Improved Lossless Audio Coding using the Noise-Shaped IntMDCT

Yoshikazu Yokotani*, Ralf Geiger†, Gerald Schuller†, Soontorn Oraintara*, and K. R. Rao*
*University of Texas at Arlington, Electrical Engineering Dept., Arlington,TX, U.S.A.
Emails: [yoshi,oraintar,rao]@uta.edu
†Fraunhofer Institute for Digital Media Technology, Ilmenau, Germany
Email: [ggr,shl]@idmt.fraunhofer.de

***Abstract*— This paper discusses approximation noise shaping to improve the efficiency of the Integer Modified Discrete Cosine Transform (IntMDCT)-based lossless audio codec. The scheme is applied to rounding operations associated with lifting steps to shape the noise spectrum towards the low frequency bands.**

In this paper, constraints on the noise shaping filter and a design procedure with the constraints are discussed. Several noise shaping filters are designed and experimental results showing the improvement are presented.

I. Introduction and Goal

The lifting scheme [1] based integer transforms are quite useful for lossless coding applications such as audio [2] and image [3] compression. These transforms are composed of approximated plane rotations, 2×2 orthogonal matrices, each of which is generally realized by two or three lifting steps associated with multiplications and rounding operations. Since every rounding operation introduces a rounding error, it is accumulated and appears as "noise floor" in the transform domain. Let us call this approximation noise. Although the noise can be cancelled by the inverse transform, for lossless coding application, it is desirable that the noise floor level is kept small since it has a significant impact on the coding efficiency.

To improve the efficiency, the multi-dimensional lifting (MDL) scheme was recently proposed [4]. This scheme significantly lowers the entire level of the noise floor by reducing the number of rounding operations as presented in [4]. In order to improve the efficiency further, reducing the noise at the high frequency bands can be considered. Typically, the signal being compressed has energy concentrated near low frequencies and decays at high frequencies. Therefore, it is possible to improve the efficiency by shaping the noise towards the low frequency bands in such a way that the noise spectrum is under the IntMDCT spectral envelope of the signal.

II. MDL Scheme-Based Stereo IntMDCT

The MDL scheme-based stereo IntMDCT [4] transforms $2N$ stereo audio samples $x_{tL}[m]$ and $x_{tR}[m]$ for $m = 0, \ldots, N-1$ into N spectral lines $X_L[k]$ and $X_R[k]$ for $k = 0, \ldots, N-1$ in the left and right channel, respectively, where $N = 1024$. The subscript L and R indicate the left and right channel, respectively. t denotes the frame number. The IntMDCT for the stereo signal, as illustrated in Fig. 1, is composed of (a) an identical sine window and time domain aliasing operation realized by the conventional three lifting steps and (b) integer discrete cosine transform of type IV (IntDCT-IV) operation realized by the MDL steps [4]. In Fig. 1 (a), only the left channel case is drawn. $cs[n]$ and $s[n]$ are the lifting coefficients where $cs[n] = (c[n]-1)/s[n]$, $c[n] = \cos\theta[n]$, $s[n] = \sin\theta[n]$, and $\theta[n] = \pi/(2N)(n+0.5)$ for $n = 0, ..., N/2-1$. e_a, e_b, and e_c is a rounding error introduced in the rounding operation associated with the first, second, and third lifting step, respectively. They are assumed to be white noise and ranged between -0.5 and 0.5, provided that all the lifting coefficients are not quantized. In Fig. 1 (b), the output of the window and time domain aliasing operation, x_{twL} and x_{twR}, are the inputs of the IntDCT-IV operation. It should be noted that the subscript t is omitted from x_{twL} and x_{twR} and the rest of the signals in Fig. 1 (b) since all the signals processed by the IntDCT-IV are in the t^{th} frame. Similar to Fig. 1 (a), the rounding errors introduced in the first, second, and third MDL step, e_1, e_2, and e_3, are assumed to have the same noise statistics.

III. Theory of the Approximation Noise Shaping Filter

This MDL scheme-based stereo IntMDCT has the same structure for the sine window and time domain aliasing operation as the conventional IntMDCT. However, as for the IntDCT-IV operation, it requires rounding operations only after each DCT-IV matrix, whereas the conventional IntMDCT requires a rounding operation after every scalar multiplication. Therefore much fewer number of rounding operations are needed, and the approximation noise can be reduced. The lowered noise level, however, is flat over the transform domain. This means that, if an input audio signal is bandlimited, only the noise occupies the bands which are outside of the bandwidth of the signal. Since the noise at the bands has to be encoded as it is, this results in a limitation in the lossless coding efficiency.

As a remedy to this problem, we apply the following conventional noise shaping scheme [5] to a rounding operation in each lifting step as illustrated in Fig.2. In this figure, $x[n]$ is an input to the lifting step which is multiplied by a scalar constant. This type of operations appears at the three lifting steps of the sine window and time domain aliasing

0-7803-8434-2/04/$20.00©2004 IEEE

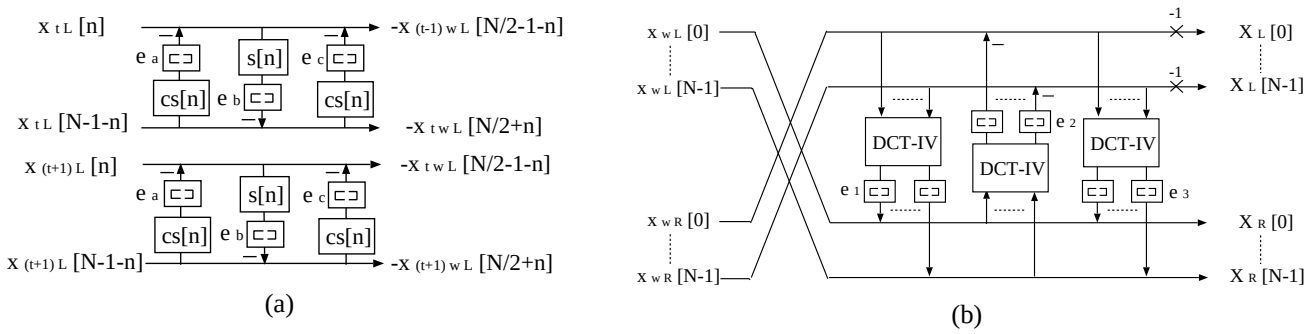


Fig. 1. The structure of the MDL scheme-based stereo IntMDCT. [] symbolizes a rounding operation. (a) the three lifting step structure for a sine window and time domain aliasing operation and (b) the MDL structure for the IntDCT-IV operation.

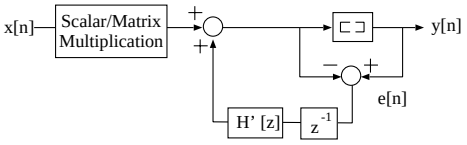


Fig. 2. A block diagram of approximation noise shaping for a lifting step. [ ] symbolizes a rounding operation.

operation. The scalar multiplication is replaced by a DCT-IV multiplication in case of MDL steps. After each multiplication, the signal is added by a filtered version of the approximation noise $e[n]$. The result is then rounded and becomes the output of the lifting step $y[n]$. The filter $h^{'}[n]$ is assumed to be causal. The approximation noise at the output is described in the z-domain as:

$$Y(z) = X(z) + (1 + z^{-1}H^{'}(z))E(z), \tag{1}$$

where $E(z)$ is the rounding noise spectrum. From (1), $E(z)$ is shaped by the filter $H(z) = 1 + z^{-1}H^{'}(z)$. Thus, reducing the noise at the high frequency bands is possible by designing $H(z)$ as a low pass filter.

Let $E[E_L^2[k]]$ and $E[E_R^2[k]]$ be the variance of approximation noise after the IntMDCT computation in the left and right channel, respectively. k indicates the spectral line index and $k = 0, \ldots, N-1$. According to [6], these variances are approximately described as follows in case of both no noise shaping and noise shaping:

Case 1:**No Noise Shaping**

$$E[E_L^2[k]] \approx \gamma[k] + \gamma_1[k] + E[e_2^2[k]], \tag{2}$$
$$E[E_R^2[k]] \approx \gamma[k] + \gamma_2[k] + E[e_3^2[k]], \tag{3}$$

where $\gamma[k]$, $\gamma_1[k]$, $\gamma_2[k]$ is an averaged variance of the approximation noise introduced in the three lifting steps in Fig. 1 (a), the first MDL step, and the second MDL step in Fig. 1 (b), respectively. They are given by (4), (5), and (6). $e_{wL}[n]$ and $e_{wR}[n]$ are the approximation noise associated with $x_{wL}[n]$ and $x_{wR}[n]$, which are illustrated in Fig. 1 (b) as the input of the IntDCT-IV computation. The variances are given by (7).

$$\gamma[k] = \frac{1}{N}\sum_{n=0}^{N-1} E[e_{wL}^2[n]] = \frac{1}{N}\sum_{n=0}^{N-1} E[e_{wR}^2[n]], \tag{4}$$
$$\gamma_1[k] = \frac{1}{N}\sum_{n=0}^{N-1} E[e_1^2[n]] = \frac{1}{12}, \tag{5}$$
$$\gamma_2[k] = \frac{1}{N}\sum_{n=0}^{N-1} E[e_2^2[n]] = \frac{1}{12}. \tag{6}$$

$$E[e_{wL}^2[n]] = E[e_{wR}^2[n]] = \begin{cases} \frac{c^2\left[\frac{N}{2}-1-n\right]+cs^2\left[\frac{N}{2}-1-n\right]+1}{12} \\ \quad \text{for} \quad n = 0, \ldots, \frac{N}{2}-1. \\ \frac{s^2\left[n-\frac{N}{2}\right]+1}{12} \\ \quad \text{for} \quad n = \frac{N}{2}, \ldots, N-1. \end{cases} \tag{7}$$

Since (4), (5), and (6) give a constant value for $k = 0, \ldots, N-1$ and the variance of e_1 and e_2 is $1/12$ constant, the approximation noise spectra E_L and E_R are approximately flat all over the IntMDCT domain.

Likewise, in case of noise shaping, $E[E_L^2[k]]$ and $E[E_R^2[k]]$ can be described as follows [6]:

Case 2:**Noise Shaping**

$$\begin{aligned} E[E_L^2[k]] &\approx |H[k]|^2(\gamma[k] + \gamma_1[k]) \\ &\quad + E[|h[n] * e_2[n]|^2] + \alpha[k](\phi[k] + \phi_1[k]) \\ &\quad + \beta[k](\psi[k] + \psi_1[k]), \end{aligned} \tag{8}$$
$$\begin{aligned} E[E_R^2[k]] &\approx |H[k]|^2(\gamma[k] + \gamma_2[k]) \\ &\quad + E[e_3^2[n]] + \alpha[k](\phi[k] + \phi_2[k]) \\ &\quad + \beta[k](\psi[k] + \psi_2[k]), \end{aligned} \tag{9}$$

where $H[k]$ is a noise shaping filter designed in the DFT of type III (DFT-III) domain. $h[n]$ is the impulse response. $*$ indicates the convolution operator and

$$\alpha[k] = Re^2(W^k H[k]) - Im^2(W^k H[k]),$$
$$\beta[k] = 2\left(Re(W^k H[k]) Im(W^k H[k])\right),$$

where $W^k = e^{-j\frac{\pi}{2N}\left(k+\frac{1}{2}\right)}$. The other terms in (8) and (9) are

written as follows:

$$\phi[k] = \frac{1}{N}\sum_{n=0}^{N-1} E[e_{wL}^2[n]]\cos\left(\frac{\pi}{N}n(2k+1)\right),$$

$$\phi_1[k] = \phi_2[k] = \frac{1}{N}\sum_{n=0}^{N-1} E[e_1^2[n]]\cos\left(\frac{\pi}{N}n(2k+1)\right),$$

$$\psi[k] = \frac{1}{N}\sum_{n=0}^{N-1} E[e_{wL}^2[n]]\sin\left(\frac{\pi}{N}n(2k+1)\right),$$

$$\psi_1[k] = \psi_2[k] = \frac{1}{N}\sum_{n=0}^{N-1} E[e_1^2[n]]\sin\left(\frac{\pi}{N}n(2k+1)\right).$$

It should be mentioned that an identical noise shaping filter is applied to rounding operations in the window and time domain aliasing operation and the first and second MDL steps [6]. (8) and (9) show that applying the noise shaping scheme shapes $\gamma[k]+\gamma_1[k]$ and $\gamma[k]+\gamma_2[k]$ in case of the left and right channel, respectively. However, this scheme introduces two extra terms associated with $\phi[k]$, $\phi_1[k]$, $\psi[k]$, and $\psi_1[k]$. Numerical evaluation confirms that each one of them is quite smaller than $\gamma[k]$ or $\gamma_1[k]$ in absolute value when k is neither nearly equal to 0 nor $N-1$ [6]. Needless to say, $|H[k]|^2 \geq \alpha[k]$ and $|H[k]|^2 \geq \beta[k]$. Therefore, the first two terms in both (8) and (9) are dominant when the value of k is neither nearly equal to 0 nor $N-1$.

It is possible to design different noise shaping filters to shape $\gamma[k]$, $\gamma_1[k]$, and $\gamma_2[k]$, individually. However, in this paper, we consider to shape these approximation noise spectra by using a fixed filter. This consideration actually makes the filter design complicated. This is because, as implied by the second term in (8), designing a filter with high attenuation at high frequencies could a large value of $E[h^2[n]]$, and as a result the second term $E[|h[n]*e_2[n]|^2]$ possibly increases the entire level of $E[E_L^2]$.

To avoid this problem, in this paper, the noise shaping scheme is not applied to the second MDL step. Consequently, (8) and (9) are re-written as

$$E[E_L^2[k]] \approx |H[k]|^2(\gamma[k]+\gamma_1[k]) + E[e_2^2[k]] + \alpha[k](\phi[k]+\phi_1[k]) + \beta[k](\psi[k]+\psi_1[k]), \quad (10)$$

$$E[E_R^2[k]] \approx |H[k]|^2\gamma[k] + E[e_2^2[k]] + E[e_3^2[k]] + \alpha[k]\phi[k] + \beta[k]\psi[k]. \quad (11)$$

IV. Reference MDCT Spectrum

In this section, we discuss the reference MDCT spectrum for the design of the noise shaping filter for the stereo IntMDCT-based lossless audio codec. The target audio signals compressed by the encoder are the audio items used in the MPEG-4 lossless audio coding task group [7]. Two different sampling frequencies, 48 kHz and 96 kHz, are used, and there are 15 items for each. All the items are quantized at 16 bit LPCM. In this paper, let us call them as 48/16 and 96/16, respectively.

Ideally, the noise shaping filter should work in such a way that the shaped noise spectrum is under the MDCT spectral envelope of each one of the target audio items. In order to do so, additionally, computation of the MDCT is necessary in advance. This is clearly not practical in terms of computational complexity. In this paper, we consider to obtain an average MDCT spectrum of the EBU-SQAM [8] (The total number of items is 70. The sampling frequency is 44.1 kHz and they are quantized at 16 bit LPCM) as a reference for the design. It should be noted that zero frames preceding and following each SQAM audio item are omitted and only the signal in the left channel of each item is used to compute the MDCT spectrum. Among 70 items, an average MDCT spectrum of audio items in each category is evaluated in terms of average bandwidth, and vocal (SQAM44-48) is chosen since the average MDCT spectrum is more bandlimited than any other one except artificial sounds (SQAM01-07) and speech signals (SQAM49-54), where none of the target items is categorized. The spectrum is shown in Fig. 3. Having the bandwidth as narrow as possible is desired in order to avoid designing a filter with wide passband which could unnecessarily amplify the approximation noise outside of the bandwidth of each target audio item.

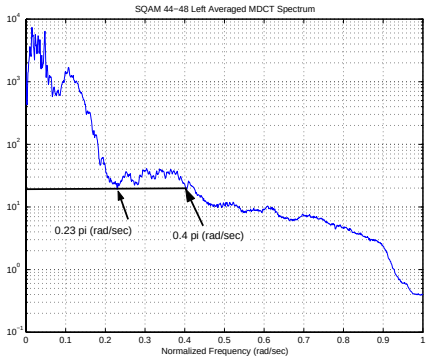


Fig. 3. An average MDCT spectrum of the left channel of vocal (SQAM44-48). A solid line indicates an experimentally determined minimum magnitude level within the passband of noise shaping filters.

Interestingly, the spectral level is less than around 0.5 when the frequency is more than 0.95π (rad/sec). This indicates the quantization noise due to LPCM occupies this band. Thus, reducing the approximation noise at this band is expected to be especially effective.

V. Constraints on the Filter Design

In this section, based on the reference spectrum obtained in the previous section, the constraints on the filter design are described.

A. Passband Gain

In order to determine the gain, it is at first necessary to know the approximation noise level. In (10) and (11), (4) and (5) are shaped by the noise shaping scheme. In fact, they can be numerically computed by using (7), and we have $\gamma[k]+\gamma_1[k] \approx 0.210$ and $\gamma[k] \approx 0.127$. Since they are actually noise power, by taking a square root of each, the noise level is approximately 0.458 and 0.356 for the left and right channel, respectively. Since the DC level of the reference

MDCT spectrum shown in Fig. 3 is about 400, the passband gain can be at most $400/0.458 \approx 872.8$. However, this value is practically too high to be used as the passband gain of the filter to shape the noise in each frame. We determine that the minimum absolute spectral level within the passband is 20 experimentally as it is shown by a solid line in Fig. 3.

In order to prevent degradation of the coding efficiency due to the amplified noise at the passband, the noise should not change the spectral envelope, in other word, the most significant bit (MSB) of the spectrum. Since the minimum absolute spectral value is 20, the maximum noise level at the passband has to be less than 15. Consequently, the maximum passband gain should be less than $15/0.458 \approx 32.75$. In this paper, three different passband gains, 5, 10, and 20, are used.

B. Stopband Width and Stopband Energy

Let $\overline{f_{Br}}$ be the bandwidth of the reference MDCT spectrum. The dimension is rad/sec. Then, the stopband width of the filter is determined by $\pi - \overline{f_{Br}}$ (rad/sec). As mentioned earlier, we determine the minimum spectral level at the passband is 20, and based on this decision, the possible range of the value of $\overline{f_{Br}}$ is set to be from $f_{low} = 0.23\pi$ (rad/sec) to $f_{high} = 0.4\pi$ (rad/sec). $\overline{f_{Br}}$ is experimentally determined within the range for each passband gain.

Once the stopband width is determined, the stopband energy has to be minimized to lower the approximation noise at the stopband as much as possible. This minimization is done by having a larger weight for the band occupied by the quantization noise than the weight for elsewhere in the stopband. The frequency range of the band occupied by the quantization noise, f_{Bq}, is $0.95\pi \sim \pi$ (rad/sec) for the $48/16$ items and $0.475\pi \sim \pi$ (rad/sec) for the $96/16$ items, respectively.

VI. Design of the Noise Shaping Filters

Table I summarizes the filters designed for $48/16$ and $96/16$ target items. For each sampling frequency, three filters are designed. H_1, H_2, and H_3 for $48/16$ items and H_4, H_5, and H_6 for $96/16$ items. In each sampling frequency, these filters are different on the passband gain such as 5, 10, and 20. The value of $\overline{f_{Br}}$ is experimentally determined to be 0.4π (rad/sec) for H_1, 0.23π (rad/sec) for the others in case of $48/16$ items, and 0.23π (rad/sec) for all the filters in case of $96/16$ items. The stopband energy within the band f_{Bq} is shown in parenthesis in the tables. It can be observed that the energy is almost the same for $48/16$ case but, in case of $96/16$ case, the higher the order of the filter is the less the energy. Fig. 4 (a) and (b) show their frequency responses in the DFT-III domain.

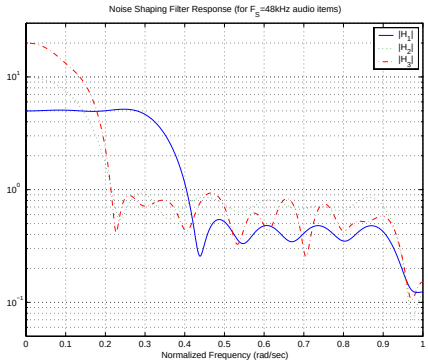
VII. Implementation Results

In this section, each of the designed filters is realized in the implementation of the IntMDCT based non-scalable codec [9]. This codec is only composed of the IntMDCT cascaded by an entropy coder. In this paper, a context-based arithmetic encoder [10] is used. Input audio items are $48/16$ and $96/16$ target audio items.

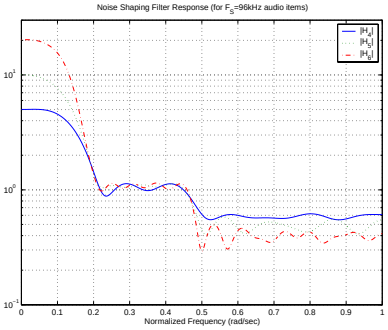
TABLE I

A summary of the filters for $48/16$ items (top) and for $96/16$ items (bottom). (() denotes the stopband energy within f_{Bq}.)

$48/16$	passband gain	stopband	order	stopband energy
H_1	5	$0.4\pi \sim \pi$	14	108.32(0.95)
H_2	10	$0.23\pi \sim \pi$	16	428.48(0.66)
H_3	20	$0.23\pi \sim \pi$	26	348.67(1.02)
$96/16$	passband gain	stopband	order	stopband energy
H_4	5	$0.23\pi \sim \pi$	18	496.77(190.9)
H_5	10	$0.23\pi \sim \pi$	20	456.44(134.08)
H_6	20	$0.23\pi \sim \pi$	32	411.9(89.87)



(a)



(b)

Fig. 4. Frequency response of the designed noise shaping filters for (a) $48/16$ audio items and (b) $96/16$ audio items.

First of all, the effect of the noise shaping is visually confirmed. Fig. 5 shows the noise power spectrum without noise shaping scheme and with the noise shaping filter H_6. The input audio signal is $96/16$ of "avemaria.wav". For the shaped noise power spectrum, the actual data obtained from the implementation and the theoretical curve computed by using (10) and (11) match well at all the bands except very low frequency bands.

Secondly, improvement in the lossless coding efficiency due to the noise shaping is evaluated. Table II shows the improvement in the estimated entropy and encoded file size compared to the no noise shaping case. All the values are shown in percent. The reason why the estimated entropy is also shown is to confirm the improvement due to not the

adaptiveness of the entropy coder used in this implementation but the noise shaping scheme. The equation to calculate the estimated entropy $\hat{H}$ is given by

$$\hat{H} = \sum_{fr=1}^{N_{fr}} \sum_{k=0}^{N-1} \log_2(2 * |X[k]| + 1), \qquad (12)$$

where N_{fr} is the number of frames and $X[k]$ is the IntMDCT spectrum at the spectral line index k.

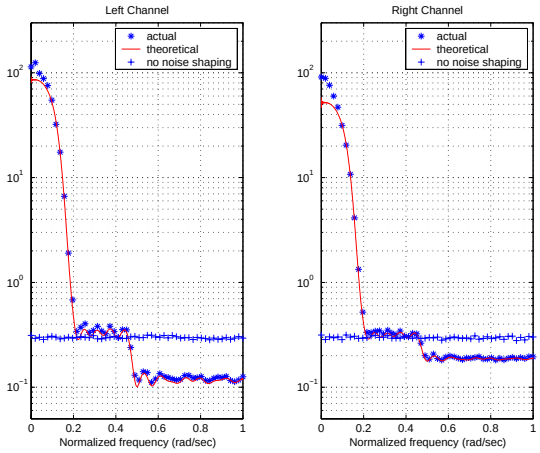


Fig. 5. The shaped noise power spectrum of "avemaria.wav" by using H_6. $*$: actual and $-$: theoretical values. $+$ indicates the noise power spectrum without noise shaping.

From Table II, several observations can be made. First of all, the improvement in the coding efficiency due to the noise shaping scheme is confirmed in terms of both the estimated entropy and actual encoded file size except a few particular audio items. Compared to the coding results in case of a simple filter $H(z) = 1+z^{-1}$ (0.05 percent for $48/16$ items and 0.5 percent for $96/16$ items on average) [6], a further improvement can be seen in Table II. Especially, the improvement is obtained for $96/16$ items. This is because the input signal spectrum is more bandlimited compared to the $48/16$ case due to the upsampling, and more importantly reduction of the quantization noise can be believed to contribute to this improvement mainly. Secondly, the value of $\overline{f_{Br}}$ and the passband gain have to be carefully chosen especially for audio items which have a lot of short attacks such as "cymbal.wav". Thirdly, the filter H_2 and H_5 improve the coding efficiency best for $48/16$ and $96/16$ items, respectively.

TABLE II

IMPROVEMENT IN ESTIMATED ENTROPY AND LOSSLESSLY ENCODED FILE SIZE (IN PERCENT).

48/16	H_1		H_2		H_3	
	est. entropy	file size	est. entropy	file size	est. entropy	file size
avemaria	0.208	0.216	0.211	0.203	0.251	0.237
blackandtan	0.083	0.085	0.068	0.069	0.076	0.080
broadway	0.038	0.035	0.031	0.029	0.035	0.032
cherokee	0.076	0.094	0.074	0.088	0.085	0.099
clarinet	0.114	0.116	0.108	0.108	0.116	0.113
cymbal	-1.964	-1.743	0.272	0.457	-0.413	-0.047
dcymbals	0.022	0.020	0.015	0.016	0.016	0.017
etude	0.128	0.133	0.124	0.115	0.137	0.129
flute	0.151	0.144	0.141	0.128	0.155	0.147
fouronsix	0.222	0.233	0.179	0.188	0.199	0.210
haffner	0.028	0.031	0.027	0.028	0.027	0.030
mfv	-0.593	-0.351	0.339	0.311	0.295	0.326
unfo	0.184	0.180	0.142	0.134	0.162	0.158
violin	0.065	0.064	0.060	0.056	0.063	0.059
waltz	0.121	0.127	0.098	0.107	0.110	0.116
average	-0.074	-0.041	0.126	0.136	0.088	0.114
96/16	H_4		H_5		H_6	
	est. entropy	file size	est. entropy	file size	est. entropy	file size
avemaria	0.958	0.892	1.098	1.033	1.129	1.095
blackandtan	1.256	1.323	1.484	1.559	1.631	1.709
broadway	0.295	0.285	0.335	0.333	0.368	0.358
cherokee	1.273	1.358	1.514	1.604	1.647	1.741
clarinet	1.617	1.599	1.902	1.872	2.065	2.035
cymbal	-0.217	-0.104	-1.007	-0.723	-2.746	-2.065
dcymbals	0.559	0.622	0.662	0.732	0.730	0.805
etude	0.741	0.694	0.854	0.807	0.896	0.865
flute	1.627	1.496	1.921	1.765	2.052	1.901
fouronsix	1.632	1.634	1.947	1.943	2.102	2.111
haffner	1.113	1.144	1.308	1.342	1.437	1.472
mfv	0.238	0.242	0.101	0.184	-0.473	-0.145
unfo	1.421	1.436	1.691	1.703	1.845	1.856
violin	1.321	1.293	1.566	1.521	1.702	1.655
waltz	1.348	1.405	1.598	1.660	1.741	1.804
average	1.012	1.021	1.132	1.156	1.075	1.147

VIII. CONCLUSION

In this paper, the approximation noise shaping was discussed to improve the lossless coding efficiency of the stereo IntMDCT-based lossless audio codec. The effect of the noise shaping was visually confirmed, and the coding results have shown the improvement especially for $96/16$ audio items.

REFERENCES

[1] I. Daubechies and W. Sweldens, "Factoring Wavelet Transforms into Lifting Steps," Tech. Rep., Bell Laboratories, Lucent Technologies, 1996.

[2] R. Geiger, J. Herre, J. Koller, and K. Brandenburg, "IntMDCT - A Link between Perceptual and Lossless Audio Coding," *Proc. ICASSP*, vol. 2, pp. 1813–1816, 2002.

[3] J. Liang and T. D. Tran, "Fast multiplierless approximations of the DCT with the lifting scheme," *IEEE Trans. on Signal Processing*, vol. 49, no. 12, pp. 3032–3044, December 2001.

[4] R. Geiger, Y. Yokotani, and G. Schuller, "Improved Integer Transforms for Lossless Audio Coding," in *Proc. of the Asilomar Conf. on Signals, Systems, and Computers*, 2003.

[5] M. Gerzon and P. G. Craven, "Optimal Noise Shaping and Dither of Digital Signals," *87th AES Convention*, Oct. 1989, Preprint 2822.

[6] Y. Yokotani, S. Oraintara, R. Geiger, G. Schuller, and K. R. Rao, "Approximation Noise Analysis for Transform-based Lossless Audio Coding," *Submitted to IEEE Globecom2004*.

[7] ISO/IEC JTC1/SC29/WG11 Moving Pictures Expert Group, "Final Call for Proposals on MPEG-4 Lossless Audio Coding," October 2002, Shanghai, China, N5208.

[8] European Broadcasting Union, "SQAM-Sound Quality Assessment Material Recordings for Subjective Tests," April 1988.

[9] R. Geiger, T. Sporer, J. Koller, and K. Brandenburg, "Audio Coding based on Integer Transforms," in *111th AES Convention*, New York, 2001, Preprint 5471.

[10] T. Qiu, "Lossless audio coding based on high order context modeling," *IEEE Fourth Workshop on Multimedia Signal Processing*, pp. 575–580, 2001.