Downloaded on February 23,2023 at 00:54:33 UTC from IEEE Xplore. Restrictions apply.

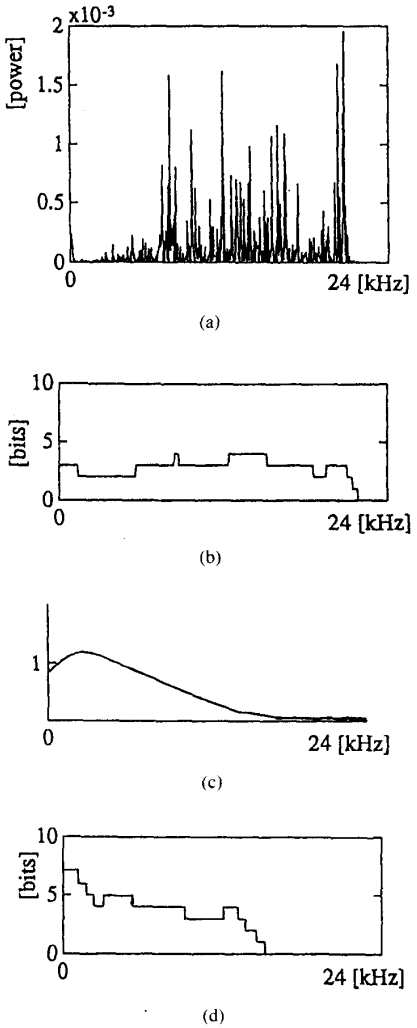



Fig. 6. Effects of the bit assignment with a psychoacoustic weighting (cabasa). (a) Transformed coefficients. (b) Bit assignment without psychoacoustic weighting. (c) A weighting function example. (d) Bit assignment with psychoacoustic weighting.

are defined based on simultaneous psychoacoustic characteristics, so that they are large for the lower frequency coefficients and small for the higher frequency ones. In this CODEC, the same function as the threshold of audibility (absolute threshold [17]) is employed as the weighting. In comparison between Fig. 6(b) and (d), bits for the higher frequency range are devoted to the lower frequency range to improve the lowband quality, which is psychoacoustically more important than that of the highband.

This modification, however, reduces the bandwidth of the decoded signals like a lowpass filter. Accordingly, the weighting is applied only to signals with a wide spectrum. The variance over the transformed spectrum is used as the measure of spectrum distribution. A small variance indicates that the spectrum is wide and nearly flat. When the variance is below the predetermined threshold, weighting

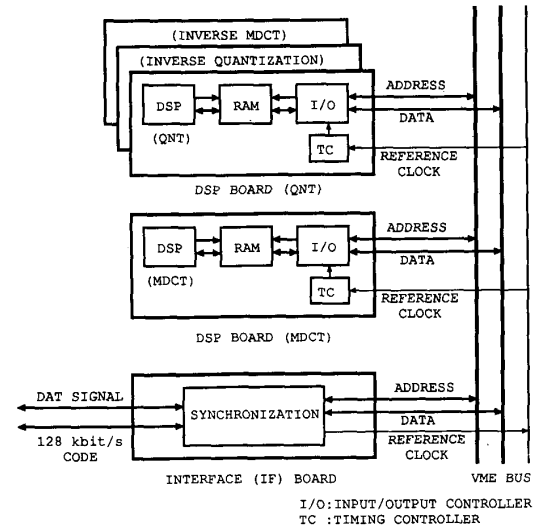



Fig. 7. Multiprocessor system configuration for the codec.

is applied to the bit assignment. Otherwise, no modification is made to the bit assignment.

## IV. Hardware Implementation

The proposed CODEC has been realized for evaluation of a variety of source signals and efficient optimization of parameters for the coding algorithm. The CODEC, based on a multiprocessor system configuration as shown in Fig. 7, is composed of an interface (IF) board and four DSP boards per channel for coding and decoding. The IF board is used for an optical and coaxial DAT digital audio interface. The board is synchronized to the digital input audio signal at a rate of 32–48 kHz and generates the DSP synchronization signal and the 128 kb/s clock used for the encoded bit stream. The board also converts the serial signal for DSP's to the digital audio signal, and vice versa.

To achieve the same high-accuracy calculation as computer simulations, a 32-bit floating-point DSP, NEC's $\mu$PD77230 [18], is employed. The capability of floating-point operations eliminates operation steps for scaling, overflow detection, and corrections. An 8$k$ word (32 bit) high-speed memory is prepared on the board as external memory, which is connected to the 32-bit high-speed VME bus for communication between the boards.

Taking into account the number of operations and functional assignment, the coding and decoding algorithms are divided into MDCT, quantization (QNT), inverse quantization and inverse MDCT, and loaded into every DSP board. Between the boards for MDCT and QNT and the boards for inverse quantization and inverse MDCT, 512 transformed coefficients for an unit block and information about the selected block size are conveyed via the VME bus. The DSP's and controllers for bus output timing and addressing are synchronized to a reference clock supplied from the IF board.

Authorized licensed use limited to: George Mason University. Downloaded on February 23,2023 at 00:54:33 UTC from IEEE Xplore. Restrictions apply.

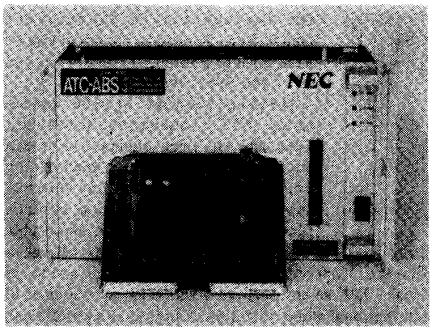



Fig. 8. Front view of the codec and a DSP board.

Fig. 8 depicts a front view of the CODEC and a DSP board.

## V. Subjective and Objective Test Results

Subjective and objective tests for ATC with adaptive block size MDCT (ATC–ABSM) have been carried out based on the hardware developed. The test signals include the sounds from castanets, a cabasa, a piano, a female voice, and a full orchestra. The sources of these signals are compact discs and the sampling frequency was digitally converted from 44.1 to 48 kHz.

For performance comparison, ATC with DCT (ATC–DCT) and subband ADPCM (SB–ADPCM) were also evaluated. The coding bit rates for those algorithms were all set at 128 kb/s for a monophonic signal. The block size $N$ for ATC–ABSM was selected from among 256, 512, and 1024, while that of ATC–DCT was fixed at 512. The coding algorithm for SB–ADPCM was the one described in CCITT Recommendation G.722 [19], which is the only standard for efficient audio coding. The bandwidth for audio signals for SB–ADPCM was limited to 15 kHz and the sampling rate was converted to 32 kHz.

Table I shows subjective evaluation test results with the mean opinion scores (MOS) [6] with loud speakers in a listening room. Source and reconstructed signals were evaluated at random sequences by 20 listeners based on a 1–5 integer point scale. The best is 5 points and the worst is 1 point.

ATC–ABSM shows better coding quality than ATC–DCT for all kinds of signals. Higher scores for castanets and a piano indicate the coding quality improvement by pre-echo suppression and block boundary noise reduction owing to the adaptive-block-size MDCT. The effects of weighted bit assignment are confirmed by a higher score by 1.5 points for a cabasa sound. In comparison of ATC–ABSM with SB–ADPCM, their coding quality is almost equivalent. However, quality degradation was perceived for a piano with SB–ADPCM due to the band limitation. As the scores of ATC–ABSM are almost the same as those of the original signals, for all the sources, the coding quality of ATC–ABSM is comparable to that of the original signals.

To objectively estimate the coding quality, computer simulation was carried out. Table II shows the averaged segmental SNR. Individual segmental SNR values were calculated every 256 samples. The results show that the SNR's for ATC–ABSM are about 0.5 dB smaller than those for ATC–DCT. This is because a smaller block size was selected in some periods to reduce pre-echoes and improve the subjective coding quality. In the comparison between ATC–ABSM and SB–ADPCM, ATC–ABSM achieved a higher SNR for a cabasa and a female voice, in spite of the fact that ATC–ABSM encodes wider bandwidth signals. This indicates that the coding efficiency of ATC–ABSM is superior to that for SB–ADPCM.

TABLE I
Subjective Evaluation Test Results (MOS)

| | Source | ATC–ABSM | ATC–DCT | S–ADPCM |
|---|---|---|---|---|
| Cabasa | 4.0 | 3.6 | 2.1 | 3.8 |
| Castanets | 4.2 | 4.0 | 3.0 | 4.0 |
| Female Voice | 4.4 | 4.4 | 4.3 | 4.4 |
| Orchestra | 3.8 | 3.5 | 3.0 | 3.6 |
| Piano | 4.5 | 4.5 | 4.1 | 4.0 |

TABLE II
Averaged SNR for ATC–ABSM, ATC–DCT, and S–ADPCM

| | ATC–ABSM | ATC–DCT | S–ADPCM |
|---|---|---|---|
| Cabasa | 10.5 | 11.0 | 9.5 |
| Castanets | 23.2 | 23.6 | 23.4 |
| Female Voice | 30.5 | 31.5 | 27.4 |
| Orchestra | 37.7 | 37.5 | 41.7 |
| Piano | 38.7 | 38.9 | 42.1 |

## VI. Conclusion

A 128 kb/s hi-fi audio CODEC with an improved ATC algorithm has been developed. The CODEC employs an adaptive block size MDCT which is equipped with a newly designed overlapped window set. The MDCT calculation is accomplished by a fast computation algorithm. The CODEC also features a weighted bit assignment algorithm based on psychoacoustic characteristics. The hardware is realized by a multiprocessor system composed of DSP boards. Subjective tests have proved that the developed CODEC achieves high coding quality with pre-echo and block boundary noise suppression. The quality degradation for flat and widely distributed spectrum signals is also reduced. The primary version of the CODEC was evaluated by ISO/MPEG. This CODEC allows high-quality audio recording and transmission over low bit rate channels.

## Acknowledgment

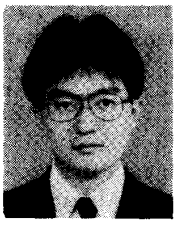
The authors would like to thank T. Ishiguro, Vice President of NEC Corporation, and B. Hirosaki, Manager of Network Research Laboratory, C&C Systems Research Laboratories (former Manager of Communications Re-

Authorized licensed use limited to: George Mason University. Downloaded on February 23,2023 at 00:54:33 UTC from IEEE Xplore. Restrictions apply.

search Laboratory), for their guidance, continuous encouragement, and invaluable comments on this paper. They are also indebted to M. Yamamoto, General Manager of C&C Systems Research Laboratories, for his encouragement and support.

## REFERENCES

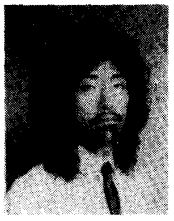
[1] K. H. Brandenburg *et al.*, "Fast signal processor encodes 48 kHz/16 bit audio into 3 bit in real time," in *Proc. ICASSP'88*, May 1988, pp. 2528-2531.
[2] J. D. Johnston, "Perceptual transform coding of wideband stereo signals," in *Proc. ICASSP'89*, May 1989, pp. 1993-1996.
[3] Y. Mahieux *et al.*, "Transform coding of audio signals using correlation between successive transform blocks," in *Proc. ICASSP'89*, May 1989, pp. 2021-2024.
[4] M. Iwadare *et al.*, "A robust 384 kbit/s stereo hifi audio codec for ISDN applications," in *Proc. GLOBECOM'89*, May 1989, pp. 1952-1956.
[5] H. G. Musmann, "The ISO audio coding standard," in *Proc. GLOBECOM'90*, Dec. 1990, pp. 0511-0517.
[6] N. S. Javant *et al.*, *Digital Coding of Waveforms*. Englewood Cliffs, NJ: Prentice-Hall, 1984.
[7] R. Zelinski and P. Noll, "Adaptive transform coding of speech signals," *IEEE Trans. Sign. Process.*, vol. ASSP-25, pp. 299-309, Aug. 1977.
[8] ——, "Approaches to adaptive transform coding at low bit rates," *IEEE Trans. Sign. Proces.*, vol. ASSP-27, pp. 89-95, Feb. 1979.
[9] L. L. Elliott, "Backward masking," *J. Acoust. Soc. Am.*, no. 34, pp. 1108-1115, 1962.
[10] M. R. Schroeder *et al.*, "Optimizing digital speech coders by exploiting masking properties of human ear," *J. Acoust. Soc. Am.*, vol. 66, pp. 1647-1652, 1979.
[11] J. M. Tribolet *et al.*, "Frequency domain coding of speech," *IEEE Trans. Sig. Proces.*, vol. ASSP-27, no. 5, pp. 512-530, Oct. 1979.
[12] J. Princen and A. B. Bradley, "Analysis/synthesis filter bank design based on time domain aliasing cancellation," *IEEE Trans. Sig. Proces.*, vol. ASSP-34, no. 5, pp. 1153-1161, Oct. 1986.
[13] A. Sugiyama *et al.*, "Adaptive transform coding with an adaptive block size (ATC-ABS)," in *Proc. ICASSP'90*, Apr. 1990, pp. 1093-1096.
[14] A. Sugiyama *et al.*, "On block-size selection algorithms for adaptive transform coding with an adaptive block size (ATC-ABS)," presented at Proc. 1990 IEICEJ Autumn Nat. Convent., Oct. 1990.
[15] T. Mochizuki *et al.*, "Constraint conditions for multiple-blocksize modified DCT," Tech. Group Rep. IEICEJ, CAS-90-9/DSP90-13, pp. 55-60, June 1990 (in Japanese).
[16] M. Iwadare *et al.*, "On a modified discrete cosine transform (MDCT) and its fast algorithm," Tech. Group Rep. IEICEJ, CAS-90-9/DSP90-13, pp. 49-54, June 1990 (in Japanese).
[17] ASPEC (Source: AT&T Bell Labs., Deutsche Thomson Brandt and Fraunhofer Gesellshaft), Doc. No. 89/205, ISO-IEC JTC1/SC2/WG8 MPEG-AUDIO, Oct. 18, 1989.
[18] T. Nishitani *et al.*, "Advanced single-chip signal processor," in *Proc. ICASSP'86*, Apr. 1986, pp. 409-412.
[19] CCITT REC. G.722, CCITT Blue Book, Melbourne, 1988.



**Masahiro Iwadare** (M'87) was born in Osaka, Japan, on February 5, 1962. He received the B.S. and M.S. degrees in electrical engineering from the University of Tokyo, Tokyo, Japan, in 1984 and 1986, respectively.

He joined the NEC Corporation, Kanagawa, Japan, in 1986, where he is currently an Assistant Manager in the C&C Systems Research Laboratories. His research interests include digital audio coding systems.

Mr. Iwadare is a member of the Institute of Electrical Engineers of Japan and the Institute of Electronics, Information and Communication Engineers.



**Akihiko Sugiyama** (M'85) received the B.Eng. and M.Eng. degrees in electrical engineering from Tokyo Metropolitan University, Tokyo, Japan, in 1979 and 1981, respectively.

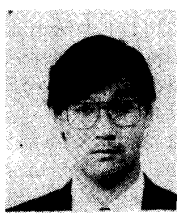
He joined the C&C Systems Research Laboratories, NEC Corporation, Kawasaki, Japan, in 1981 and has been engaged in research on production system evaluation methodologies, signal processor applications to transmission terminals, subscriber loop transmission systems, adaptive filter applications, and hi-fi audio coding. In the 1987 academic year, he was on leave at the Faculty of Engineering and Computer Science, Concordia University, Montreal, Canada, as a Visiting Scientist. Since 1989, he has been involved in the activities of ISO/MPEG-Audio for international standardization of high-quality audio data compression. His current interests lie in the areas of digital signal processing and circuit theory.

Mr. Sugiyama is a member of the Institute of Electronics, Information and Communications Engineers of Japan. He was awarded the 1987 Shinohara Memorial Academic Encouragement Award by the Institute of Electronics, Information and Communication Engineers of Japan.

**Fumie Hazu** received the B.S. degree in physics from Nihon University, Tokyo, Japan, in 1985.

She joined NEC Corporation, Kanagawa, Japan, in 1985. She is now working on digital coding of high-quality audio signals.

Ms. Hazu is a member of the Institute of Electronics, Information and Communication Engineers of Japan, and the Acoustical Society of Japan.



**Akihiro Hirano** (M'89) was born in Miyazu-shi, Kyoto, Japan, on August 1, 1964. He received the B.S. and M.S. degrees in electronics engineering from Kanazawa University, Kanazawa, Japan, in 1987 and 1989, respectively.

He joined the NEC Corporation, Kawasaki, Japan in 1989, where he is a Research Engineer in the C&C Systems Research Laboratories. He has been engaged in digital signal processing, especially echo cancellers and compilers for digital signal processors.

**Takao Nishitani** (M'82) received the B.S. and M.S. degrees in electronics engineering from Osaka University, Osaka, Japan in 1971 and 1973, respectively.

In 1973, he joined the NEC Corporation, Kawasaki, Japan, where he is currently the Senior Manager of the C&C Systems Research Laboratories. He has been engaged in management and research on speech and video bandwidth compression and on LSI signal processing architectures. He contributed to the development of the NEC signal processing chips 7720 and 77230. His ADPCM work contributed to CCITT Recommendation G.721 of a PCM-ADPCM transcoder.

Mr. Nishitani is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) of Japan. He received the 1980 Conference Paper Award from the IEICE of Japan, and the 1989 Research Award form the Japanese Government.

Authorized licensed use limited to: George Mason University. Downloaded on February 23,2023 at 00:54:33 UTC from IEEE Xplore. Restrictions apply.