MULTIPLIERLESS FAST ALGORITHM FOR DCT VIA FAST HARTLEY TRANSFORM

Taizo Suzuki[1], *Yuichi Tanaka*[2], *Masaaki Ikehara*[3], *and Hirotomo Aso*[1]

[1]EEE Dept., College of Eng., Nihon University, Koriyama, Fukushima 963-8642, Japan
[2]IS Dept., Utsunomiya University, Utsunomiya, Tochigi 321-8585, Japan
[3]EEE Dept., Keio University, Yokohama, Kanagawa 223-8522, Japan
Email: {taizo, aso}@ee.ce.nihon-u.ac.jp, tanaka@is.utsunomiya-u.ac.jp, ikehara@tkhm.elec.keio.ac.jp

ABSTRACT

Discrete cosine transform (DCT) is known as efficient frequency transform, and when it is implemented on software/hardware, multiplier is undesirable for faster implementation. This paper presents a realization of multiplierless fast DCT for lossy image/video coding on arbitrary devices. First, the proposed DCT is constructed by using fast Hartley transform (FHT). Next, the redundancy of the structure is eliminated by using several characteristics of rotation matrix. Then, multiplierless DCT is obtained by approximating rotation matrices to multiplierless lifting structures with adders and bit-shifters. Finally, the proposed DCT is validated by comparing with the conventional DCTs in image coding.

Index Terms— Discrete cosine transform (DCT), fast Hartley transform (FHT), multiplierless lifting structure

1. INTRODUCTION

The discrete cosine transform (DCT) [1] has satisfactory performance in terms of energy compaction capability, and many fast DCT algorithms with efficient software/hardware implementations have been proposed. The DCT has found wide applications in image/video processing and other related signal processing fields. It has become the heart of many international standards such as JPEG, MPEG and H.26x family [2–4].

Since the conventional DCT usually maps integer values to floating-point ones, the computation cost and the power consumption cannot be neglected, especially the cost of floating-point multipliers. When implementing the DCT in mobile devices, the issue of power consumption is the most important. It is worth to come up with new algorithms for the DCT so that the dependence on floating-point multipliers can be reduced or completely eliminated [5–8].

In this paper, we proposes a fast DCT by using fast Hartley transform (FHT) [9] and multiplierless lifting structures for lossy image/video coding on any devices. The DCT can be implemented using only adders and bit-shifters, i.e., no multipliers. Also, it inherits all desirable DCT characteristics such as high coding gain, no DC leakage and symmetric basis functions, and has better coding performance than the conventional methods, which are integer DCT (IntDCT) based on Walsh-Hadamard transform (WHT) [5] and binary DCT (BinDCT) based on Chen's factorization [6].

Notations: **I**, **J** and $\{\cdot\}^T$ are an identity matrix, a reversal identity matrix and transpose of a matrix, respectively.

This work was supported in part by Utsunomiya University Grant for Exploratory Research.

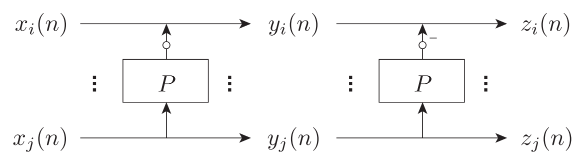

Fig. 1. A lifting structure (white circle: rounding operation).

2. REVIEW

2.1. Discrete Cosine Transform (DCT)

This paper only describes types DCT-II and -III which are commonly called DCT and inverse DCT (IDCT) [1], respectively. These are often used for the transform part in image/video coding such as JPEG, MPEG and H.26x family [2–4]. The (m, n)-elements of M-channel DCT and IDCT matrices $\mathcal{C}$ and $\mathcal{D}$ are defined as

$$[\mathcal{C}]_{m,n} = \sqrt{\frac{2}{M}}\, c_m \cos\left(\frac{m\,(n+1/2)\,\pi}{M}\right)$$
$$[\mathcal{D}]_{m,n} = \sqrt{\frac{2}{M}}\, c_n \cos\left(\frac{(m+1/2)\,n\pi}{M}\right)$$

where $\mathcal{D} = \mathcal{C}^{-1} = \mathcal{C}^T$, $0 \leq m, n \leq M-1$,

$$c_m = \begin{cases} \frac{1}{\sqrt{2}} & (m=0) \\ 1 & (m \neq 0) \end{cases} \quad \text{and} \quad c_n = \begin{cases} \frac{1}{\sqrt{2}} & (n=0) \\ 1 & (n \neq 0) \end{cases},$$

respectively. For simplicity, let us define $M = 2^n$ $(n \in \mathbb{N})$.

2.2. Multiplierless Lifting Structure

The lifting structure [10], also known as the ladder structure, is a special type of lattice structure. It is implemented by cascading elementary matrices - identity matrices with a single nonzero off-diagonal element.

Fig. 1 shows a basic lifting structure. It is expressed by

$$y_j(n) = x_j(n), \;\; y_i(n) = x_i(n) + \mathrm{round}\{Px_j(n)\}$$
$$z_j(n) = y_j(n), \;\; z_i(n) = y_i(n) - \mathrm{round}\{Py_j(n)\}$$

where $\mathrm{round}\{\cdot\}$ and P are rounding operation and a lifting coefficient, respectively. Thus the lifting structure with rounding operation can achieve integer-to-integer transform. Also, the lifting matrix and its inverse matrix in this case are represented as

$$\begin{bmatrix} 1 & P \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & P \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -P \\ 0 & 1 \end{bmatrix},$$

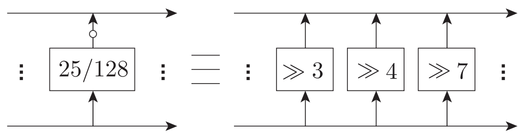

Fig. 2. An approximation from multiplier with dyadic value to bit-shifter and adder in a lifting structure ($\gg n$: n bit-shifter).

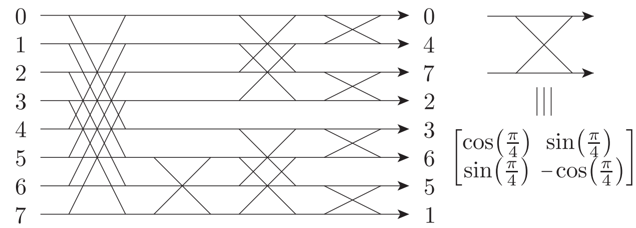

Fig. 3. 8-channel FHT $\hat{\mathcal{H}}$.

respectively.

For a high-speed implementation, lifting coefficients are required to approximate floating-point values to software/hardware-friendly dyadic values such as $k/2^n$ ($k,\ n \in \mathbb{N}$) which can be implemented by only adders and bit-shifters. The dyadic representation enables to perform fast implementation in a real time software encoder and reduce the circuit size. For example, a coefficient $25/128$ can be operated as

$$\frac{25}{128} = \frac{2^4+2^3+2^0}{2^7} = \frac{1}{2^3} + \frac{1}{2^4} + \frac{1}{2^7}.$$

Hence, the lifting structure with its coefficient $25/128$ and a rounding operation is replaced to the summation of 3, 4 and 7 bit-shifters illustrated in Fig. 2. It is clear that the perfect reconstruction in lifting structure is always kept even if floating-point values are approximated to dyadic values.

3. MULTIPLIERLESS FAST DCT VIA FHT

In this section, we derive a multiplierless fast DCT for lossy image/video coding by the following procedures.

1. A novel DCT structure is introduced based on FHT.
2. All of redundant rotation matrices in the structure are eliminated by using several characteristics of rotation matrix.
3. Multiplierless fast DCT is obtained by replacing several rotation matrices to multiplierless lifting structures and moving the scaling factors to the quantization part.

3.1. DCT via FHT

The FHT is a tool for the frequency analysis, design and implementation of digital signal processing algorithms and systems. It is strictly symmetrical concerning the transformation and its inverse. The (m, n)-elements of M-channel FHT matrix $\mathcal{H}$, which is the symmetric orthogonal matrix as $\mathcal{H}^{-1} = \mathcal{H}^T = \mathcal{H}$, are defined as [9]

$$[\mathcal{H}]_{m,n} = \frac{1}{\sqrt{M}} \operatorname{cas}\left(\frac{2mn\pi}{M}\right) \quad \text{where} \quad \operatorname{cas}\theta = \cos\theta + \sin\theta. \tag{1}$$

Since $\mathcal{H}$ in (1) is not according to the frequency band order, let $\mathcal{H}$ with the correct order be $\hat{\mathcal{H}}$. In case of $M = 8$, it is simplified by rotation matrices as shown in Fig. 3.

A new representation of DCT is constructed by only one FHT and $(M/2 - 1)$ rotation matrices as follows:

$$\mathcal{C} = \begin{bmatrix} 1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{JCJ} & \mathbf{0} & \mathbf{JS} \\ \mathbf{0} & \mathbf{0} & 1 & \mathbf{0} \\ \mathbf{0} & -\mathbf{SJ} & \mathbf{0} & \mathbf{C} \end{bmatrix} \hat{\mathcal{H}}^T \tag{2}$$

where

$$[\mathbf{C}]_{k,k} = \cos\left(\frac{(k+1)\pi}{2M}\right) \quad \text{and} \quad [\mathbf{S}]_{k,k} = \sin\left(\frac{(k+1)\pi}{2M}\right)$$
$$\text{for} \quad 0 \le k \le \frac{M}{2} - 2$$

as shown at the top of Fig. 4.

3.2. Redundant Rotation Matrices Elimination

Note that DCT via FHT in (2) has several redundant rotation matrices. In case of $M = 8$, 3 rotation matrices can be eliminated as shown in the middle of Fig. 4 by using the characteristics of rotation matrix as

$$\mathbf{\Theta}_4^{\pi}\mathbf{\Theta}_4^{\pi} = \mathbf{I} \quad \text{and} \quad \operatorname{diag}\{1, -1\}\mathbf{\Theta}_8^{\pi}\mathbf{\Theta}_4^{\pi} = \mathbf{\Theta}_8^{\pi}$$

where

$$\mathbf{\Theta}_b^a = \begin{bmatrix} \cos\left(\frac{a}{b}\right) & \sin\left(\frac{a}{b}\right) \\ \sin\left(\frac{a}{b}\right) & -\cos\left(\frac{a}{b}\right) \end{bmatrix}.$$

This non-redundant structure has 13 rotation matrices with $\pi/4$, $3\pi/16$, $\pi/8$ and $\pi/16$ angles.

3.3. Multiplierless Approximation

Multiplierless DCT is achieved by the following procedures. First, the rotation matrices marking angles in the middle of Fig. 4 are factorized into the scaled lifting structures [6] as

$$\begin{bmatrix} c(\theta) & s(\theta) \\ -s(\theta) & c(\theta) \end{bmatrix} = \begin{bmatrix} c(\theta) & 0 \\ 0 & \frac{1}{c(\theta)} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -s(\theta)c(\theta) & 1 \end{bmatrix} \begin{bmatrix} 1 & t(\theta) \\ 0 & 1 \end{bmatrix} \tag{3}$$

where $c(\theta) = \cos\theta$, $s(\theta) = \sin\theta$ and $t(\theta) = \tan\theta$. However, the only rotation matrix with $\pi/4$ angle drawn by bold lines in the middle of Fig. 4 is factorized into the normal lifting structure as

$$\begin{bmatrix} c(\theta) & s(\theta) \\ s(\theta) & -c(\theta) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{1-c(\theta)}{s(\theta)} & 1 \end{bmatrix} \begin{bmatrix} 1 & s(\theta) \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{c(\theta)-1}{s(\theta)} & 1 \end{bmatrix}.$$

Next, all of the floating-point values in the lifting coefficients are approximated to dyadic values such as $k/2^n$. Then, rotation matrices marking no angle in the middle of Fig. 4 are represented by

$$\begin{bmatrix} c\left(\frac{\pi}{4}\right) & s\left(\frac{\pi}{4}\right) \\ s\left(\frac{\pi}{4}\right) & -c\left(\frac{\pi}{4}\right) \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \tag{4}$$

Finally, the scaling factors $\operatorname{diag}\{c(\theta), 1/c(\theta)\}$ and $1/\sqrt{2}$ in (3) and (4) are moved to the quantization part, respectively. Consequently, the proposed DCT can be implemented by only adders and

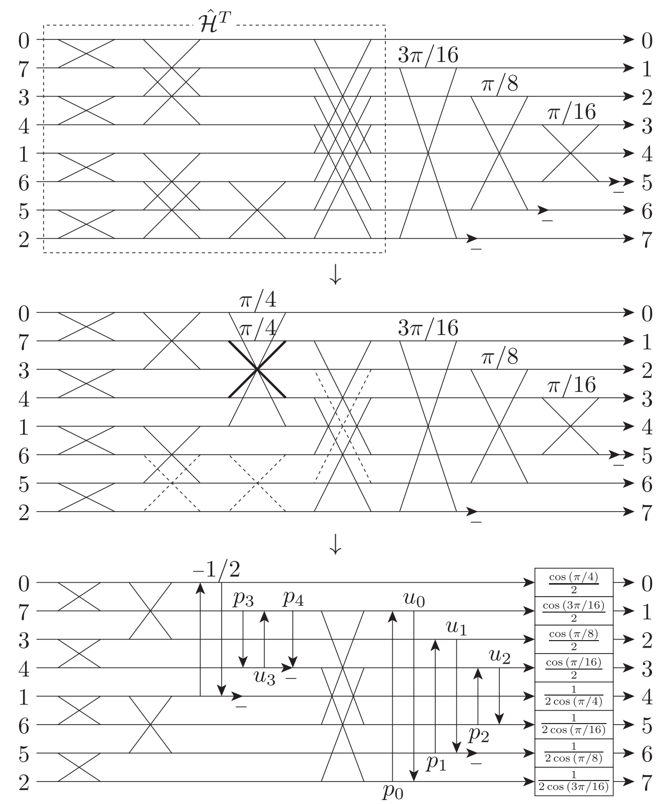

Fig. 4. The 8-channel proposed DCT: (top) DCT via FHT, (middle) redundant rotation matrices elimination, (bottom) its multiplierless approximation.

bit-shifters without multipliers as shown in the bottom of Fig. 4 and Table 1.[1] The dyadic values in Table 1 were empirically-determined.

4. EXPERIMENTAL RESULTS

4.1. Coding Gain and Frequency Response

Coding gain is one of the most important factors to be considered for a transform used in compression applications. A transform with higher coding gain compacts more energy into a fewer number of coefficients. As a result, higher objective performances such as peak signal-to-noise ratio (PSNR) would be achieved after quantization. Since the coding gain of the DCT approximates the optimal Karhunen-Loéve transform (KLT) closely, it is desired that the multiplierless DCT has similar coding gain to that of the original DCT. The biorthogonal coding gain is defined as [11]

$$C_g = 10 \log_{10} \frac{\sigma_x^2}{\left(\prod_{k=0}^{M-1} \sigma_{x_i}^2 \parallel f_i \parallel^2\right)^{\frac{1}{M}}}$$

where σ_x^2, $\sigma_{x_i}^2$ and $\parallel f_i \parallel^2$ are the variances of the input signal, the variance of the i-th subband signal and the norm of the i-th synthesis basis function, respectively. In this paper, we assume that input signal $x(n)$ is the AR(1) process with the intersample autocorrelation coefficient $\rho = 0.95$ in common use. Also, the comparison of the number of multipliers, adders, bit-shifters and coding gain C_gs is shown in Table 2 for 8-channel DCTs. Despite less adders, the coding gain of our DCT is higher than the conventional methods IntDCT [5] and BinDCT [6] which have type-C1 coefficients in [7]. The float DCT is based on Chen's factorization [12].[2]

Table 1. Lifting coefficients of the proposed DCT.

	floating-point value	dyadic value
p_0	$\tan\left(\frac{3\pi}{16}\right)$	$\frac{1}{2^1}+\frac{1}{2^3}$
u_0	$-\sin\left(\frac{3\pi}{16}\right)\cos\left(\frac{3\pi}{16}\right)$	$-\frac{1}{2^1}+\frac{1}{2^5}$
p_1	$\tan\left(\frac{\pi}{8}\right)$	$\frac{1}{2^2}+\frac{1}{2^3}+\frac{1}{2^5}$
u_1	$-\sin\left(\frac{\pi}{8}\right)\cos\left(\frac{\pi}{8}\right)$	$-\frac{1}{2^1}+\frac{1}{2^3}+\frac{1}{2^5}$
p_2	$\tan\left(\frac{\pi}{16}\right)$	$\frac{1}{2^3}+\frac{1}{2^4}$
u_2	$-\sin\left(\frac{\pi}{16}\right)\cos\left(\frac{\pi}{16}\right)$	$-\frac{1}{2^2}+\frac{1}{2^4}$
p_3	$\frac{\cos(\pi/4)-1}{\sin(\pi/4)}$	$-\frac{1}{2^1}+\frac{1}{2^4}+\frac{1}{2^5}$
u_3	$\sin\left(\frac{\pi}{4}\right)$	$1-\frac{1}{2^2}-\frac{1}{2^4}$
p_4	$\frac{1-\cos(\pi/4)}{\sin(\pi/4)}$	$\frac{1}{2^1}-\frac{1}{2^4}$

Table 2. The comparison of the number of multipliers, adders and bit-shifters and the coding gain C_g for 8-channel DCTs.

	DCT [12]	IntDCT [5, 7]	BinDCT [6, 7]	DCT Prop.
multiplier	52	0	0	0
adder	26	45	40	38
bit-shifter	0	17	23	21
C_g	8.8259	8.7240	8.8244	8.8250

Also, Fig. 5 compares the frequency responses of float DCT, IntDCT, BinDCT and the proposed DCT. Briefly speaking, the characteristics of the frequency responses of all of DCT are preserved without loss of the regularity, which is an important property for image compression [11].

4.2. Lossy Image Coding Performance

In this subsection, the proposed DCT is validated in lossy image coding by PSNR:

$$\text{PSNR [dB]} = 10 \log_{10} \left(\frac{255^2}{\text{MSE}}\right)$$

where MSE is the mean squared error. The test images are 512×512 8-bit grayscale images, *Barbara*, *Boat*, *Goldhill*, *Lena* and *Pepper*. The set partitioning in hierarchical trees (SPIHT) progressive image transmission algorithm [13] was used to encode the transformed images.

In Table 3, we compare the lossy image coding results in PSNR. Fig. 6 illustrates the comparison of a particular area of the image *Lena* when the compression ratio is 1:32. In especially low bit rate, it is obvious that the proposed DCT achieves comparable or even better performance on perceptual visual quality of reconstructed image against the conventional methods.[3]

[1] According to [6], we use the floating-point values of the scaling factors, which are always combined with the quantization steps and rounded to integers in practical implementations.

[2] Since [8] is DCT for lossy-to-lossless mode, the proposed DCT for only lossy mode is not compared with it in this paper.

[3] Since DCT in [12] is an ideal DCT structure, it basically shows the best coding performance.

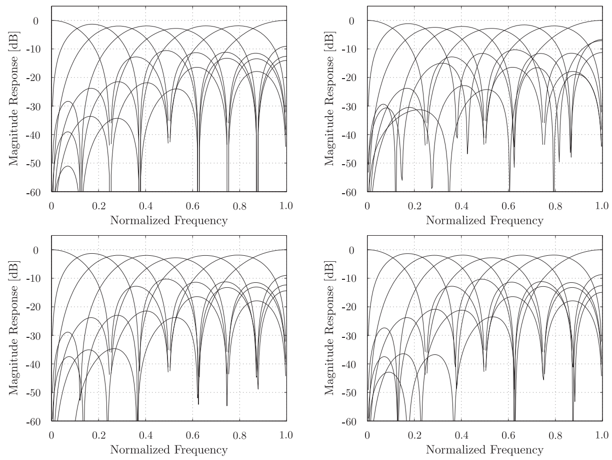

Fig. 5. The frequency responses of 8-channel DCTs: (top-left) float DCT [12], (top-right) IntDCT [5, 7], (bottom-left) BinDCT [6, 7], (bottom-right) the proposed DCT.

Fig. 6. Comparison of a particular area of the image *Lena* when the compression ratio is 1:32 : (left-to-right) float DCT [12], IntDCT [5, 7], BinDCT [6, 7] and the proposed DCT.

5. CONCLUSION

We presented an efficient multiplierless fast algorithm for discrete cosine transform (DCT). First, the DCT structure is introduced by focusing on fast Hartley transform (FHT). Second, all of redundant rotation matrices in the structure are eliminated by using the characteristics of rotation matrix. Then, the proposed DCT is obtained by replacing several rotation matrices to multiplierless lifting structures and moving the scaling factors to the quantization part. As a result, the proposed DCT achieves comparable or better image coding performance in low bit rates in spite of its less adders and bit-shifters than other multiplierless DCTs.

Table 3. Lossy image coding results by PSNR [dB] (the boldface values show the best coding performance except for DCT in [12]).

Test images	Comp. ratio	DCT [12]	IntDCT [5, 7]	BinDCT [6, 7]	DCT Prop.
Barbara	1:64	24.45	24.31	**24.44**	24.42
	1:32	26.95	26.68	**26.88**	**26.88**
	1:16	30.70	29.96	30.45	**30.47**
Boat	1:64	26.06	25.90	26.02	**26.03**
	1:32	28.66	28.43	28.58	**28.60**
	1:16	31.91	31.51	31.55	**31.64**
Goldhill	1:64	27.36	27.16	27.31	**27.39**
	1:32	29.38	29.17	29.25	**29.31**
	1:16	31.99	31.66	31.60	**31.69**
Lena	1:64	28.71	28.46	28.65	**28.69**
	1:32	31.91	31.48	31.69	**31.78**
	1:16	35.64	34.83	34.86	**35.04**
Pepper	1:64	28.22	28.02	28.18	**28.19**
	1:32	31.43	31.16	31.27	**31.31**
	1:16	34.49	34.00	33.92	**34.04**

6. REFERENCES

[1] K. R. Rao and P. Yip, *Discrete Cosine Transform Algorithms*, Academic Press, 1990.

[2] ISO/IEC 10918-1, *Information technology - Digital compression and coding of continuous-tone still images: Requirements and Guidelines*.

[3] ISO/IEC 14496-2, *Information technology - Coding of audio-visual objects - Part 2: Visual*.

[4] ISO/IEC 14496-10, *Information technology - Coding of audio-visual objects - Part 10: Advanced Video Coding*.

[5] Y. J. Chen, S. Oraintara, and T. Q. Nguyen, "Video compression using integer DCT," in *Proc. of ICIP'00*, Vancouver, British Columbia, Canada, Sep. 2000.

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

[7] Y. J. Chen, S. Oraintara, T. D. Tran, K. Amaratunga, and T. Q. Nguyen, "Multiplierless approximation of transforms with adder constraint," *IEEE Signal Process. Lett.*, vol. 9, no. 11, pp. 344–347, Nov. 2002.

[8] T. Suzuki and M. Ikehara, "Integer discrete cosine transform via lossless Walsh-Hadamard transform with structural regularity for low-bit-word-length," *IEICE Trans. Fundamentals*, vol. E93-A, no. 4, pp. 734–741, Apr. 2010.

[9] H. J. Meckelburg and D. Lipka, "Fast Hartley transform algorithm," *IET Electron. Lett.*, vol. 21, no. 8, pp. 311–313, Apr. 1985.

[10] I. Daubechies and W. Sweldens, "Factoring wavelet transforms into lifting steps," *J. Fourier Anal. Appl.*, vol. 4, no. 3, pp. 245–267, 1998.

[11] G. Strang and T. Nguyen, *Wavelets and Filter Banks*, Wellesley-Cambridge Press, 1996.

[12] W. H. Chen, C. H. Smith, and S. C. Fralick, "A fast computational algorithm for the discrete cosine transform," *IEEE Trans. Commun.*, vol. 25, no. 9, pp. 1004–1009, Sep. 1977.

[13] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, no. 3, pp. 243–250, June 1996.