RESIZING OF IMAGES IN THE DCT SPACE BY ARBITRARY FACTORS

Jayanta Mukhopadhyay

Dept. of Computer Science & Engineering,
Indian Institute of Technology,
Kharagpur, INDIA 721302.
Email: jay@cse.iitkgp.ernet.in

Sanjit K. Mitra *

Dept. of Electrical & Computer Engineering,
University of California,
Santa Barbara, CA, 93106, USA.
Email: mitra@ece.ucsb.edu

ABSTRACT

In this work we describe algorithms for the resizing by arbitrary factors of images represented in the block DCT space. The algorithms are designed by exploiting spatial relationship of the block DCT coefficients and sub-band correlation of the DCT coefficients. An interesting application of these methods shown is the conversion in the compressed domain of images (video frames) from one format to another.

1. INTRODUCTION

Image resizing operation is often required for display, storage, and transmission of images. In Internet applications, browsing of images may require transmission of the same image at varying resolutions for different specifications of the display and communication network at the client ends. Another important application of image resizing operation is in the transcoding of images/videos from one data format to the other (e.g. HDTV to NTSC). Usually, resizing operation is performed in the spatial domain. However, as most images are stored in the compressed format, it is preferable to perform the resizing operation directly in the compressed domain. This reduces the computational overhead associated with decompression and compression operations with the compressed stream.

As the DCT-based JPEG standard is widely used for image compression, a number of approaches have been advanced to resize the images in the DCT space [1], [2], [3], [4], [5], [6]. In particular, Dugad and Ahuja [7] have suggested a simple fast computation technique for halving and doubling of images using their low frequency components. Later, Mukherjee and Mitra [8] proposed some modifications to their scheme.

Usually, image resizing operations are meant for either image halving or image doubling operations. However there are a few efforts in recent times for performing these operations in general cases. In [9] image resizing operations with arbitrary factors of $\frac{L}{M} \times \frac{L}{M}$ (where L and M are positive integers) are carried out in the DCT domain by making use of the multiplication-convolution properties of the discrete trigonometric transforms [10]. Our approach distinctly differs from this work and it is general enough to accommodate arbitrary resizing with factors $\frac{P}{Q} \times \frac{R}{S}$, where P, Q, R and S are positive integers. In our work we have used the spatial relationships of the DCT coefficients between a block and its sub-blocks as developed by Jiang and Feng [11]. Using their approach it is possible to decompose (recompose) the DCT block(s). Fortunately, their techniques for recomposition and decomposition of DCT blocks (coupled with sub-band approximation of DCT coefficients) can also be used in image resizing. We have shown how these could be carried out by exploiting the spatial relationship and sub-band correlation of the DCT coefficients. One of the major contributions in the proposed schemes is the identification of computational steps providing a general framework for image resizing operations. The approach is general enough to perform resizing operations with arbitrary factors, namely with integral and rational factors. A practical application of these methods shown here is the conversion of images (video frames) from one format to another in the compressed domain itself. It is demonstrated by converting a HDTV frame to a NTSC one.

2. SUB-BAND DISCRETE COSINE TRANSFORM COMPUTATION

We review here briefly the approximate computation of DCT coefficients from its sub-bands. In addition, we summarize the observations with regard its relevance to image resizing made by us previously [8]. The DCT of a 2-D image

*Work supported in part by a University of California MICRO grant with matching support from Philips Research Laboratories, and in part by Microsoft Corp.

0-7803-8554-3/04/$20.00 ©2004 IEEE.

$\{x(m,n), 0 \leq m \leq N-1, 0 \leq n \leq N-1\}$ is given by:

$$C(k,l) = \frac{2}{N}.\alpha(k).\alpha(l).\sum_{m=0}^{N-1}\sum_{n=0}^{N-1}(x(m,n) \cos(\frac{(2m+1)\pi k}{2N})\cos(\frac{(2n+1)\pi l}{2N})), \quad 0 \leq k,l \leq N-1. \tag{1}$$

where $\alpha(p)$ is given by:

$$\alpha(p) = \begin{cases} \sqrt{\frac{1}{2}} & \text{for } p = 0 \\ 1 & \text{otherwise,} \end{cases} \tag{2}$$

where $p = k, l$.

The *low-low* sub-band $x_{LL}(m,n)$ of the image is obtained as:

$$x_{LL}(m,n) = \frac{1}{4}\{x(2m,2n) + x(2m+1,2n) + x(2m,2n+1) + x(2m+1,2n+1)\}, \quad 0 \leq m,n \leq \frac{N}{2}-1. \tag{3}$$

Let $\overline{C_{LL}}(k,l), 0 \leq k,l \leq \frac{N}{2}-1$ be the 2D DCT of $x_{LL}(m,n)$. Then the *sub-band approximation* of DCT of $x(m,n)$ is given by:

$$C(k,l) = \begin{cases} 2\cos(\frac{\pi k}{2N})\cos(\frac{\pi l}{2N})\overline{C_{LL}}(k,l), & 0 \leq k,l \leq \frac{N}{2}-1, \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

It should be noted that for small values of k and l, cosine terms can be approximated as 1. Hence, Eq. (4) can be further simplified as:

$$C(k,l) = \begin{cases} 2\overline{C_{LL}}(k,l), & 0 \leq k,l \leq \frac{N}{2}-1, \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

We refer the above approximation as *low-pass truncated approximation* of the sub-band coefficients.

3. RECOMPOSITION AND DECOMPOSITION OF DCT BLOCKS

In this section we briefly describe the technique for recomposition and decomposition of DCT blocks. The details could be found in [11]. For convenience, we first discuss the spatial relationships of the DCT blocks in 1-dimension. Let there be M blocks of the N-point DCT of a sequence $\{x(n)\}, 0 \leq n \leq MN-1$. Hence the DCT of the pth block can be expressed as follows:

$$C_p(k) = \sqrt{\frac{2}{N}}\alpha(k)\sum_{n=0}^{N-1} x(pN+n)cos(\frac{(2n+1)\pi k}{2N}), \quad 0 \leq p \leq M-1, 0 \leq k \leq N-1. \tag{6}$$

On the other hand, the MN-point DCT of $x(n)$ is expressed as:

$$C(k) = \sqrt{\frac{2}{M \times N}}\alpha(k)\sum_{n=0}^{M \times N-1} x(n)\cos(\frac{(2n+1)\pi k}{2 \times M \times N}), \quad 0 \leq k \leq MN-1. \tag{7}$$

In both of the above two equations $\alpha(k)$ is given by Eq. (2).

Jiang and Feng showed that a block DCT transformation as expressed by Eq. (6) is nothing but orthonormal expansion of the sequence $\{x(n)\}$ with a set of $M \times N$ basis vectors, each of which is derived from the basis vectors of the N-point DCT. Hence there exists an invertible linear transformation from M blocks of the N-point DCT transform to the usual MN-point DCT transform. In other words, for a sequence of N-point DCT blocks $\{C_i^{(N)}\}, i = 0, 1, ...M-1$, the corresponding composite DCT $C^{(MN)}$ (MN-point DCT), there exists a matrix $A_{(M,N)}$ of size $MN \times MN$ such that

$$C^{(M.N)} = A_{(M,N)}.[C_0^{(N)} C_1^{(N)} \ldots C_{M-1}^{(N)}]^T, \tag{8}$$

where the super-script "T" denotes the transpose of a matrix.

The analysis in the 1-dimension can be easily extended to the 2-dimension. In this case, for an $L \times M$ adjacent DCT blocks, a composite DCT block can be formed by the following equation:

$$C^{(LN \times MN)} = A_{(L,N)}.\begin{bmatrix} C_{0,0}^{(N\times N)} & C_{0,1}^{(N\times N)} & \cdots & C_{0,M-1}^{(N\times N)} \\ C_{1,0}^{(N\times N)} & C_{1,1}^{(N\times N)} & \cdots & C_{1,M-1}^{(N\times N)} \\ . & . & \cdots & . \\ . & . & \cdots & . \\ C_{L-1,0}^{(N\times N)} & C_{L-1,1}^{(N\times N)} & \cdots & C_{L-1,M-1}^{(N\times N)} \end{bmatrix}.A_{(M,N)}^T \tag{9}$$

Similarly, for decomposing a DCT block $C^{(L.N \times M.N)}$ to $L \times M$ DCT blocks of size $N \times N$ each, following expression is used:

$$\begin{bmatrix} C_{0,0}^{(N\times N)} & C_{0,1}^{(N\times N)} & \cdots & C_{0,M-1}^{(N\times N)} \\ C_{1,0}^{(N\times N)} & C_{1,1}^{(N\times N)} & \cdots & C_{1,M-1}^{(N\times N)} \\ . & . & \cdots & . \\ . & . & \cdots & . \\ C_{L-1,0}^{(N\times N)} & C_{L-1,1}^{(N\times N)} & \cdots & C_{L-1,M-1}^{(N\times N)} \end{bmatrix} = A_{(L,N)}^{-1}.C^{(L.N \times M.N)}.A_{(M,N)}^{-1^T} \tag{10}$$

It may be noted that the conversion matrices and their inverses are sparse. Hence, lesser number of multiplications and additions of two numbers are needed than those required in the multiplication two full matrices.

4. RESIZING WITH ARBITRARY FACTORS

In this section we propose algorithms for resizing by arbitrary factors in the block DCT space. First, we present algorithms with integer conversion factors and then for more general cases with arbitrary rational conversion factors.

4.1. Resizing with Integer Conversion Factors

While down-sampling an image a number of blocks are first recomposed to a single block. Finally, the recomposed block is sub-sampled to a smaller block using the sub-band DCT approximation. Hence in this case, if an image is to be reduced to a size by a factor of $L \times M$ (where L and M are two positive integers), one may consider $L \times M$ number of $N \times N$ DCT blocks and convert them into a block of $(LN \times MN)$-point DCT. We have used the sub-band approximation (extended from Eq. (5))in Eq. (11). Let $C_{LL}(k,l), 0 \leq k \leq N-1, 0 \leq l \leq N-1$ be the DCT coefficients of a $N \times N$ block of the down-sampled image. Then the DCT coefficients of a $LN \times MN$ block are approximated by the following equation.

$$C(k,l) = \begin{cases} \sqrt{LM} C_{LL}(k,l), & 0 \leq k,l \leq N-1 \\ 0, & \text{otherwise.} \end{cases} \quad (11)$$

The algorithm for down-sampling of images by $L \times M$ is described below.

4.1.1. Algorithm LM_Down-sampling (LMDS)

Input: 8×8 block based DCT encoding of an image.
Output: 8×8 block based DCT encoding of the $L \times M$ down-sampled image.
Begin

For every adjacent $L \times M$ blocks, $\{ B_{ij}, 0 \leq i \leq L-1, 0 \leq j \leq M-1 \}$ of the input image do the following:
{

(a) Compose a $LN \times MN$ DCT block according to equation (9).
(b) Get the resulting $N \times N$-DCT coefficients of the down-sampled image from equation (11).

}

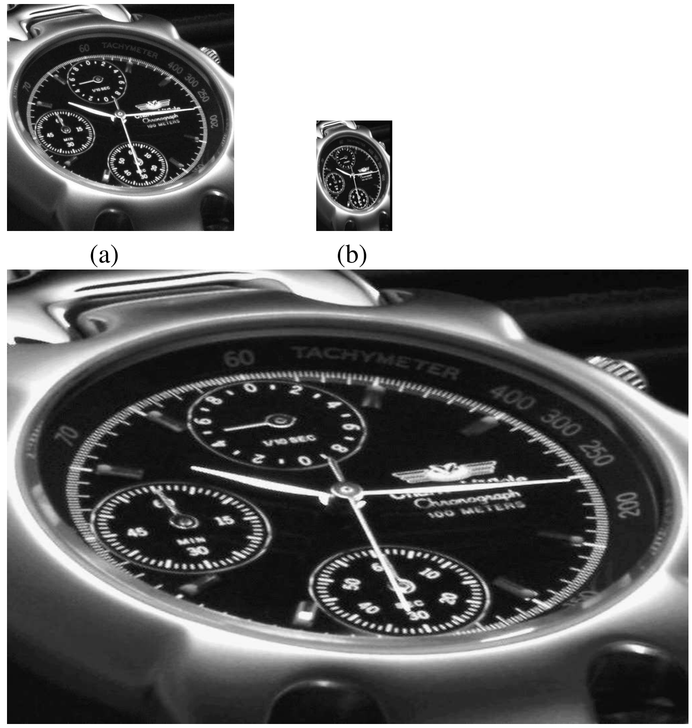
End *LM_Down-sampling (LMDS)*

One may observe that for $L = 2$ and $M = 2$, the operation is nothing but the image halving operation and the approach significantly differs from the earlier methods [7], [8]. Equation (11) may also be used for converting a $N \times N$ DCT block of an image to a $LN \times MN$ DCT block of the up-sampled image. After conversion, one may use DCT-block decomposition to obtain $L \times M$ numbers of $N \times N$-DCT blocks in the up-sampled image. The algorithm is described below:

4.1.2. Algorithm LM_Up-sampling (LMUS)

Input: $N \times N$ block based DCT encoding of an image.
Output: $L \times M$ Up-sampled image in the compressed domain.
Begin

For each $N \times N$ block B do the following:



(a) (b) (c)

Fig. 1. Image Resizing by a factor of 2×3: (a) Original, (b) down-sampled image, and (c) up-sampled image

(a) Approximate B to a $LN \times MN$ DCT block as follows:

$$\hat{B}^{(LN \times MN)} = \begin{bmatrix} \sqrt{LM}.B & Z_{(N,(M-1)N)} \\ Z_{((L-1)N,N)} & Z_{((L-1)N,(M-1)N)} \end{bmatrix}. \quad (12)$$

where $Z_{(a,b)}$ denotes a matrix of $a \times b$ zero elements.

(b) Decompose $\hat{B}^{(LN \times MN)}$ into $L \times M$ number of $N \times N$ DCT blocks using Eq. (10).

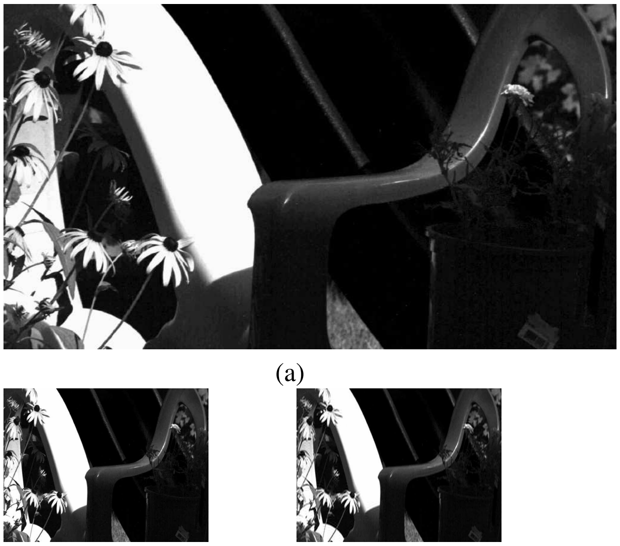
End *LM_Up-sampling(LMUS)*

Interestingly, the algorithm for image doubling as proposed in [8] is a special case (for $L = 2$ and $M = 2$) of the above one.

Typical examples of image resizing operations with a factor of 2×3 are shown in the Figure 1.

4.2. Resizing with Rational Conversion Factors

One can also use the integral factor resizing methods for resizing with a factor of rational number in both directions.



(a)

(b) (c)

Fig. 2. Conversion of a HDTV frame (1080×1920) to a NTSC frame (480×640) (a) HDTV (b) NTSC (up-sampling followed by down-sampling) (c) NTSC (down-sampling followed by up-sampling)

Let an image be resized by a factor of $\frac{P}{Q} \times \frac{R}{S}$, where P, Q, R and S are positive integers. This could be carried out by up-sampling an image with the integral factor of $P \times R$, followed by a down-sampling operation with the factor of $Q \times S$. One may also carry out the down-sampling operation (by a factor of $Q \times S$) before the up-sampling one (with $P \times R$ factor). This will speed up the computation. In Figure 2, we present a typical example of conversion of a HDTV frame[1] (of size 1020×1920) to a NTSC one (of size 480×640). We have used both the approaches, namely up-sampling followed by down-sampling and its reverse. Though the first one is computationally expensive, it provides a visually better quality picture than the latter one.

5. CONCLUSION

In this work we have discussed resizing algorithms for images represented in the block DCT space. The algorithms exploit the spatial relationship of the block DCT coefficients and sub-band correlation of the DCT coefficients. The proposed approach is general enough to accommodate resizing operations with arbitrary factors, namely with integral and rational factors. An interesting application of these methods demonstrated here is the conversion of images (video frames) to one format to another in the compressed domain itself.

[1]Obtained from *http://www.siliconimaging.com/hdtv-images.htm*

6. REFERENCES

[1] B.C.Smith and L. Rowe, "Algorithms for manipulating compressed images," *IEEE Comput. Graph. Applicat. Mag.*, vol. 13, no. Sept., pp. 34–42, 1993.

[2] N. Merhav and V. Bhaskaran, "Fast algorithms for DCT-domain image down-sampling and for inverse motion compensation," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 7, no. 6, pp. 468–476, 1997.

[3] S.-F.Chang and D. G. Messerschmitt, "Manipulation and composition of MC-DCT compressed video," *IEEE J. Selected Areas Commun.*, vol. 13, no. 1, pp. 1–11, 1995.

[4] A. Vetro, H. Sun, P. DaGraca, and T. Poon, "Minimum drift architectures for 3-layer scalable DTV decoding," *IEEE Trans. on Consumer Electronics*, vol. 44, no. 3, pp. 527–536, August, 1998.

[5] Q. Hu and S. Panchanathan, "Image/video spatial scalability in compressed domain," *IEEE Trans. on Ind. Electron.*, vol. 45, no. Feb., pp. 23–31, 1998.

[6] A. Neri, G. Russo, and P. Talone, "Inter-block filtering and downsampling in DCT domain," *Image Communication*, vol. 6, no. Aug., pp. 303–317, 1994.

[7] R. Dugad and N. Ahuja, "A fast scheme for image size change in the compressed domain," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 11, no. 4, pp. 461–474, 2001.

[8] J. Mukherjee and S.K. Mitra, "Image resizing in the compressed domain using subband DCT," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 12, no. 7, pp. 620–627, July 2002.

[9] H.W. Park, Y.S. Park, and S.K. Oh, "L/M-image folding in block DCT domain using symmetric convolution," *IEEE Trans. on Image Processing*, vol. 12, no. 9, pp. 1016–1034, Sept., 2003.

[10] S.A. Martucci, "Symmetric convolution and the discrete sine and cosine transforms," *IEEE Trans. on Signal Processing*, vol. 42, no. 5, pp. 1038–1051, May, 1994.

[11] J. Jiang and G. Feng, "The spatial relationships of DCT coefficients between a block and its sub-blocks," *IEEE Trans. on Signal Processing*, vol. 50, no. 5, pp. 1160–1169, May 2002.