



Audio Engineering Society Convention Paper 5471

Presented at the 111th Convention
2001 September 21–24 New York, NY, USA

This convention paper has been reproduced from the author's advance manuscript, without editing, corrections, or consideration by the Review Board. The AES takes no responsibility for the contents. Additional papers may be obtained by sending request and remittance to Audio Engineering Society, 60 East 42nd Street, New York, New York 10165-2520, USA; also see www.aes.org. All rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.

Audio Coding based on Integer Transforms

Ralf Geiger*, Thomas Sporer*, Jürgen Koller*, Karlheinz Brandenburg**/**

*Fraunhofer Institut für Integrierte Schaltungen, Arbeitsgruppe für Elektronische Medientechnologie

**Ilmenau Technical University

Am Helmholtzring 1, D-98693 Ilmenau, Germany

{ggr,spo,klr,bdg}@emt.iis.fhg.de

ABSTRACT

Most of the current audio coding schemes use transforms like the Modified Discrete Cosine Transform (MDCT) to calculate a blockwise frequency representation of the audio signal. Since these transforms usually produce floating point values even for integer input samples, a quantization process is necessary to achieve a reduction of data rate. This paper presents a new transform with perfect reconstruction that produces integer output values. The transform is called IntMDCT and is derived from the MDCT preserving most of its attractive properties. It provides a good spectral representation of the audio signal, critical sampling and overlapping of blocks. A lossless audio coding scheme may be built by simply cascading IntMDCT with an entropy coding scheme.

INTRODUCTION

Today audio coding is used for many applications both in the consumer and the professional market. The upcoming of lossless coding and the increased precision of 24 bit linear audio make rounding errors a serious issue for implementers. Most of the current audio coding schemes use transforms resp. filterbanks to get a blockwise frequency representation of the audio signal. These transforms usually produce floating point values even for integer input samples. So quantization is necessary to achieve a reduction of data rate. When applying

these transforms to lossless audio coding, either the quantization has to be fine enough to allow neglecting the resulting error, or the error signal has to be coded additionally in time domain [1], [2], [3].

An optimal transform for lossless audio coding should have the following properties:

Perfect reconstruction

By applying forward and inverse transform the input signal should be reconstructed without error.

Discrete spectral values

The transform should produce a discrete range of output values for discrete input values to enable a reduction of data rate without quantization.

Low range of spectral values

The range of spectral values should be as low as possible to achieve a high coding gain.

Good frequency selectivity

Tonal input signals should result in compaction of energy to a low number of coefficients.

Fast Algorithm

The transform should provide an algorithm that is at least as fast as algorithms for established transforms.

A promising approach for meeting these requirements is introduced in [4] by the lifting scheme. This technique allows to approximate Givens Rotations by mapping integers to integers in a reversible way. Therefore every transform that can be decomposed into Givens Rotations can be approximated by a lossless integer transform. For transforms focusing on image coding this technique was already used several times. In [5] an 8-point lossless Discrete Cosine Transform (DCT) is obtained by this idea. In [6] an 8-point lossless Lapped Orthogonal Transform (LOT) is described. In [8], [9], [10] this technique is further refined to get fast multiplierless approximations of DCT and LOT used for image coding. The lifting scheme can also be utilized for the Fast Fourier Transform (FFT), as shown in [11].

Recently the lifting scheme was initially utilized for perceptual audio coding [12]. An Integer Discrete Cosine Transform is used to remove inter-channel redundancy of a multichannel audio signal in a lossless way after quantization of MDCT coefficients of individual channels.

In this paper we will show that the MDCT itself can also be decomposed into Givens-Rotations and the lifting scheme can be applied.

This paper is organized as follows: After a short review of the Modified Discrete Cosine Transform a decomposition of this transform into Givens rotations is presented. Then the lifting scheme is introduced, which allows to approximate the decomposed transform by a reversible integer transform. The performance of this integer transform for audio coding is evaluated and some possible entropy coding schemes are presented. Finally additional coding tools are considered.

THE MODIFIED DISCRETE COSINE TRANSFORM

The Modified Discrete Cosine Transform (MDCT) is widely used in modern audio coding schemes. It provides critical sampling, overlapping of blocks and good frequency selectivity. To achieve critical sampling in combination with overlapping blocks a subsampling in frequency domain is performed. This subsampling introduces aliasing in time domain which is cancelled by an overlap and add of two succeeding blocks in the synthesis filterbank. This technique introduced in [13], [14] is called Time Domain Aliasing Cancellation (TDAC).

For a block t $2N$ time domain samples $x_t(k)$, $k = 0, \dots, 2N-1$ are used to calculate N spectral lines $X_t(m)$, $m = 0, \dots, N-1$. Two succeeding blocks overlap by 50%, so each block processes N new time domain samples. For a smooth overlapping of blocks a window $w(k)$, $k = 0, \dots, 2N-1$ is used.

The MDCT formula is given by

$$X_t(m) = \sqrt{\frac{2}{N}} \sum_{k=0}^{2N-1} w(k)x_t(k) \cos\left(\frac{\pi}{4N}(2k+1+N)(2m+1)\right)$$

$$m = 0, \dots, N-1$$

The formula for the inverse MDCT is

$$y_t(k) = w(k) \sqrt{\frac{2}{N}} \sum_{m=0}^{N-1} X_t(m) \cos\left(\frac{\pi}{4N}(2k+1+N)(2m+1)\right)$$

$$k = 0, \dots, 2N-1$$

By applying forward and inverse MDCT a time domain aliasing error is introduced. This error is cancelled by adding the outputs of the inverse MDCT of two succeeding blocks t and $t+1$ in the overlapping part:

$$x'_t(k) = y_t(N+k) + y_{t+1}(k)$$

$$k = 0, \dots, N-1$$

To ensure this time domain aliasing cancellation the windows of two succeeding blocks have to fulfill certain conditions in their overlapping part. A sufficient condition for time domain aliasing cancellation is:

$$w(k)^2 + w(N+k)^2 = 1$$

$$w(k) = w(2N-1-k)$$

$$k = 0, \dots, N-1 \quad (1)$$

An example for a window fulfilling this condition is a sine window

$$w(k) = \sin\left(\frac{\pi}{4N}(2k+1)\right)$$

$$k = 0, \dots, 2N-1$$

MDCT BY DCT-IV AND GIVENS ROTATIONS

An MDCT with a window length of $2N$ can be reduced to a Discrete Cosine Transform of Type IV (DCT-IV) with a length of N . This is achieved by performing Time Domain Aliasing (TDA) explicitly in time domain and consecutively applying the DCT-IV.

If we define the time domain aliased signal

$$\tilde{x}_t(k), \quad k = 0, \dots, N-1$$

by

$$\begin{aligned} \tilde{x}_t(k) &= w\left(\frac{N}{2} + k\right)x_t\left(\frac{N}{2} + k\right) \\ &\quad - w\left(\frac{N}{2} - 1 - k\right)x_t\left(\frac{N}{2} - 1 - k\right) \\ \tilde{x}_t(N-1-k) &= w\left(\frac{3N}{2} + k\right)x_t\left(\frac{3N}{2} + k\right) \\ &\quad + w\left(\frac{3N}{2} - 1 - k\right)x_t\left(\frac{3N}{2} - 1 - k\right) \end{aligned} \quad (2)$$

$$\begin{aligned} \tilde{x}_t(N-1-k) &= w\left(\frac{3N}{2} + k\right)x_t\left(\frac{3N}{2} + k\right) \\ &\quad + w\left(\frac{3N}{2} - 1 - k\right)x_t\left(\frac{3N}{2} - 1 - k\right) \end{aligned} \quad (3)$$

$$k = 0, \dots, \frac{N}{2} - 1$$

the formula for the MDCT reduces to

$$X_t(m) = \sqrt{\frac{2}{N}} \sum_{k=0}^{N-1} -\tilde{x}_t(N-1-k) \cos\left(\frac{\pi}{4N}(2k+1)(2m+1)\right)$$

$$m = 0, \dots, N-1$$

which is the application of a length N DCT-IV to

$$-\tilde{x}_t(N-1-k), \quad k = 0, \dots, N-1$$

The left half of the window for block t overlaps with the right half of block $t-1$. From equation (3) it follows that this part of the input signal is used for the MDCT of block $t-1$ by

$$\begin{aligned} \tilde{x}_{t-1}(N-1-k) &= w\left(\frac{3N}{2} + k\right)x_t\left(\frac{N}{2} + k\right) \\ &\quad + w\left(\frac{3N}{2} - 1 - k\right)x_t\left(\frac{N}{2} - 1 - k\right) \\ &= w\left(\frac{N}{2} - 1 - k\right)x_t\left(\frac{N}{2} + k\right) \\ &\quad + w\left(\frac{N}{2} + k\right)x_t\left(\frac{N}{2} - 1 - k\right) \end{aligned}$$

$$k = 0, \dots, \frac{N}{2} - 1$$

Combining this with equation (2) for block t we see that in the overlapping part of the two succeeding blocks $t-1$ and t the time domain signal $x_t(k), k = 0, \dots, N-1$ is prepared for application of DCT-IV by

$$\begin{pmatrix} \tilde{x}_t(k) \\ \tilde{x}_{t-1}(N-1-k) \end{pmatrix} = \begin{pmatrix} w\left(\frac{N}{2} + k\right) & -w\left(\frac{N}{2} - 1 - k\right) \\ w\left(\frac{N}{2} - 1 - k\right) & w\left(\frac{N}{2} + k\right) \end{pmatrix} \begin{pmatrix} x_t\left(\frac{N}{2} + k\right) \\ x_t\left(\frac{N}{2} - 1 - k\right) \end{pmatrix}$$

$$k = 0, \dots, \frac{N}{2} - 1$$

From the TDAC condition in equation (1) it follows that

$$w\left(\frac{N}{2} + k\right)^2 + w\left(\frac{N}{2} - 1 - k\right)^2 = 1$$

so for certain angles

$$\alpha_k = \arctan \frac{w\left(\frac{N}{2} - 1 - k\right)}{w\left(\frac{N}{2} + k\right)}$$

$$k = 0, \dots, \frac{N}{2} - 1$$

this preprocessing in time domain can be written as an application of Givens rotations

$$\begin{pmatrix} \cos \alpha_k & -\sin \alpha_k \\ \sin \alpha_k & \cos \alpha_k \end{pmatrix}$$

For the inverse MDCT the same procedure can be applied in reversed order. The inverse DCT-IV is the DCT-IV itself. The rotations applied for windowing and time domain aliasing are reverted by applying rotations with angles $-\alpha_k, k = 0, \dots, \frac{N}{2} - 1$. The whole process is illustrated in Figure 1.

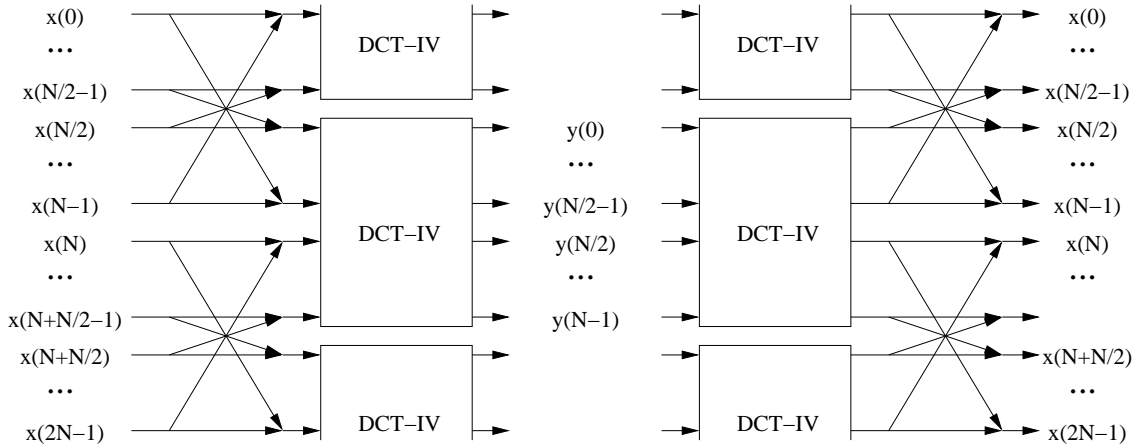


Fig. 1: Decomposition of MDCT and inverse MDCT into Givens rotations and DCT-IV

With this decomposition of MDCT it is easy to see that the window shape can be chosen individually in each frame as described in [16]. Based on rotations this window shape adaptation can be performed by changing the rotation angles for combined windowing and time domain aliasing in each frame. For perfect reconstruction it is only necessary to choose the negative angles of each frame in the inverse transform. So a window shape sequence like the one presented in [17] and illustrated in figure 2 is possible.

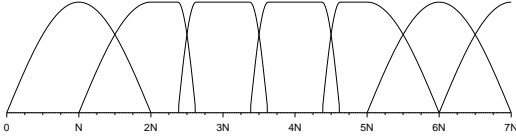


Fig. 2: Typical window shape sequence for MDCT

DCT-IV BY GIVENS ROTATIONS

The Discrete Cosine Transform of Type IV (DCT-IV) with length N is given by

$$X_t(m) = \sqrt{\frac{2}{N}} \sum_{k=0}^{N-1} x(k) \cos\left(\frac{\pi}{4N}(2k+1)(2m+1)\right) \\ m = 0, \dots, N-1$$

The coefficients of DCT-IV build an orthonormal $N \times N$ matrix. Every orthonormal $N \times N$ matrix can be decomposed into $\frac{N(N-1)}{2}$ Givens rotations [18]. But this decomposition is not unique. Other decompositions using a lower number rotations are possible. Some fast algorithms for DCT-IV focus on reducing the number of these rotations to a magnitude of $O(N \log_2 N)$. A possible decomposition is described in [19]. In [21] another decomposition of DCT-IV into Givens rotation is described implicitly by presenting a fast algorithm for the MDCT.

THE LIFTING SCHEME

The application of a Givens Rotation

$$\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$$

is illustrated in figure 3.

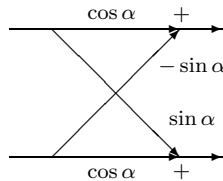


Fig. 3: Givens rotation

This Givens rotation can be decomposed into three lifting steps:

$$\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} = \begin{pmatrix} 1 & \frac{\cos \alpha - 1}{\sin \alpha} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \sin \alpha & 1 \end{pmatrix} \begin{pmatrix} 1 & \frac{\cos \alpha - 1}{\sin \alpha} \\ 0 & 1 \end{pmatrix}$$

This decomposition is illustrated in figure 4.

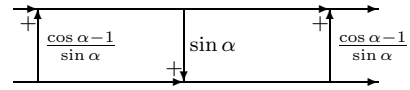


Fig. 4: Givens rotation by three lifting steps

We can now include a rounding function

$$r : \mathbb{R} \rightarrow \mathbb{Z}$$

into each of these lifting steps to get an integer approximation.

The application of the second lifting step

$$(x_1, x_2) \mapsto (x_1, x_2 + x_1 \sin \alpha)$$

for example is approximated by

$$(x_1, x_2) \mapsto (x_1, x_2 + r(x_1 \sin \alpha))$$

In this map the first component is not modified. So $r(x_1 \sin \alpha)$ can still be calculated after applying this map. So the inverse map can be built by

$$(x'_1, x'_2) \mapsto (x'_1, x'_2 - r(x'_1 \sin \alpha))$$

Therefore the integer approximation of the lifting step can be inverted without introducing any error. Applying this approximation to each of the three lifting steps we get an integer approximation of the Givens rotation. This rounded rotation can be reverted without introducing an error by applying the inverse rounded lifting steps in reverse order using the same rounding function.

If the rounding function r is odd symmetric the inverse rounded rotation is identical to the rounded rotation with angle $-\alpha$

$$\begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}$$

Figure 5 illustrates the inverse rotation by lifting steps.

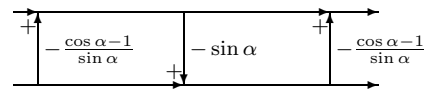


Fig. 5: Inverse Givens rotation by three lifting steps

THE INTEGER MODIFIED DISCRETE COSINE TRANSFORM (INTMDCT)

Replacing each Givens-Rotation of the MDCT decomposition described above by these rounded rotations, the output values stay integer, when integer input values are used. Nevertheless the whole process is invertible by applying the inverse rotations in reverse order. So we have an integer approximation

of the MDCT preserving perfect reconstruction. We call it the Integer Modified Discrete Cosine Transform (IntMDCT).

PERFORMANCE OF INTMDCT

This new transform produces integer output values instead of floating point values. It provides perfect reconstruction, so no error is introduced by applying forward and inverse transform. This transform is derived from the Modified Discrete Cosine Transform (MDCT). Therefore it preserves most properties of the MDCT: It has an overlapping structure providing better frequency selectivity than non-overlapping block transforms. Due to Time Domain Aliasing Cancellation (TDAC) critical sampling is maintained, so the total number of spectral values representing an audio signal does not exceed the number of input samples.

To study the frequency selectivity of IntMDCT it has to be considered that the result heavily depends on the level of the input signal. Due to the rounding in the rotation steps nonlinearities are included. So it is not possible to see this transform as an application of FIR filters and to compute the frequency responses. Therefore we try to get an impression of the frequency selectivity of IntMDCT by comparing the IntMDCT spectrum of certain input signals with the MDCT spectrum.

Figures 6 and 7 show the absolute values of MDCT and IntMDCT spectrum of a 1 kHz sine wave with a level of -20dB (SQAM01, [22]). For this signal the MDCT achieves a better rejection at high frequencies than the IntMDCT. Here the

IntMDCT reaches the limit of resolution for integer values and rounding errors of cascaded rounded rotations pile up. The absolute range for this rounding errors stays constant for most of the input signals. So the frequency selectivity of IntMDCT depends on the level of the input signal. For sine waves with a high level it is still comparable to the frequency selectivity of the MDCT.

For normal audio signals containing more than one frequency the rounding errors do not affect the spectrum as much as for sine waves. In figure 8 the absolute values of MDCT and IntMDCT spectrum of a part of Carl Orff's Carmina Burana (SQAM64, [22]) are compared in one plot together with the difference values. The difference values are not correlated with the spectral values, they have a constant order of magnitude in the whole spectral domain.

From a perceptual point of view the spectra in figure 8 are equal for most of the frequency bands. For audio signals containing a certain energy in each frequency band the difference between MDCT and IntMDCT is masked. So it may also be considered to use IntMDCT as an approximation of MDCT for perceptual audio coders.

Another interesting property of IntMDCT is a certain kind of energy preservation. Due to the overlapping structure an energy preservation on a block by block basis like the one described by Parseval's Theorem is not given. Energy can be distributed unequally between two succeeding blocks. But the averaged energy per block is maintained because in the

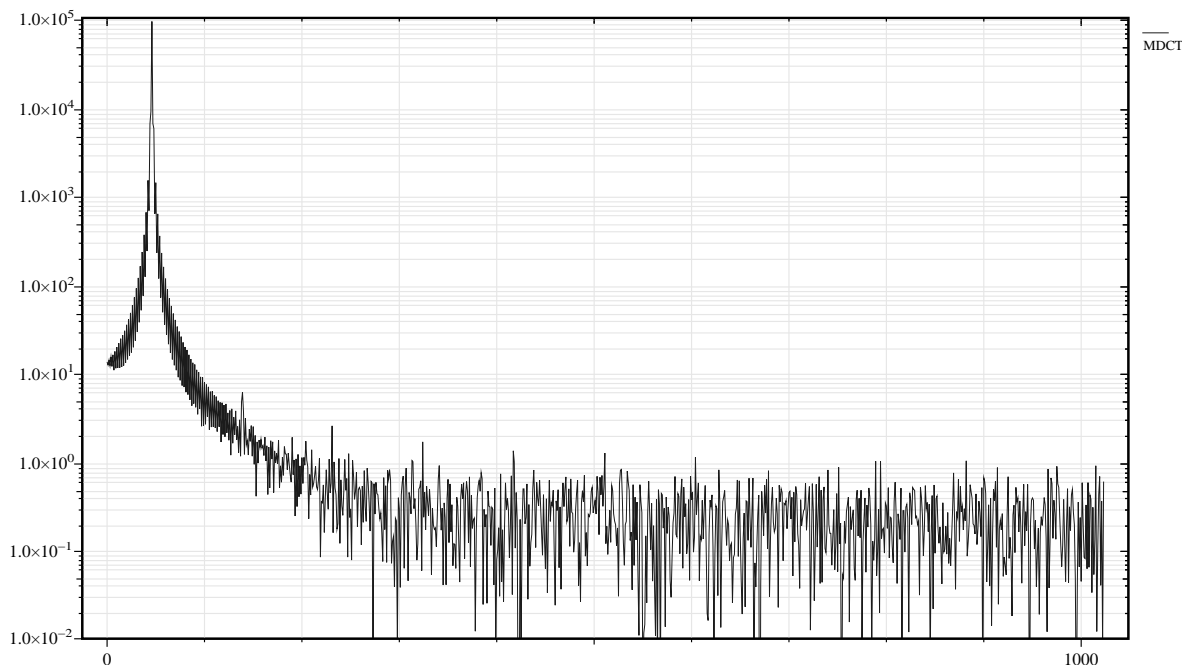


Fig. 6: Absolute values of MDCT spectrum, length 1024, sine window, SQAM01, Sine 1kHz -20dB

complete process only rounded Givens Rotations are applied which roughly preserve energy. So the range of integer spectral values does not exceed the range of input values by far. The additional dynamics in the range of spectral values compared with dynamics of the input signal only results from the energy compaction property of IntMDCT.

FAST ALGORITHM

The algorithm for IntMDCT is essentially based on fast algorithms for DCT-IV resp. MDCT using as low number of rotations. Givens rotations require four floating-point multiplications when applied directly for MDCT. Based on the lifting scheme only three floating-point multiplications are required for each rotation of IntMDCT. But on the other hand butterflies

$$\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

are calculated without multiplications for MDCT. For IntMDCT these butterflies have to be implemented as rounded Givens rotations with an angle of $\pi/4$ to ensure the energy preservation described above. This leads to three additional floating-point multiplications for each butterfly. So overall the computational complexity of MDCT and IntMDCT is roughly comparable when the lifting steps of IntMDCT are implemented by floating-point multiplications and roundings.

But the lifting scheme offers the possibility to further reduce computational complexity without losing the perfect

reconstruction property. This is achieved by approximating the floating-point lifting coefficients by dyadic numbers ($k \cdot 2^m, k, m \in \mathbb{Z}$) and performing the floating-point multiplications by shift and addition operations. This multiplierless approximation was introduced for image coding applications in [8], [10].

ENTROPY CODING

Concepts for entropy coding

IntMDCT provides a good spectral representation of the audio signal while staying in the integer domain. When applied to tonal parts of an audio signal this results in a good energy compaction. So an efficient lossless coding scheme can be built by simply cascading IntMDCT with an entropy coding scheme. This coding scheme should fit to the properties of the IntMDCT values. In contrast to entropy coding schemes for transform coding described in [23] and [1] the spectral values to be coded are not dynamically scaled to certain quantization step sizes. So a wide range of values has to be considered.

To adapt to different statistics and ranges of the integer spectrum the spectral domain is decomposed into bands adapted to the Bark scale. One possible decomposition is described in [23] using approximately two bands per Bark. For each band a different Huffman code book can be used. Possible lengths of codebooks can be from one up to e.g. 4096. Values greater than the maximum value can be coded by stacked coding, as described in [1].

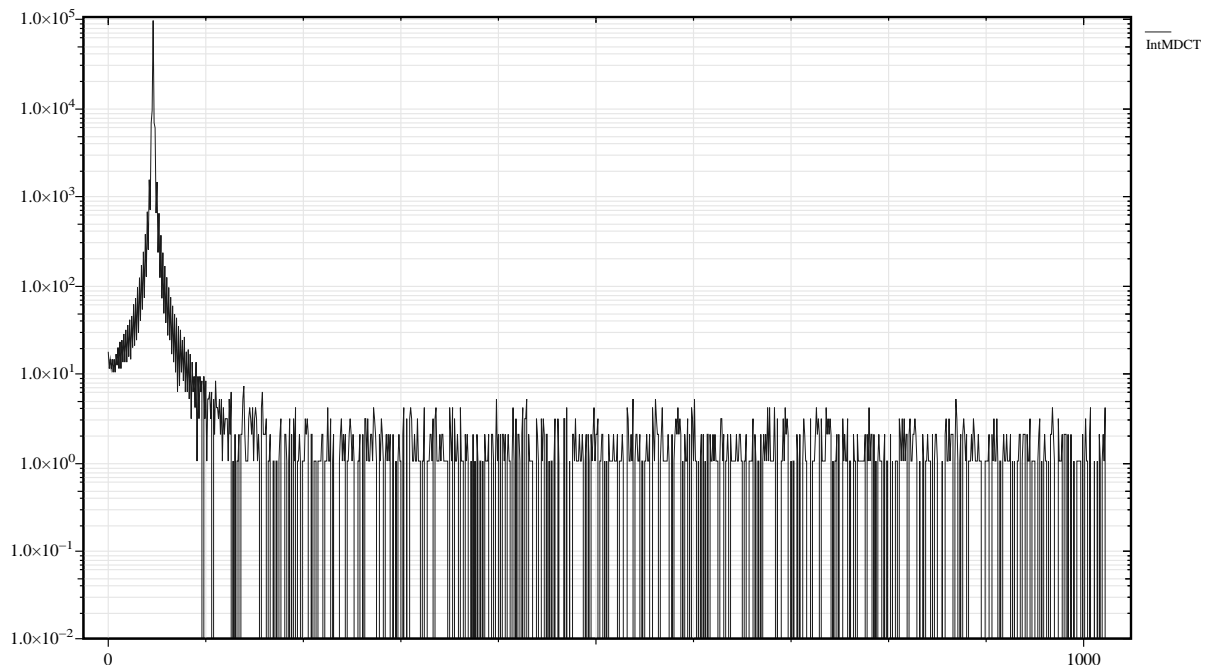


Fig. 7: Absolute values of IntMDCT spectrum, length 1024, sine window, SQAM01, Sine 1kHz -20dB

Due to the the absence of scaling another coding scheme may be considered: When most of the spectral lines of one band have to be coded using escape values, stacked coding can be very inefficient. It could be more convenient to scale down all values by a certain power of 2 until they fit to the desired codebook and additionally code the omitted LSBs. Compared with the alternative of using bigger codebooks this technique saves memory for storing codebooks. It is assumed to be appropriate because no additional coding gain will be achieved by codebooks exceeding the dynamic range of spectral values to be coded. As an interesting side effect a near lossless coder may be built by simply omitting some of the LSBs.

Results of entropy coding

First results for the compression efficiency are obtained using the following setup: For IntMDCT a frame length of 1024 samples and a sine window is used. The entropy coding scheme is implemented using eight huffman codebooks with lengths from one up to 16384 together with stacked coding. The codebook can be switched individually for each band.

The sound material used for testing comes from the SQAM compact disc [22]. These items have shown to be very critical for perceptual audio coding and have often been used as a reference for lossless audio coding.

Encoding all tracks an average data rate of 4.9 bit per sample is achieved. But for a realistic estimation of lossless coding efficiency for other audio signals it has to be considered that

the SQAM items contain lots of zero samples at the beginning and at the end of each track. Therefore frames which only contain zero samples are omitted in the following results. Encoding all tracks with zero frames omitted the average data rate increased to 5.6 bit per sample.

In figure 9 the average bit rates for individual SQAM items are presented. Especially for the artificial signals (tracks 3-7) and some of the single instruments items (tracks 8-43) a high coding gain is achieved. The worst case item for this compression scheme is Carl Orff's Carmina Burana (track 64) with an average bit rate of 9.1 bit per sample. This complex item contains choir and orchestra and has a very rich spectrum, see figure 8.

Besides the average data rate it is also important to know which maximum data rate usually occurs. In these test results the highest peak data rates measured were 14.9 bit per sample for track 31 (cymbal) and track 65 (orchestra, R. Strauss), and 13.9 bit per sample for track 27 (castanets). In all these items the peak data rates occur at transient parts of the signal.

ADDITIONAL CODING TOOLS

To enhance the performance of the lossless coding scheme described above two additional coding tools may be considered:

Linear Prediction in Frequency Domain

With the technique of entropy coding in spectral domain a

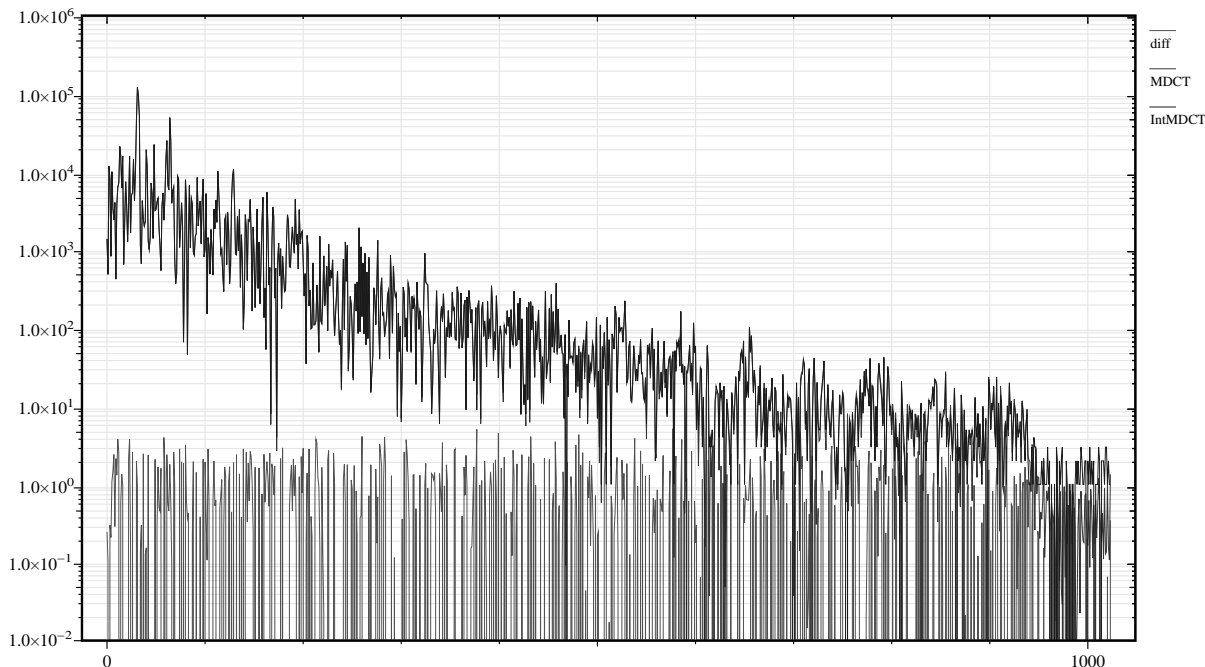


Fig. 8: Absolute values of MDCT, IntMDCT and difference spectrum, length 1024, sine window, SQAM64, Orff

high coding gain can be reached especially for tonal signals. For transient parts of the signal the coding gain is low due to the flat spectrum of transient signals. As described in [24], [25] this flatness can be exploited by applying linear prediction in frequency domain. Two alternatives are described there. One uses an open loop predictor, the other uses a closed loop predictor. The first alternative is also known as Temporal Noise Shaping (TNS). The quantization after the prediction lead to an adaption of the resulting quantization noise to the temporal structure of the audio signal and therefore prevents preechos in perceptual audio coders. This technique is used in MPEG-2 AAC [23]. For lossless audio coding the second alternative is more appropriate because the closed loop prediction allows perfect reconstruction of the input signal. When applying this technique to the IntMDCT spectrum a rounding to integer values has to be performed after each step of the prediction filter to stay in the integer domain. By using the inverse filter and the same rounding the original spectrum can be reconstructed perfectly.

Joint Stereo Coding

To use the redundancy between two channels mid-side-coding can be applied in a lossless way by applying a rounded rotation with angle $\pi/4$. Compared with the alternative of just calculating sum and difference of left and right channel the rounded rotation has the advantage of preserving energy. The usage of joint stereo coding can be switched on and off for each band, as done in [23]. Other rotation angles may also be considered to reduce redundancy between two channels in a more flexible way.

For multichannel signals the lossless redundancy reduction scheme based on Integer Discrete Cosine Transform described in [12] may be considered.

CONCLUSIONS

In this paper we have presented a new integer transform for audio coding. This transform is derived from the Modified

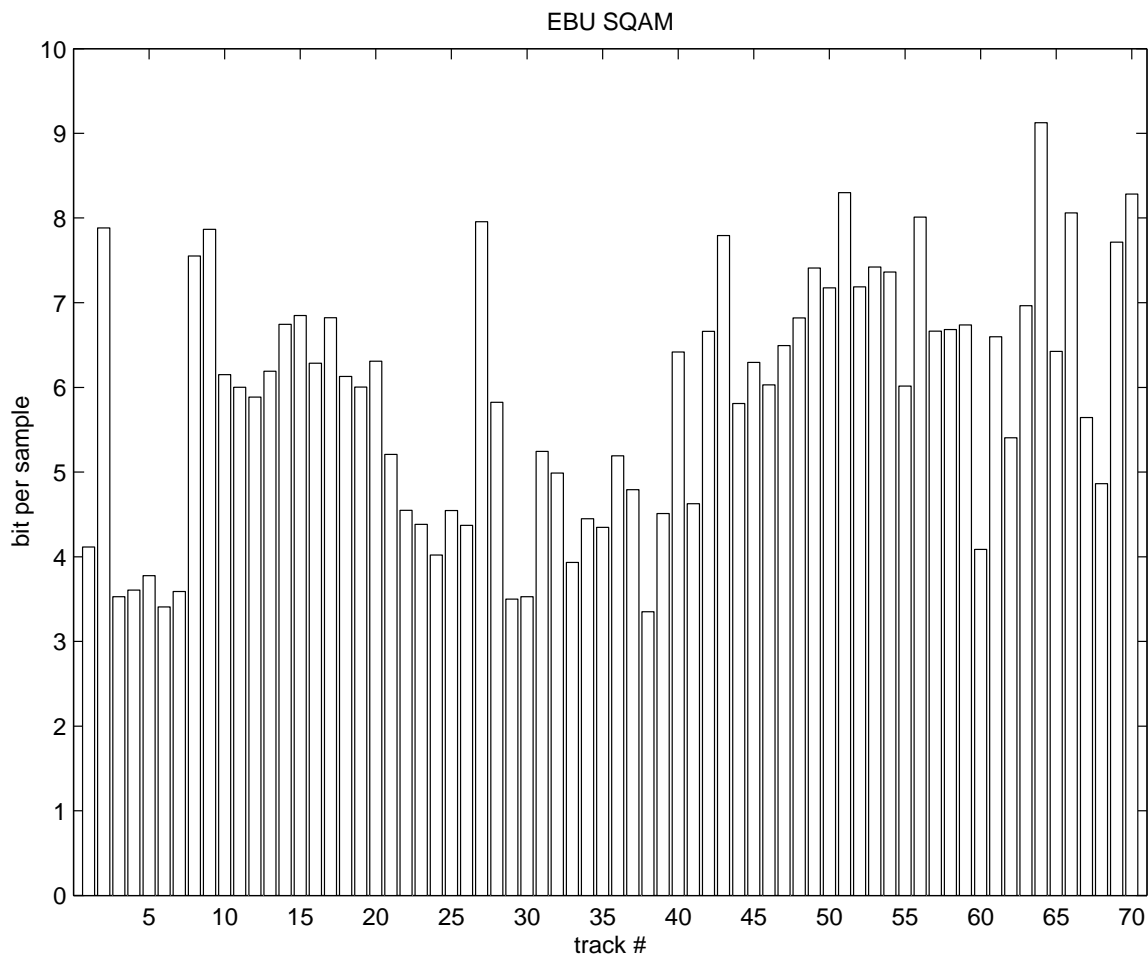


Fig. 9: Average bit rates for SQAM items, zero frames omitted

Discrete Cosine Transform using the lifting scheme. This IntMDCT preserves most of the attractive properties of MDCT: It provides perfect reconstruction, overlapping of blocks, critical sampling, good frequency selectivity and a fast algorithm. Additionally IntMDCT only produces integer output values for integer input samples. So a lossless audio coder can be built by cascading IntMDCT with an entropy coding scheme. This lossless audio coding scheme provides good compression efficiency.

ACKNOWLEDGEMENTS

The authors would like to thank Jürgen Herre for helpful remarks, Steffen Markert and Jens Hirschfeld for helping to perform entropy coding tests, and all the other colleagues at Fraunhofer Institute who supported this work.

REFERENCES

- [1] J. Koller, T. Sporer, K. Brandenburg: "Robust Coding of High Quality Audio Signals", 103rd AES-Convention, New York 1997, preprint 4621
- [2] J. Koller, T. Sporer, K. Brandenburg: "Improving Lossless Audio Coding", AES 17th International Conference, Florence 1999
- [3] M. Purat, T. Liebchen, P. Noll: "Lossless Transform Coding of Audio Signals", 102nd AES-Convention, Munich 1997, preprint 4414
- [4] I. Daubechies, W. Sweldens: "Factoring Wavelet Transforms into Lifting Steps", Preprint, Bell Laboratories, Lucent Technologies, 1996
- [5] K.Komatsu, K.Sezaki: "Reversible Discrete Cosine Transform", IEEE ICASSP98, vol.3, pp. 1769-1772, May 1998
- [6] K.Komatsu, K.Sezaki: "Design of Lossless LOT and Its Performance Evaluation", IEEE ICASSP2000, vol.4, pp.2119-2122, 2000
- [7] K.Komatsu, K.Sezaki: "Design of Lossless Block Transforms and Filter Banks for Image Coding", IEICE Transactions, Vol.E82-A No.8 p.1656-1664, August 1999
- [8] J. Liang, T. D. Tran: "Fast multiplierless approximations of the DCT with the lifting scheme", submitted to IEEE Trans. on Signal Processing, Feb. 2001
- [9] T. D. Tran: "The LiftLT: fast lapped transforms via lifting steps", IEEE Signal Processing Letters, vol. 7, pp. 145-149, Jun. 2000
- [10] T. D. Tran: "The BinDCT: fast multiplierless approximation of the DCT", IEEE Signal Processing Letters, vol. 7, pp. 141-145, Jun. 2000
- [11] S. Oraintara, Y. Chen, T. Nguyen: "Integer Fast Fourier Transform (INTFFT)", IEEE ICASSP2001, 2001
- [12] Y. Wang, M. Vilermo, M. Väänänen, L. Yaroslavsky: "A Multichannel Audio Coding Algorithm for Inter-Channel Redundancy Removal", AES 110th Convention, May 2001, Amsterdam, Netherlands, preprint 5295
- [13] J. Princen, A. Bradley: "Analysis/Synthesis Filter Bank Design Based on Time Domain Aliasing Cancellation", IEEE Transactions, ASSP-34, No.5, Oct 1986, pp. 1153-1161
- [14] J. Princen, A. Johnson, A. Bradley: "Sub-band/Transform Coding Using Filter Bank Designs Based on Time Domain Aliasing Cancellation", Proc. of the ICASSP 1987, pp 2161-2164
- [15] H. S. Malvar: "Signal Processing with Lapped Transforms", Artech House, 1992
- [16] B. Edler: "Codierung von Audiosignalen mit überlappender Transformation und adaptiven Fensterfunktionen", Frequenz, Vol. 43, pp. 252-256, 1989 (in German)
- [17] E. Allamanche, R. Geiger, J. Herre, T. Sporer: "MPEG-4 Low Delay Audio Coding based on the AAC Codec", 106th AES Convention, Munich 1999, preprint 4929
- [18] P. P. Vaidyanathan: "Multirate Systems and Filter Banks", Prentice Hall, Englewood Cliffs, 1993
- [19] Z. Wang: "Fast Algorithms for the Discrete W Transform and for the Discrete Fourier Transform", IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-32, No. 4, pp. 803-816, 1984
- [20] Z. Wang: "On Computing the Discrete Fourier and Cosine Transforms", IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-33, No. 4, pp. 1341-1344, 1985
- [21] T. Sporer, K. Brandenburg, B. Edler: "The use of multi-rate filter banks for coding of high quality digital audio", 6th European Signal Processing Conference (EUSIPCO), Amsterdam, June 1992, Vol.1 pp. 211-214
- [22] European Broadcasting Union (EBU): "Sound quality assessment material (SQAM) - Recordings for subjective tests"
- [23] ISO/IEC JTC1/SC29/WG11 MPEG, International Standard ISO/IEC 13818-7 "Generic Coding of Moving Pictures and Associated Audio: Advanced Audio Coding", 1997
- [24] J. Herre, J. D. Johnston: "Enhancing the Performance of Perceptual Audio Coders by Using Temporal Noise Shaping (TNS)", 101st AES Convention, Los Angeles 1996, preprint 4384
- [25] J. Herre, J. D. Johnston: "Exploiting Both Time and Frequency Structure in a System That Uses an Analysis/Synthesis Filterbank with High Frequency Resolution", 103rd AES Convention, New York 1997, preprint 4519