Acknowledgments: The authors wish to acknowledge Corning Optical Fibres UK for supplying the fibre, Agilent Sweden for providing the test equipment, and D. Wake from University College London, for many helpful discussions and suggestions.

20 January 2003
Electronics Letters Online No: 20030369
DOI: 10.1049/el:20030369

R.E. Schuh and E. Sundberg (*TeliaSonera AB, SE-201 20 Malmö, Sweden*)

E-mail: ralf.x.schuh@telia.se

A. Alping (*Ericsson AB, SE-431 84 Mölndal, Sweden*)

References

1 SCHUH, R.E., WAKE, D., VERRI, B., and SUNDBERG, E.: 'Penalty free WCDMA radio signal transmission over fibre'. Proc. Microwave Photonics, MWP'00, Oxford, UK, September 2000, pp. 39-42
2 IEEE 802.3 Gigabit Ethernet Standard: '1000Base-SX for short wavelength transceivers (850 nm) and multimode fibre'
3 VCSEL Model 1A448 SC-2A from Zarlink Semiconductors
4 Photo receiver Model 8C478 SC-2A from Zarlink Semiconductors
5 3GPP TS 25.141 document, Technical Specification Group Radio Access Networks: 'Base station conformance testing (FDD)'

Efficient image downconversion for mixed field/frame-mode macroblocks

Sungjoo Suh, Seong Soo Chun, Myung-Hee Lee and Sanghoon Sull

An efficient image downconversion algorithm in the compressed domain for mixed field/frame-mode macroblocks is proposed. A 16 × 16 field/frame-mode macroblock is converted into an 8 × 8 reduced block in the discrete cosine transform (DCT) domain using a modified inverse DCT (IDCT) kernel. Experimental results show that the proposed algorithm provides the downconverted image quality similar to an existing method with significantly lower computational cost.

Introduction: Several image/video resizing algorithms have recently been proposed to resize the image/video directly in the DCT domain [1, 2]. Those DCT domain methods assumed that a DCT-encoded image contains only frame-mode macroblocks. However, a digitally broadcast video stream based on MPEG-2 contains mixed field/frame-mode macroblocks in a frame picture especially when there are large motion components between top and bottom fields. Thus there is a need for efficiently resizing the mixed field/frame-mode macroblocks in the DCT domain. Recently, Yim and Isnardi [3] proposed a DCT domain image resizing scheme for mixed field/frame-mode macroblocks using field/frame-mode resizing transformation matrices similar to the matrices in [1]. The method uses all the DCT coefficients despite the fact that high frequency DCT coefficients are usually unnecessary for image downconversion, and thus is not computationally efficient. In this Letter we propose an efficient image downconversion algorithm for mixed field/frame-mode macroblocks. The proposed method is based on a modified IDCT kernel providing stronger lowpass filtering while maintaining the efficiency of simple DCT domain truncation.

Modified IDCT: A common method for image downconversion averages neighbouring pixels. Let $x(n)$, $n=0, 1, \ldots, N-1$ be an N-point data sequence with N even and $x_R(n)$, $n=0, 1, \ldots, N/2-1$ be an $N/2$-point reduced data sequence. Then, $x_R(n)$ can be written as $x_R(n) = [x(2n)+x(2n+1)]/2$ where $n=0, \ldots, N/2-1$. Suppose that $X(k)$, $k=0, 1, \ldots, N-1$ is the DCT of $x(n)$. By expressing $x(2n)$ and $x(2n+1)$ in terms of $X(k)$ and applying a simple mathematical manipulation, we can derive

$$x_R(n) = \sum_{k=0}^{N-1} \alpha(k)X(k)\cos\left(\frac{k\pi}{2N}\right)C_{N/2}^{k(2n+1)} \qquad (1)$$

where $C_N^{k(2n+1)} = \cos((2n+1)k\pi/2N)$, $\alpha(0)=\sqrt{(1/N)}$ and $\alpha(k) = \sqrt{(2/N)}$ for $1 \le k \le N-1$.

Since the most significant energy is usually compacted in the low frequency range $k \in \{0, 1, \ldots, N/2-1\}$ and $\cos(k\pi/2N)$ is decreasing with the increasing value of k, we can ignore the high frequency coefficients in (1). Thus, $x_R(n)$ can be approximated as

$$x_R(n) \simeq \sum_{k=0}^{N/2-1} \alpha(k)\cos\left(\frac{k\pi}{2N}\right)X(k)C_{N/2}^{k(2n+1)} \qquad (2)$$

For example, when $N=8$, (2) describes the reduction of an 8-point data sequence to a 4-point data sequence and can be expressed as a matrix multiplication. Suppose that $\mathbf{x_R}$ is a column vector containing the 4-point reduced data sequence and $\hat{\mathbf{X}}$ is a column vector containing the first four low frequency components of an 8-point DCT data sequence. Then, we have the following:

$$\mathbf{x_R} = M_4^t\hat{\mathbf{X}} \qquad (3)$$

where $M_4 \overset{def}{=} DT_4$ and T_4 is a 4-point DCT matrix including normalising constants $\alpha(k)$ with $N=8$, and D is a 4×4 diagonal matrix defined as

$$D = \mathrm{diag}\left\{1, \cos\left(\frac{\pi}{16}\right), \cos\left(\frac{2\pi}{16}\right), \cos\left(\frac{3\pi}{16}\right)\right\} \qquad (4)$$

Note that M_4^t provides stronger lowpass filtering than the simple truncation IDCT matrix T_4^t used in [2] by zeroing out high frequency components and taking averages of neighbouring pixels. Thus, the modified IDCT method is an approximation of averaging filter for downsampling from the DCT domain to the spatial domain.

Downconversion with frame-mode macroblock: A 16 × 16 frame-mode macroblock $\mathbf{B}_F$ consists of four 8 × 8 DCT blocks. Let $\hat{\mathbf{B}}_F$ be an 8 × 8 block composed of the four submatrices each of which represents the 4 × 4 low frequency components of the corresponding 8 × 8 DCT block of $\mathbf{B}_F$. Then, $\hat{\mathbf{B}}_F$ can be written as

$$\hat{\mathbf{B}}_F = \begin{bmatrix} \hat{\mathbf{B}}_{11}\hat{\mathbf{B}}_{12} \\ \hat{\mathbf{B}}_{21}\hat{\mathbf{B}}_{22} \end{bmatrix} \qquad (5)$$

Suppose that $\hat{\mathbf{P}}_{11}$, $\hat{\mathbf{P}}_{12}$, $\hat{\mathbf{P}}_{21}$ and $\hat{\mathbf{P}}_{22}$ are reduced pixel blocks obtained from each $\hat{\mathbf{B}}_{11}$, $\hat{\mathbf{B}}_{12}$, $\hat{\mathbf{B}}_{21}$ and $\hat{\mathbf{B}}_{22}$, respectively. Then, the reduced pixel block $\hat{\mathbf{P}}_F$ in the spatial domain is described by extending (3) to two-dimension as follows:

$$\hat{\mathbf{P}}_F = \begin{bmatrix} \hat{\mathbf{P}}_{11}\hat{\mathbf{P}}_{12} \\ \hat{\mathbf{P}}_{21}\hat{\mathbf{P}}_{22} \end{bmatrix} = \begin{bmatrix} M_4^t\hat{\mathbf{B}}_{11}M_4 & M_4^t\hat{\mathbf{B}}_{12}M_4 \\ M_4^t\hat{\mathbf{B}}_{21}M_4 & M_4^t\hat{\mathbf{B}}_{22}M_4 \end{bmatrix} \qquad (6)$$

Let $\hat{\mathbf{B}}$ be the DCT of $\hat{\mathbf{P}}_F$ The reduced 8 × 8 DCT block downconverted from a 16 × 16 frame-mode macroblock can be obtained as

$$\begin{aligned} \hat{\mathbf{B}} = T_8\hat{\mathbf{P}}_{\mathbf{F}}T_8^t = [T_LT_R]&\begin{bmatrix} M_4^t\hat{\mathbf{B}}_{11}M_4 & M_4^t\hat{\mathbf{B}}_{12}M_4 \\ M_4^t\hat{\mathbf{B}}_{21}M_4 & M_4^t\hat{\mathbf{B}}_{22}M_4 \end{bmatrix}\begin{bmatrix} T_L^t \\ T_R^t \end{bmatrix} \\ &= (T_LM_4^t)\hat{\mathbf{B}}_{11}(T_LM_4^t)^t + (T_LM_4^t)\hat{\mathbf{B}}_{12}(T_RM_4^t)^t \\ &\quad + (T_RM_4^t)\hat{\mathbf{B}}_{21}(T_LM_4^t)^t + (T_RM_4^t)\hat{\mathbf{B}}_{22}(T_RM_4^t)^t \end{aligned} \qquad (7)$$

where T_8 is an 8-point DCT matrix including normalising constants $\alpha(k)$ with $N=8$, and T_L and T_R are two submatrices representing the first and last four columns of T_8, respectively. By decomposing $T_LM_4^t$ and $T_RM_4^t$ into $C+D$ and $C-D$ with two very sparse matrices C and D, we have

$$\hat{\mathbf{B}} = (X+Y)C^t + (X-Y)D^t \qquad (8)$$

where

$$X = C(\hat{\mathbf{B}}_{11} + \hat{\mathbf{B}}_{21}) + D(\hat{\mathbf{B}}_{11} - \hat{\mathbf{B}}_{21})$$
$$Y = C(\hat{\mathbf{B}}_{12} + \hat{\mathbf{B}}_{22}) + D(\hat{\mathbf{B}}_{12} - \hat{\mathbf{B}}_{22})$$

Note that the matrices C and D, the elements of which are not shown here, are different from those in [2] since we use the modified IDCT matrix M_4^t employing the diagonal matrix D instead of the simple IDCT truncation matrix T_4^t.

Downconversion with field-mode macroblock: In the spatial domain, a field-mode macroblock consists of two upper 8 × 8 blocks representing even lines and two bottom 8 × 8 blocks representing odd lines before de-interlacing. Let $\mathbf{P}_f$ be a field-mode macroblock in the spatial domain before de-interlacing with four 8 × 8 sub-matrices, denoted by

$$\mathbf{P}_f = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{bmatrix} \tag{9}$$

Since the averages of neighbouring even and odd lines provide vertical downsampling, the vertically downconverted pixel block can be obtained by averaging the upper and bottom blocks. Suppose that $\tilde{\mathbf{P}}_{11}$, $\tilde{\mathbf{P}}_{12}$, $\tilde{\mathbf{P}}_{21}$ and $\tilde{\mathbf{P}}_{22}$ are horizontally reduced pixel blocks from $\mathbf{P}_{11}$, $\mathbf{P}_{12}$, $\mathbf{P}_{21}$ and $\mathbf{P}_{22}$, respectively. Let $\hat{\mathbf{P}}_f$ be a downconverted pixel block in both vertical and horizontal directions. Then, we have

$$\hat{\mathbf{P}}_f = \frac{1}{2}[\,\tilde{\mathbf{P}}_{11} + \tilde{\mathbf{P}}_{21} \quad \tilde{\mathbf{P}}_{12} + \tilde{\mathbf{P}}_{22}\,] \tag{10}$$

Let $\hat{\mathbf{B}}_f$ be the DCT of $\hat{\mathbf{P}}_f$ and $\tilde{\mathbf{B}}_{ij}$ be the left 8 × 4 DCT subblock taken from 8 × 8 DCT block of each $\mathbf{P}_{ij}$ for $i, j = 1, 2$. Here, we need only four low frequency DCT coefficients per each row for horizontal downconversion. Thus, the reduced 8 × 8 DCT block downconverted from a 16 × 16 field-mode macroblock can be written as

$$\begin{aligned} \hat{\mathbf{B}}_f &= T_8\hat{\mathbf{P}}_f T_8^t \\ &= \frac{1}{2}T_8[T_8^t(\tilde{\mathbf{B}}_{11} + \tilde{\mathbf{B}}_{21})M_4 T_8^t(\tilde{\mathbf{B}}_{12} + \tilde{\mathbf{B}}_{22})M_4]T_8^t \\ &= \frac{1}{2}((\tilde{\mathbf{B}}_{11} + \tilde{\mathbf{B}}_{21})M_4 T_L^t + (\tilde{\mathbf{B}}_{12} + \tilde{\mathbf{B}}_{22})M_4 T_R^t) \\ &= \frac{1}{2}((\tilde{\mathbf{B}}_{11} + \tilde{\mathbf{B}}_{12} + \tilde{\mathbf{B}}_{21} + \tilde{\mathbf{B}}_{22})C' \\ &\quad + (\tilde{\mathbf{B}}_{11} + \tilde{\mathbf{B}}_{21} - \tilde{\mathbf{B}}_{12} - \tilde{\mathbf{B}}_{22})D') \end{aligned} \tag{11}$$

Note that an image with field-mode macroblocks can be downconverted using very sparse matrices C and D, resulting in lower computational overhead. The simple DCT truncation in the DCT domain could cause visual artifacts such as aliasing and block artifacts. Further, since a field-mode macroblock is vertically downsized by averaging neighbouring pixels, it is desirable to horizontally downsize in the same way to avoid visual artifacts caused by the difference in filtering characteristics between vertical and horizontal downconversions. Our proposed method yields a good compromise between computational overhead and visual quality by approximating the process of downsampling or averaging in the DCT domain.

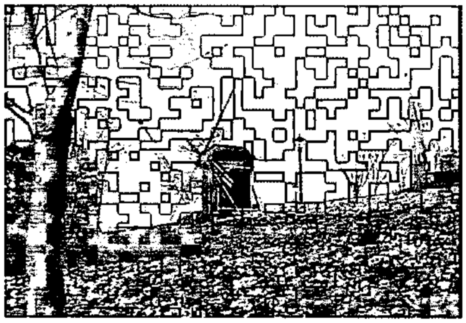

Fig. 1 *Test image with randomly generated field/frame-mode macroblocks*

Experimental results: In the experiment, we implemented Yim's method described in [3] using the fast 1-D IDCT algorithm [4]. Compared with Yim's method, our algorithm offered the PSNR value within 0.1 dB over a wide range of images. Fig. 1 shows the flower garden image with mixed field/frame-mode macroblocks randomly generated where the intensity values of the field-mode macroblocks are decreased for display purpose. The downsized images in the DCT domain using Yim's method and the proposed method are shown in Figs. 2 and 3. The quality of the two resized images appears similar without any noticeable visual artifact. Table 1 indicates that the computational complexity of the proposed scheme is much lower than Yim's method for downconversion of field/frame-mode macroblocks.

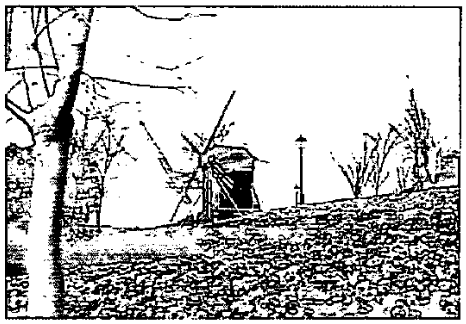

Fig. 2 *Downconverted image using Yim's method*

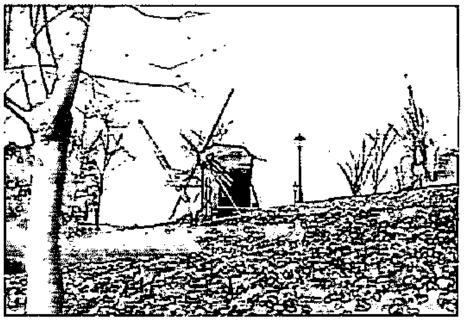

Fig. 3 *Downconverted image using proposed method*

Table 1: Complexity comparison for downconversion of field/frame-mode macroblock

		Multiplication	Addition
Field-mode macroblock	Proposed method	160	208
	Yim's method	264	888
Frame-mode macroblock	Proposed method	320	320
	Yim's method	880	2512

Conclusion: We have proposed an efficient algorithm for image downconversion for mixed field/frame-mode macroblocks in the DCT domain. The proposed method is based on the modified IDCT kernel that efficiently approximates average operations for downsampling. The experimental results show that the proposed method achieves significant computational savings compared with the existing method while providing the similar visual quality.

17 January 2003
Electronics Letters Online No: 20030362
DOI: 10.1049/el:20030362

Sungjoo Suh, Seong Soo Chun, Myung-Hee Lee and Sanghoon Sull (*Department of Electronics and Computer Engineering, Korea University, 5-1 Anam-dong, Songbuk-gu, Seoul, 136-701, Korea*)

References

1 CHANG, S.F., and MESSERSCHMITT, D.G.: 'Manipulation and compositing of MC-DCT compressed video', *IEEE J. Sel. Areas Commun.*, 1995, **13**, (1), pp. 1–11

2 DUGAD, R., and AHUJA, N.: 'A fast scheme for image size change in the compressed domain', *IEEE Trans. Circuits Syst. Video Technol.*, 2001, **11**, (4), pp. 461–474

3 YIM, C., and ISNARDI, M.A.: 'An efficient method for DCT-domain image resizing with mixed field/frame-mode macroblocks', *IEEE Trans. Circuits Syst. Video Technol.*, 1999, **9**, (5), pp. 696–700

4 LOEFFER, C., and LIGTENBERG, A.: 'Pratical fast 1-D DCT algorithms with 11 multiplications'. IEEE Conf. Acoustics, Speech and Signal Processing, Glasgow, UK, May 1989, Vol. 2, pp. 988–991