DOWN-SAMPLING IN DCT DOMAIN USING LINEAR TRANSFORM WITH DOUBLE-SIDED MULTIPLICATION FOR IMAGE/VIDEO TRANSCODING

Xiang Yu, En-hui Yang, and Haiquan Wang

Dept. of Electrical and Computer Engineering
University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada

ABSTRACT

This paper proposes a designing framework for downsampling compressed images/video frames with arbitrary ratio in the discrete cosine transform (DCT) domain. We first derive a set of DCT-domain down-sampling methods which can be represented by a linear transform with double-sided matrix multiplication (LTDS) in the DCT domain, and show that the set contains a wide range of methods with various complexity and visual quality. Then, based on a pre-selected spatial-domain method, we formulate an optimization problem for finding an LTDS to approximate the given spatial domain method for achieving the best trade-off between the visual quality and the complexity. By selecting a spatial-domain reference method with the popular Butterworth lowpass filtering and bicubic interpolation, the proposed framework discovers LTDSs with better visual quality and lower computational complexity as saving 20%$\sim$70% execution time when compared with state-of-the-art methods in the literature.

Index Terms— DCT, down-sampling, transcoding.

1. INTRODUCTION

As the number of network users with diverse devices is steadily increasing, there has been a great momentum in the multimedia industry for supporting content display in diverse devices all over the network. A big challenge, however, is that these devices have various display resolutions ranging from large resolutions such as those of high-definition TV to small resolutions such as those of smart phones. This leads to research on transcoding, which involves in automatic reformatting. In this paper, we are concerned with transcoding with spatial resolution down-conversion due to its applications to wireless communications. In particular, we will focus on transcoding compressed image/video in the DCT domain because most image/video data to be shared over the network are originally captured with high resolution and coded using DCT-based techniques such as JPEG, DV, etc.

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada under Grants RGPIN203035-02 and RGPIN203035-06, by the Premier's Research Excellence Award, by the Ontario Distinguished Researcher Award, and by the Canada Research Chairs Program.

A straightforward method for down-sampling DCT images is to concatenate inverse DCT, spatial-domain down-sampling and DCT[9]. However, it is not suitable for some real-world transcoding applications because of its high complexity. To tackle the complexity issue, it is desired to perform down-sampling in the DCT domain directly. One category of methods for doing this is referred to as DCT coefficient manipulation [5], [6]. The manipulation methods generally encounter a problem of ringing effect[5], particularly for images with sharp contrast, because truncating DCT coefficients is equivalent to filtering by an ideal filter with a very narrow transition band[9], which leads to ringing effect.

Another category of DCT-domain methods are based on a transform model corresponding to a reference method selected in the spatial domain[1][2][8]. E.g., in [1], an efficient DCT-domain transform was developed based on a specific spatial domain method with the nearest neighbor approach. The omitting of lowpass filtering leads to an efficient DCT-domain transform, but it also limits the quality. On the other hand, a DCT-domain method based on a spatial-domain method with a freely-selected lowpass filter still has a high complexity for software implementation [2].

In this paper, we propose a designing framework for an efficient DCT-domain down-sampling transform corresponding to any given spatial-domain down-sampling method. First, a linear transform with double-sided matrix multiplication for DCT-domain down-sampling is derived corresponding to an arbitrarily selected spatial-domain down-sampling method. An optimization problem is then formulated to minimize a joint cost of the visual quality and the complexity based on the linear transform. By modeling the linear transform as a multiple layer network, a so-called structural learning with forgetting algorithm [7] is used for training, resulting in a desired trade-off between the quality and the complexity. In our simulations, the popular and well-recognized spatial-domain method with Butterworth lowpass filter and bicubic interpolation is used as the reference. The proposed framework discovers LTDSs with good subjective visual quality and low computational complexity with significant saving of execution time when compared with other methods.

The paper is organized as follows. Section 2 derives LTDSs. An optimization problem is formulated in Section 3 to find

an LTDS approximation for a given spatial-domain method, while solution is provided in Section 4. Experiments are discussed in Section 5, and conclusion is drawn in Section 6.

2. DCT-DOMAIN LINEAR TRANSFORM WITH DOUBLE-SIDED MATRIX MULTIPLICATION

This section derives a result that for a wide range of spatial-domain down-sampling methods with certain properties, a concatenation of inverse DCT, spatial-domain down-sampling and DCT can be implemented as an LTDS in the DCT domain.

Denote $\boldsymbol{t}_{\text{SS}}$ as a DCT matrix. Consider to down-sample an $\text{M}\times\text{N}$ DCT image $\boldsymbol{C}_{\text{MN}}$. First, apply inverse DCT to recover the original image $\boldsymbol{X}_{\text{MN}}$ as

$$\boldsymbol{X}_{\text{MN}} = \boldsymbol{t}'_{\text{SS}} \boxdot \boldsymbol{C}_{\text{MN}} \boxdot \boldsymbol{t}_{\text{SS}}, \tag{1}$$

where $\boxdot$ denotes block-wise multiplications. Then, a spatial-domain method is used to down-sample $\boldsymbol{X}_{\text{MN}}$ to an $\text{I}\times\text{J}$ image $\boldsymbol{x}_{\text{IJ}}$. Finally, a DCT-domain down-sampling result is

$$\boldsymbol{V}_{\text{IJ}} = \boldsymbol{t}_{\text{SS}} \boxdot \boldsymbol{x}_{\text{IJ}} \boxdot \boldsymbol{t}'_{\text{SS}}. \tag{2}$$

Consider a general spatial-domain down-sampling method. First apply a 2D low-pass filter based on the 2D discrete Fourier transform (DFT), i.e.,

$$\tilde{\boldsymbol{X}}_{\text{MN}} = \boldsymbol{A}^*_{\text{MM}}((\boldsymbol{A}_{\text{MM}} \cdot \boldsymbol{X}_{\text{MN}} \cdot \boldsymbol{B}_{\text{NN}}) \otimes \boldsymbol{F}_{\text{MN}}) \cdot \boldsymbol{B}^*_{\text{NN}}, \tag{3}$$

where $\tilde{\boldsymbol{X}}_{\text{MN}}$ is the low-pass filtering output, $\boldsymbol{A}_{\text{MM}}$ is a standard $\text{M}\times\text{M}$ DFT transform matrix and the $*$ represents the conjugate. $\boldsymbol{B}_{\text{NN}}$ is a standard N×N DFT transform matrix. $\boldsymbol{F}_{\text{MN}}$ is a low-pass filtering matrix in the DFT domain. The symbol $\otimes$ denotes element-wise multiplications. Then, consider to implement a 2D interpolation with two 1D linear interpolations along the width and the height separately. Denote $\boldsymbol{E}_{\text{IM}}$ and $\boldsymbol{G}_{\text{NJ}}$ as the interpolation matrixes. The down-sampling output is

$$\boldsymbol{x}_{\text{IJ}} = \boldsymbol{E}_{\text{IM}} \cdot \tilde{\boldsymbol{X}}_{\text{MN}} \cdot \boldsymbol{G}_{\text{NJ}}. \tag{4}$$

Combining equations (1) to (4), we have

$$\boldsymbol{V}_{\text{IJ}} = \boldsymbol{t} \boxdot [\boldsymbol{E}_{\text{IM}} \boldsymbol{A}^*_{\text{MM}} [(\boldsymbol{A}_{\text{MM}} (\boldsymbol{t}' \boxdot \boldsymbol{C}_{\text{MN}} \boxdot \boldsymbol{t}) \boldsymbol{B}_{\text{NN}}) \otimes \boldsymbol{F}_{\text{MN}}] \boldsymbol{B}^*_{\text{NN}} \boldsymbol{G}_{\text{NJ}}] \boxdot \boldsymbol{t}', \tag{5}$$

where $\boldsymbol{t} = \boldsymbol{t}_{\text{SS}}$. The $\boxdot$ in (5) can be replaced by applying a result of $\boldsymbol{t} \boxdot \boldsymbol{C}_{\text{MN}} = \boldsymbol{T}_{\text{MM},\boldsymbol{t}} \cdot \boldsymbol{C}_{\text{MN}}$, where $\boldsymbol{T}_{\text{MM},\boldsymbol{t}} = \text{diag}(\boldsymbol{t}, \cdots, \boldsymbol{t})$. The $\otimes$ in (5) may be removed under one condition that the lowpass filter takes a form of

$$\boldsymbol{F}_{\text{MN}} = L_{\text{M1}} \cdot R_{1\text{N}}. \tag{6}$$

Consequently, (5) can be rewritten as a linear transform,

$$\boldsymbol{V}_{\text{IJ}} = \boldsymbol{D}^{\text{o}}_{\text{IM}} \cdot \boldsymbol{C}_{\text{MN}} \cdot \boldsymbol{W}^{\text{o}}_{\text{NJ}}, \tag{7}$$

$$\text{where} \quad \begin{aligned} \boldsymbol{D}^{\text{o}}_{\text{IM}} &= \boldsymbol{T}_{\text{II},\boldsymbol{t}} \boldsymbol{E}_{\text{IM}} \boldsymbol{A}^*_{\text{MM}} \boldsymbol{L}_{\text{MM}} \boldsymbol{A}_{\text{MM}} \boldsymbol{T}_{\text{MM},\boldsymbol{t}'}, \\ \boldsymbol{W}^{\text{o}}_{\text{NJ}} &= \boldsymbol{T}_{\text{NN},\boldsymbol{t}} \boldsymbol{B}_{\text{NN}} \boldsymbol{R}^*_{\text{NN}} \boldsymbol{B}_{\text{NN}} \boldsymbol{G}_{\text{NJ}} \boldsymbol{T}_{\text{JJ},\boldsymbol{t}'}, \end{aligned}$$

and $\boldsymbol{L}_{\text{MM}}$ and $\boldsymbol{R}_{\text{NN}}$ are diagonal matrixes with diagonal elements being L_{M1} and $R_{1\text{N}}$, respectively.

Motivated by (7), we call any DCT domain transform in the form of $\boldsymbol{D}_{\text{IM}} \cdot \boldsymbol{C}_{\text{MN}} \cdot \boldsymbol{W}_{\text{NJ}}$ a DCT-domain LTDS, where $\boldsymbol{D}_{\text{IM}}$ and $\boldsymbol{W}_{\text{NJ}}$ are arbitrary matrixes with respective dimensions. We are interested in the set S of all DCT-domain LTDSs. Clearly, as shown in (7), the set contains all spatial domain methods satisfying (4) and (6). Some LTDSs in the set have high complexity while others have low complexity. Given any spatial-domain method, in the following we want to find its LTDS approximation in the set, which gives the best trade-off between the quality and complexity.

3. PROBLEM FORMULATION

Fix a spatial-domain method A. Applying A to any DCT image $\boldsymbol{C}_{\text{MN}}$, one gets a corresponding down-sampled DCT image $\boldsymbol{V}_{\text{IJ}}$. Our purpose is to find an LTDS in the set S that, when applied to $\boldsymbol{C}_{\text{MN}}$, outputs an image with similar quality to $\boldsymbol{V}_{\text{IJ}}$ with a low complexity.

3.1. Visual Quality Measurement

Quality measure is the very basis to formulate our optimization problem for finding optimal LTDSs corresponding to a pre-selected spatial-domain down-sampling method. There have been two major objective quality measures used in the literature for image down-sampling. The first one is to measure the quality with a reference image obtained using a standard spatial-domain down-sampling method [1]. The second one is to up-sample the down-sampled image to the original resolution. Then, the quality is measured by the MSE between the up-sampled image and the original one [5].

In this paper, we apply the first measure, which will naturally allow us to approach the visual performance of a pre-selected spatial-domain method by setting the down-sampling output of the pre-selected method as the reference. Then, we may be able to approach the best visual quality by pre-selecting a spatial-domain method with the best visual quality if there is one. Since we are interested in viewing a down-sampled image in its own resolution, there is no up-sampling process involved, and hence the first measure is more appropriate than the second measure for our purpose.

Specifically, the quality measure in the paper is, for any $g(\cdot) \in S$, the quality of $g(\boldsymbol{C}_{\text{MN}})$ is measured by $|\boldsymbol{V}_{\text{IJ}} - g(\boldsymbol{C}_{\text{MN}})|^2$, where $\boldsymbol{V}_{\text{IJ}}$ is obtained using a given spatial-domain method.

3.2. A complexity model

In general, the complexity for computing $\boldsymbol{D}_{\text{IM}} \cdot \boldsymbol{C}_{\text{MN}} \cdot \boldsymbol{W}_{\text{NJ}}$ is related to two facts. The first is the number of non-zero elements in $\boldsymbol{D}_{\text{IM}}$ and $\boldsymbol{W}_{\text{NJ}}$. The second is how multiplications may be implemented by additions and shifts. Initially, we consider a complexity model of $r_{\text{f}} = |\boldsymbol{D}_{\text{IM}}| + |\boldsymbol{W}_{\text{NJ}}|$, where $|\cdot|$ defines the l_1 norm of a matrix. By a learning with forgetting stage of SLF[7], the minimization of $|\boldsymbol{D}_{\text{IM}}| + |\boldsymbol{W}_{\text{NJ}}|$ will force as many as possible elements to be zero.

The complexity model r_{f} is further adjusted by considering to implement a multiplication with a series of additions

and shifts. First, we introduce a quantization procedure as follows,

$$\mathrm{Q}(x) = \sum_{i=-2}^{i=15} a_i \cdot 2^{-i}, \;\; a_i \in \{1, -1, 0\} \tag{8}$$

$$\text{where } \{a_i\} = \arg \min_{|x - \sum(a_i 2^{-i})| \le |x|\eta} \sum |a_i|,$$

where η is a small constant, and $x \in (-8, 8)$ [1]. Essentially, this quantization procedure leads to an approximation within a given neighboring region with the minimal number of ones in the binary representation. Thus, the corresponding multiplication may be implemented with the minimal number of shifts and additions. Overall, the complexity model becomes,

$$\begin{aligned} r_{\mathrm{q}} &= (|\boldsymbol{D}_{\mathrm{IM}}|_{|d_{im}|<d_0} + |\boldsymbol{W}_{\mathrm{NJ}}|_{|w_{nj}|<w_0}) \\ &\quad +\rho \cdot (|\boldsymbol{D}_{\mathrm{IM}} - Q(\boldsymbol{D}_{\mathrm{IM}})| + |\boldsymbol{W}_{\mathrm{NJ}} - Q(\boldsymbol{W}_{\mathrm{NJ}})|), \end{aligned} \tag{9}$$

where ρ is a constant, and $Q(\boldsymbol{D}_{\mathrm{IM}})$ and $Q(\boldsymbol{W}_{\mathrm{NJ}})$ mean to apply $Q(\cdot)$ to each element of $\boldsymbol{D}_{\mathrm{IM}}$ and $\boldsymbol{W}_{\mathrm{NJ}}$.

3.3. Joint Optimization of Visual Quality and Complexity

A joint optimization problem is then formulated as follows,

$$\min_{\boldsymbol{D}_{\mathrm{IM}}, \boldsymbol{W}_{\mathrm{NJ}}} ||\boldsymbol{D}_{\mathrm{IM}} \boldsymbol{C}_{\mathrm{MN}} \boldsymbol{W}_{\mathrm{NJ}} - \boldsymbol{V}_{\mathrm{IJ}}||^2 + \lambda \cdot r. \tag{10}$$

Clearly, knowledge of down-sampling in the spatial domain has been utilized to set up a benchmark of the quality for designing a down-sampling algorithm in the DCT domain. The objective is to find an LTDS with a desired trade-off between the visual quality and the complexity of r, which takes value of r_{f} or r_{q}.

4. PROBLEM SOLUTION

The optimization problem (10) is solved by training the 3-layer network shown in Figure 1 using the SLF algorithm [7]. Specifically, the learning procedure includes two stages, i.e., learning with forgetting and learning with selective forgetting, with $r = r_{\mathrm{f}}$ and $r = r_{\mathrm{q}}$, respectively. Given training data $(\boldsymbol{C}_{\mathrm{MN}}, \boldsymbol{V}_{\mathrm{IJ}})$, the learning with forgetting stage is as follows:

1. Pass $\boldsymbol{C}_{\mathrm{MN}}$ forward to compute the network outputs.

$$\boldsymbol{Y}_{\mathrm{IN}} = \boldsymbol{D}_{\mathrm{IM}} \cdot \boldsymbol{C}_{\mathrm{MN}} \quad \Rightarrow \quad \boldsymbol{Z}_{\mathrm{IJ}} = \boldsymbol{Y}_{\mathrm{IN}} \cdot \boldsymbol{W}_{\mathrm{NJ}}$$

2. Compute the network error, and propagate it backward.

$$\Delta \boldsymbol{Z}_{\mathrm{IJ}} = \boldsymbol{Z}_{\mathrm{IJ}} - \boldsymbol{V}_{\mathrm{IJ}} \quad \Rightarrow \quad (\Delta \boldsymbol{Y})_{\mathrm{IN}} = (\Delta \boldsymbol{Z})_{\mathrm{IJ}} \cdot (\boldsymbol{W}^{\mathrm{t}})_{\mathrm{JN}}$$

3. Compute the learning amount for $\boldsymbol{D}$ and $\boldsymbol{W}$.

$$\begin{aligned} \Delta \boldsymbol{D} &= (\Delta \boldsymbol{Y})_{\mathrm{IN}} \cdot (\boldsymbol{C}^{\mathrm{t}})_{\mathrm{NM}} + \lambda_1 \cdot \mathrm{sgn}(\boldsymbol{D}_{\mathrm{IM}}) \\ \Delta \boldsymbol{W} &= (\boldsymbol{Y}^{\mathrm{t}})_{\mathrm{NI}} \cdot (\Delta \boldsymbol{Z})_{\mathrm{IJ}} + \lambda_1 \cdot \mathrm{sgn}(\boldsymbol{W}_{\mathrm{NJ}}) \end{aligned} \tag{11}$$

where $\mathrm{sgn}(\cdot)$ is the sign function.

[1] This range is set empirically as observation shows that elements in $\boldsymbol{D}_{\mathrm{IM}}$ and $\boldsymbol{W}_{\mathrm{NJ}}$ have magnitudes strictly smaller than 1.

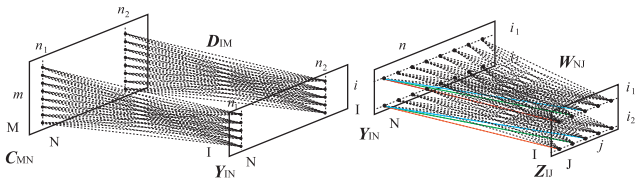

Fig. 1. A network structure and connections.

4. Learn with error propagation and forgetting.

$$\begin{aligned} \boldsymbol{D}_{\mathrm{NJ}}^{(n+1)} &= \boldsymbol{D}_{\mathrm{NJ}}^{(n)} - \alpha \cdot \Delta \boldsymbol{D} \\ \boldsymbol{W}_{\mathrm{NJ}}^{(n+1)} &= \boldsymbol{W}_{\mathrm{NJ}}^{(n)} - \alpha \cdot \Delta \boldsymbol{W} \end{aligned}$$

where α is a constant, the superscripts (n) and $(n+1)$ accord to the nth and $(n+1)$th iterations. Note that the superscripts are omitted in steps 1 to 4 for simplicity.

5. Repeat steps 1 to 4 until the decrement of J_f is smaller than a given threshold.

The above learning stage ends with a skeleton structure but a large distortion of $||\boldsymbol{D} \cdot \boldsymbol{C} \cdot \boldsymbol{W} - \boldsymbol{V}||^2$. The selective forgetting stage is then used to tune the structure for a better trade-off between the distortion and the complexity. Compared with the above procedure, the selective forgetting stage is mostly the same, except the computation of (11). See [3] for details.

The design procedure is summarized as follows,

1. Generate a training set: Choose some pictures. Given a pixel-domain down-sampling method, apply it to these images. Compute DCT transform for the original images and the down-sampled versions as the training set.

2. Learning with forgetting: Construct the 3-layer structure with $\boldsymbol{D}_{\mathrm{IM}}$ and $\boldsymbol{W}_{\mathrm{NJ}}$. Find a skeleton structure using the learning with forgetting algorithm.

3. Learning with selective forgetting: Refine $\boldsymbol{D}_{\mathrm{IM}}$ and $\boldsymbol{W}_{\mathrm{NJ}}$ with the learning with selective forgetting algorithm.

5. SIMULATION RESULTS

The proposed design algorithm has been implemented and applied to discover down-sampling methods in the DCT domain with good quality and low complexity. A spatial-domain method with the 10th order Butterworth low-pass filtering and bicubic interpolation is selected to generate reference images for evaluating the visual quality among LTDSs.

The LTDS corresponding to $d_0 = w_0 = 0.1$ makes a good choice for down-sampling by 2:1 in terms of a better quality and a lower complexity when compared with other algorithms. Table 1 provides a comprehensive comparison for our obtained LTDS with four other algorithms in the literature. The method in [5] is developed for down-sampling by a factor of 2 based on DCT coefficient manipulation while the method

Table 1. Performance comparison of five methods for down-sampling JPEG images/DV frames with a ratio of 2:1. Complexity is measured with number of operations per pixel in the original image, while computation time is reported based on our computer with 3.4Ghz P-IV CPU. The visual quality is measured by subjective criterions for a testing set of 20 images. Note that a 'yes' means that the corresponding effect shows up for some, not necessarily all, images in the whole set, while a 'no' means that the corresponding effect has not been observed for all images in the set. Visual illustration is omitted due to the page limitation and the fact that all images need to be illustrated in original sizes in order not to introduce resizing distortion. See [3] for visual illustration.

	Complexity			Visual quality			Computation time
	MUL	ADD	SHL	Ringing	Aliasing	Other artifacts	per image
LTDS with ($d_0 = w_0 = 0.1$)	0	5.16	3.66	no	no	no	1.6ms
Method in [5]	1.25	1.25	0	yes	no	no	2.1ms
L/M [6]	3.31	8.68	2.2	yes	no	no	6.3ms
Bilinear average in [8]	3.75	5.81	0.38	yes	yes	no	2.9ms
Fast algorithm in [8]	0	2.72	0.72	yes	yes	severe	0.9ms

Table 2. Performance comparison for down-sampling JPEG images/DV frames with a ratio of 3:2.

	Complexity			Visual quality	Computation time
	MUL	ADD	SHL	Ringing effect	per image
LTDS	0.26	10.4	14.3	no	3.2ms
L/M [6] (II)	1.94	8.06	0	yes	4.1ms

in [6] shares a similar spirit of DCT coefficient manipulation but it is extended to support arbitrary down-sampling ratios. The work in [8] targeted down-sampling by 2:1 with 8×8 DCT. The fast algorithm in [8] is the only one more efficient than the obtained LTDS, yet it has severe artifacts. In term of visual quality, the method in [5] is similar with the obtained LTDS as it only shows a slight ringing effect, while the obtained LTDS shows a 20% faster execution speed in our simulation.

The proposed framework is applicable for generating down-sampling algorithms with arbitrary ratios. Experiments have been conducted to generate down-sampling algorithms with a ratio of 3:2. The obtained LTDS for down-sampling by 3:2 with $d_0 = w_0 = 0.02$ is compared with the L/M method in [6], since the method in [8] is for 2:1 only and the work in [5] targets for down-sampling by a factor of 2. The result is shown in Table 2, with a focus on the complexity. Essentially, the obtained LTDS has a lower complexity than the method in [6]. Specifically, there are two algorithms proposed for down-sampling by 3:2 in [6], referred to as case I and case II, with case II being the enhanced version of case I. Experimental results by the computation time show that the obtained LTDS is more efficient than the case II algorithm.

6. CONCLUSIONS

In this paper, we have proposed a design framework for efficient down-sampling in the DCT domain by jointly optimizing the visual quality and the computational complexity. First, a DCT-domain LTDS is derived corresponding to a concatenation of inverse DCT, spatial-domain down-sampling, and DCT. Based on the LTDS model, an optimization problem is formulated to find an optimal DCT domain method for a given spatial-domain method. Then, the problem is solved using an automatic machine learning algorithm called structure learning with forgetting, resulting in a desired trade-off between the visual quality and the computational complexity. Experiments on JPEG image and DV video transcoding show that the down-sampling methods discovered by the proposed design framework outperforms other methods with a significant reduction of computation time by 20%$\sim$70% when similar or better visual quality is maintained.

7. REFERENCES

[1] N. Merhav, V. Bhaskaran, "Fast Algorithms for DCT-Domain Image Down-sampling and for Inverse Motion Compensation", *IEEE Transactions on Circuits and Systems for Video Technology*, pp.468-476, Vol. 7, No. 3, June 1997.

[2] J.B. Lee, A. Eleftheriadis, "2-D Transform-Domain Resolution Translation", *IEEE Transactions on Circuits and Systems for Video Technology*, pp.704-714, Vol. 10, No. 5, August 2000.

[3] X. Yu, E.-h. Yang, and H.Q. Wang, "Down-sampling Design in DCT Domain with Arbitrary Ratio for Image/Video Transcoding," Submited to IEEE Trans. on Image Processing.

[4] S.F. Chang, D. G. Messerschmitt, "Manipulation and Compositing of MC-DCT COmpressed Video", *IEEE Journal on Selected Areas in Communications*, ppl-11, Vol. 13, No. 1, January 1995.

[5] R. Dugad, and N. Ahuja, "A Fast Scheme for Image Size Change in the Compressed Domain", *IEEE Transactions on Circuits and Systems for Video Technology*, pp.461-474, Vol.11, No. 4, April 2001.

[6] Y.S. Park, and H.W. Park, "Arbitrary-Ratio Image Resizing using Fast DCT of Composite Length for DCT-based Transcoder" IEEE Transactions on Image Processing, pp.494-500, Vol. 15, No.2, Feb. 2006.

[7] M. Ishikawa, "Structral Learning with Forgetting" *Neural Networks*, Vol. 9, No. 3, pp.509-521, 1996.

[8] B. K. Natarajan, V. Bhaskaran, "A Fast Approximate Algorithm for Scaling Down Digital Images in the DCT Domain", in *Proc. IEEE Int. Conf. Image Processing'1995*, pp. 241-243.

[9] W. K. Pratt, *Digital Image Processing*, John Wiley & Sons, Inc, 1991.