DESIGN OF OPTIMIZED PREFILTERS FOR TIME-DOMAIN LAPPED TRANSFORMS WITH VARIOUS DOWNSAMPLING FACTORS

Masaki Onuki and Yuichi Tanaka

Graduate School of BASE, Tokyo University of Agriculture and Technology
Koganei, Tokyo, 184-8588 Japan
Email: masaki.o.0806@msp.lab.tuat.ac.jp, ytnk@cc.tuat.ac.jp

ABSTRACT

In image compression, decimation and interpolation techniques as pre-/postprocessing of a compression framework offer excellent energy compaction capability for low bit rate image coding. However, an interpolation filter cannot be defined as the inverse matrix of the given decimation filter since the former is a wide matrix and the latter is a tall one. Even without any compression, there will be some distortions in the reconstructed image. To tackle the problem, interpolation-dependent image downsampling (IDID) method produces optimized downsampled images, which leads to the optimized prefilter of a given postfilter. In this paper, we propose to integrate IDID with time-domain lapped transforms (TDLTs) to improve image coding performances.

Index Terms— Biorthogonal filter bank, lapped transform, image coding, downsampling, interpolation

1. INTRODUCTION

Block coding based on the discrete cosine transform (DCT) is very popular in image and video compression [1]-[3]. Its success stems from the DCT's excellent energy compaction capability, low computational complexity, and high degree of flexibility. However, block coding based on the DCT leads to blocking artifacts at low bit rate because correlation between neighboring blocks has not been taken into account. To overcome this problem, time-domain lapped transform (TDLT) was proposed [4]. It realizes a LT [5] with pre- and postprocessings outside the existing DCT-based coding framework, i.e., JPEG.

Image coding with decimation and interpolation is an alternative approach to achieve better performance for low bit rate image coding, where images are decimated prior to compression and the missing portions are interpolated after decompression [6]. With this approach, computational complexity is reduced at coding/decoding processes, and reconstructed image quality may be extremely improved at low bit rate.

An interpolation filter cannot be defined as the inverse matrix of the given decimation filter since the former is a wide matrix and the latter is a tall one. Therefore, the interpolated image of the decimated original image produces distortions even without compression. To reduce the amount of degradation, interpolation-dependent image downsampling (IDID) was proposed [7]. This method iteratively optimizes the decimation filter to find a convergent point. Eventually, when the iteration is finished, this offers the optimal downsampled image of a given interpolation filter. However, it does not consider distortions generated by quantization in image coding.

TDLT-based decimation/interpolation filters were proposed as under-sampled boundary pre-/postfilters [8] and over-sampled ones [9]. This method is an extension of [4] from critically-sampled systems to under-sampled/over-sampled systems. TDLTs are sufficiently considered and designed for image coding. However, it assumes the AR(1) process to design filter banks. Hence it is not optimal for a given specific signal and/or image.

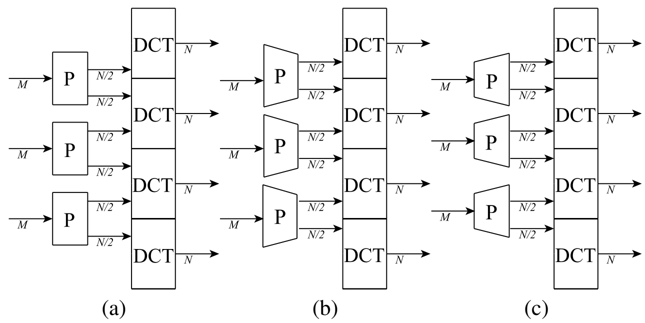

Fig. 1. Time-domain lapped transforms. The prefilter $\mathbf{P}$ is of size $N \times M$. (a) critically-sampled TDLT. (b) under-sampled TDLT ($N < M$). (c) over-sampled TDLT ($N > M$).

Recently, we proposed combined TDLTs which optimizes the downsampling factors of prefilters [10]. In this method, the optimal combination of a set of prefilters with various downsampling factors is decided by a target quality parameter QP of JPEG. In this paper, we further extend our combined TDLTs to determine the optimized prefilters as well as the optimized downsampling factors for a given portion in an input image. In the experimental results, the appropriately designed prefilters show high compression performances.

Notation: Upper case bold-face letters indicate matrices. A subscript of a matrix represents its size if there is no explicit expression of the subscript. Let $\mathbf{I}_M$, $\mathbf{J}_M$, and $\mathbf{0}_{M\times N}$ denote the $M \times M$ identity, the $M \times M$ reversal identity, and the $M \times N$ null matrix, respectively. Additionally, let $\mathbf{X}$ be input image as long as it is not explicitly declared. Superscript $\cdot^T$ and $\cdot^+$ are the transpose of the matrix and pseudo inverse matrix, respectively. Finally, diag$(\cdot)$ is a block diagonal matrix.

2. IDID

IDID derives an optimized downsampled image from a given interpolation matrix. Let $\mathbf{X}_v$, $\mathbf{Y}$, $\hat{\mathbf{X}}_v$, and $\mathbf{H}$ be the (vectorized) input image, the decimated image, the interpolated image, and the interpolation matrix, respectively. Then, we consider the expression

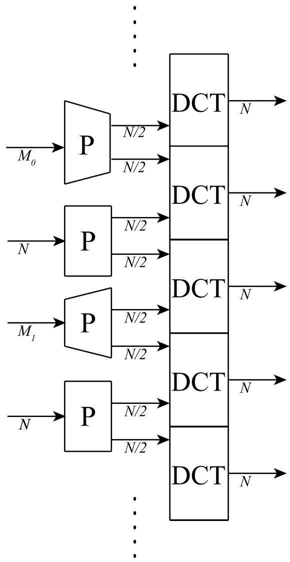

Fig. 2. Combined TDLTs.

$$\mathbf{Y}_{\rm opt} = \arg\min_{\mathbf{Y}} \|\hat{\mathbf{X}}_v - \mathbf{X}_v\|^2. \tag{1}$$

This expression derives the downsampled image that minimizes the distortion between the reconstructed image and the input image. $\hat{\mathbf{X}}_v$ is represented as

$$\hat{\mathbf{X}}_v = \mathbf{H}\mathbf{Y}. \tag{2}$$

Substituting (2) into (1), the objective function to derive the optimal downsampled image can be expressed as

$$\mathbf{J} = \|\mathbf{H}\mathbf{Y} - \mathbf{X}_v\|^2. \tag{3}$$

Setting the partial derivative of $\mathbf{J}$ to be zero, we derive

$$\frac{\partial \mathbf{J}}{\partial \mathbf{Y}} = 2\mathbf{H}^T(\mathbf{H}\mathbf{Y} - \mathbf{X}_v) = \mathbf{0}. \tag{4}$$

Then, the optimal downsampled image can be expressed as

$$\mathbf{Y} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{X}_v. \tag{5}$$

The optimization algorithm is shown below:

Step 1 Initialize $\mathbf{H}^0$ and $\mathbf{Y}^0$.
Step 2 Compute $\mathbf{H}^i$ based on $\mathbf{Y}^{i-1}$.
Step 3 Compute $\mathbf{Y}^i$ according to (5).
Step 4 Compute $E(i)$. Set i to $i+1$, then if $E(i) > T$, go to Step 2.

$E(i)$ calculates the mean squared error between $\mathbf{Y}^i$and $\mathbf{Y}^{i-1}$ and T is a threshold. From the above, the optimized downsampled image is produced. In short, it is equivalent to the optimization of the downsampling matrix for a given interpolation matrix.

3. TDLT

Let $\mathbf{P}$ and $\mathbf{T}$ denote the prefilter of size $N \times M$ and the postfilter of size $M \times N$, respectively. The analysis polyphase matrix $\mathbf{E}(z)$ and the synthesis one $\mathbf{R}(z)$ of TDLTs are represented as follows:

$$\mathbf{E}(z) = \mathbf{C}_N \mathbf{\Lambda}_N(z)\mathbf{P} \tag{6}$$

$$\mathbf{R}(z) = \mathbf{T}\mathbf{\Lambda}_N^{-1}(z)\mathbf{C}_N^T \tag{7}$$

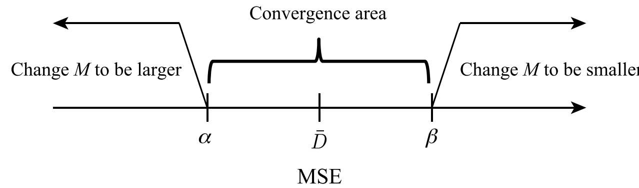

Fig. 3. The permissible range of distortion.

where $\mathbf{C}_N$ is the $N \times N$ type-II DCT matrix and $\mathbf{\Lambda}_N(z) = \begin{bmatrix} \mathbf{0}_{N/2} & \mathbf{I}_{N/2} \\ z\mathbf{I}_{N/2} & \mathbf{0}_{N/2} \end{bmatrix}$. When $M = N$, $\mathbf{P}$ and $\mathbf{T}$ are

$$\mathbf{P} = \mathbf{W}_N \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{V} \end{bmatrix} \mathbf{W}_N \tag{8}$$

$$\mathbf{T} = \mathbf{W}_N \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}^{-1} \end{bmatrix} \mathbf{W}_N, \tag{9}$$

in which $\mathbf{V}$ is an $N/2 \times N/2$ invertible matrix and

$$\mathbf{W}_N = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I}_{N/2} & \mathbf{J}_{N/2} \\ \mathbf{J}_{N/2} & -\mathbf{I}_{N/2} \end{bmatrix}. \tag{10}$$

It is the critically-sampled (CS) TDLT [4]. When $M > N$ and $M < N$, (6) and (7) become the under-sampled (US) and over-sampled (OS) TDLTs, respectively [8, 9]. In Fig. 1, three types of TDLTs are shown. $\mathbf{P}$ and $\mathbf{T}$ can be modified as follows:

$$\mathbf{P} = \mathbf{W}_N \begin{bmatrix} \mathbf{U} & \mathbf{0} \\ \mathbf{0} & \mathbf{V} \end{bmatrix} \mathbf{W}_M \tag{11}$$

$$\mathbf{T} = \mathbf{W}_M \begin{bmatrix} \mathbf{U}^+ & \mathbf{0} \\ \mathbf{0} & \mathbf{V}^+ \end{bmatrix} \mathbf{W}_N, \tag{12}$$

where $\mathbf{U}$ and $\mathbf{V}$ are $M/2 \times N/2$ matrices.

The US TDLT performs well in low bit rate image coding. In high bit rate image coding, OS TDLT outperforms the CS TDLT. Hereafter, a TDLT is specified with a label LT-NM. For example, LT-88 and LT-816 are CS TDLT and US TDLT, respectively.

4. COMBINED TDLTS FOR VARIOUS DOWNSAMPLING FACTORS

The architecture of combined TDLTs with various downsampling factors (combined TDLT in short) [10] is very simple since all the TDLTs can be represented as pre-/postfilters of the DCT. For example, a combined TDLT is shown in Fig. 2. It focuses on improving image coding performance by changing downsampling factors of TDLTs depending on characteristics of the input image.

Since TDLTs are separable transforms, downsampling factors for vertical and horizontal directions can be separately optimized. In this section, its algorithm is described for the vertical direction. The optimization for the horizontal direction can be applied straightforwardly.

The input image $\mathbf{X}$ is evenly partitioned into L nonoverlapping rows, each denoted by $\mathbf{R}_l$, where $l = 0, \ldots, L-1$. Then, we compute the distortion due to compression as follows:

$$D_l = \mathrm{MSE}(\mathbf{R}_l, \mathbf{R}'_l), \tag{13}$$

where $\mathrm{MSE}(\mathbf{x}, \mathbf{y})$ calculates mean squared error between $\mathbf{x}$ and $\mathbf{y}$ and

$$\mathbf{R}'_l = \mathbf{P}_l^+ \mathrm{IDCT_{2D}}(\mathrm{Quant}(\mathrm{DCT_{2D}}(\mathbf{P}_l\mathbf{R}_l), QP)), \tag{14}$$

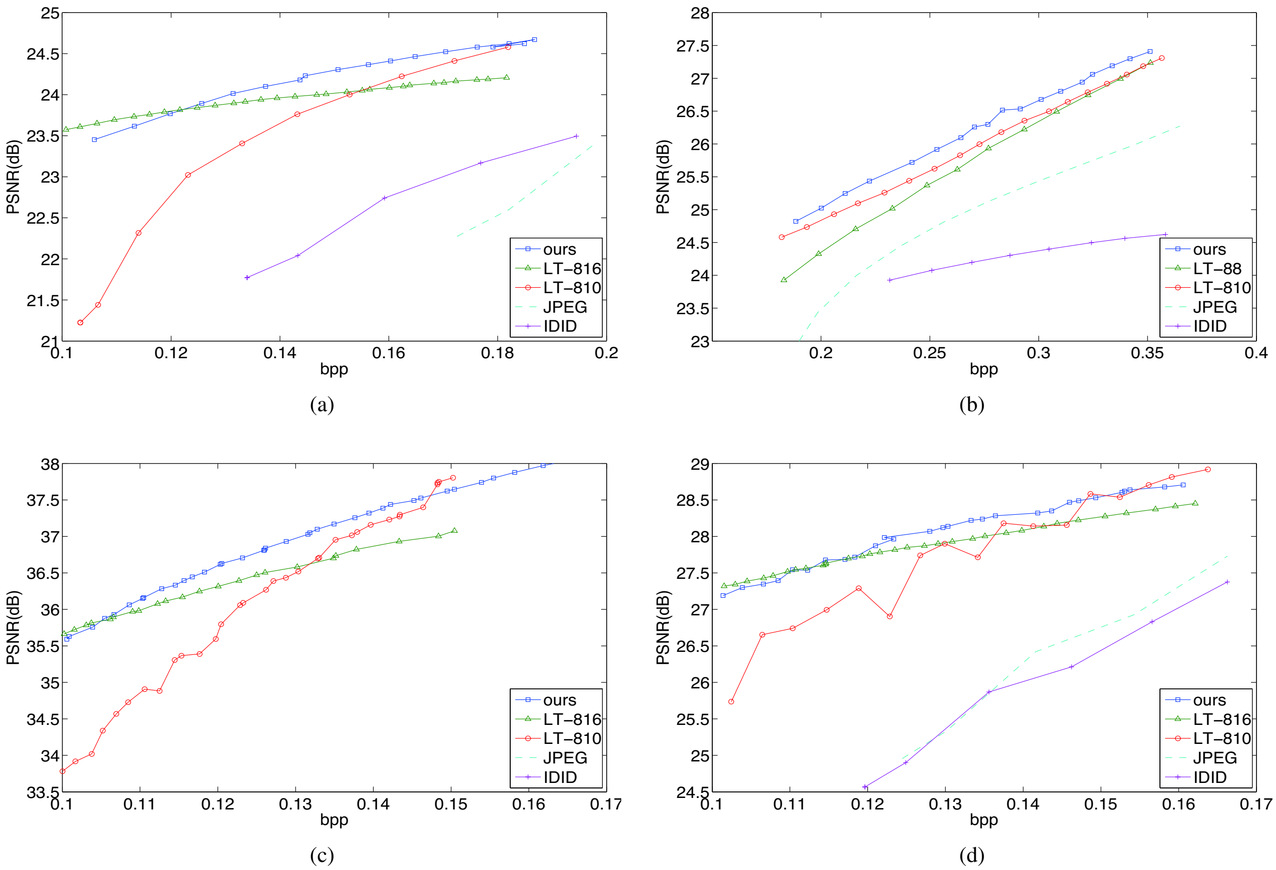

Fig. 4. R-D curves for JPEG-based compression. (a) *Babara* (low bit rate). (b) *Babara* (high bit rate). (c) *Birds* (IDID and JPEG < 32dB). (d) *Kodim20*.

where $\mathrm{Quant}(\mathbf{x}, QP)$ is quantization operator of a signal $\mathbf{x}$ by a quantization matrix of JPEG at a quality parameter QP, and

$$\mathbf{P}_l = \mathrm{diag}(\mathbf{B}_u, \mathbf{P}, \cdots, \mathbf{P}, \mathbf{B}_l), \tag{15}$$

where $\mathbf{B}_u$ and $\mathbf{B}_l$ are the upper and lower boundary transformation matrices corresponding to $\mathbf{P}$. Furthermore, $\mathrm{DCT}_{\mathrm{2D}}(\cdot)$ and $\mathrm{IDCT}_{\mathrm{2D}}(\cdot)$ are 2-D DCT and IDCT, respectively. In this method, a set of prefilters is optimized as the minimization of the joint cost function including bit rate and D_l.

5. OPTIMIZED PREFILTER FOR TDLTS WITH VARIOUS DOWNSAMPLING FACTORS

In this section, we propose a method to optimize the prefilter by integrating IDID with the combined TDLTs. As mentioned above, the pre-/postfilters ($\mathbf{P}$ and $\mathbf{T}$) of TDLTs are *not* optimized for a specific given signal, whereas IDID optimizes the prefilter depending on images *without* considering image coding. The proposed approach utilizes the advantage of both methods. Hereafter, we indicate the algorithm to obtain the postfilter-dependent optimal prefilters taking into account downsampling ratio and quantization error.

As in the above, we compute the distortion of the whole image as follows:

$$\bar{D} = \mathrm{MSE}(\mathbf{X}, \mathbf{X}'), \tag{16}$$

where $\mathbf{X}'$ is the reconstructed image. As shown in Fig. 3, when $D_l < \bar{D} - \alpha$, i.e., the distortion of a portion $\mathbf{R}'_l$ is much lower than that of the whole image, we change the prefilter size M to be larger than $\mathbf{P}_l$. For example, if LT-810 is employed as $\mathbf{P}$, this is replaced with LT-816. Whereas, when $D_l > \bar{D} + \beta$, we change the prefilter size to be smaller one. Let $\mathbf{M}$ be a vector which specifies the downsampling factor M of TDLTs. It is represented as follows:

$$\mathbf{M} = [M_0, M_1, \ldots, M_{m-1}]^T, \tag{17}$$

where $M_0 > M_1 > \ldots > M_{m-1}$ and m is the number of prefilters used for the optimization. In the proposed method, (15) is modified as follows:

$$\mathbf{P}_{l,M_s} = \mathrm{diag}(\mathbf{B}_u, \mathbf{P}_{M_s}, \cdots, \mathbf{P}_{M_s}, \mathbf{B}_l), \tag{18}$$

where $s = 0, \ldots, m-1$. Again, $\mathbf{P}_{M_s}$ is the prefilter of LT-NM_s. The postfilter $\mathbf{T}_{l,M_s}$ is similarly defined. Let $\mathbf{P}^{(i)}_{l,M_s}$ be the prefilter at the i-th iteration. From the above, the optimization algorithm is shown below:

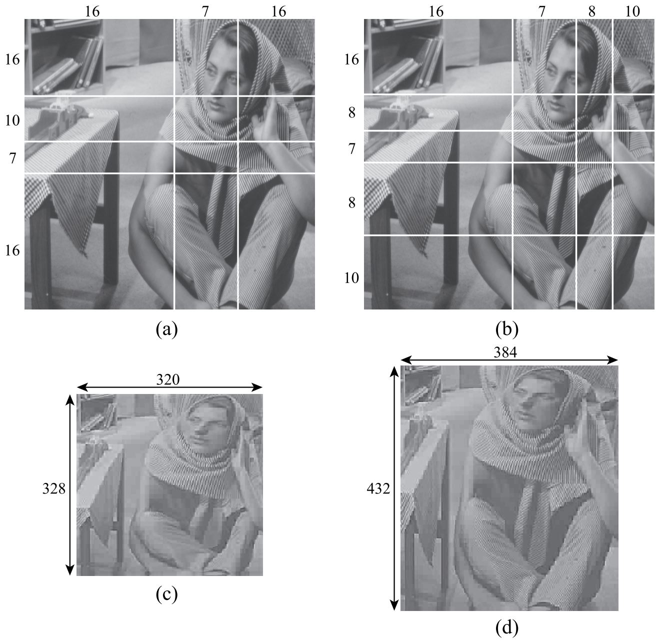

Fig. 5. Selected prefilters (specified by M) of *Barbara*. (a) in low bit rate case. (b) in high bit rate case. (c) prefiltered image of (a). (d) prefiltered image of (b).

Step 1 Initialize $\mathbf{P}_{l,M_0}^{(0)}$ and $\bar{D}$.
Step 2 Compute D_l with (13).
Step 3 If $D_l < \bar{D} - \alpha$ and $D_l > \bar{D} + \beta$, update s to be $s-1$ and $s+1$, respectively. Else, the optimized prefilter becomes $\mathbf{P}_{l,M_s}^{(i)}$.
Step 4 As IDID, derive the optimized $\mathbf{P}_{l,M_s}^{(i)}$ (denoted as $\hat{\mathbf{P}}_{l,M_s}^{(i)}$) from $\mathbf{T}_{l,M_s}^{(i)}$ and update $\mathbf{P}_{l,M_s}^{(i+1)}$ to be $\hat{\mathbf{P}}_{l,M_s}^{(i)}$ and set i to $i+1$. Then go to Step 2.

As in the above, this algorithm iteratively optimizes a set of prefilters.

6. EXPERIMENTAL RESULTS

In this section, image compression performances are compared between TDLTs and the proposed method. In the experiment for low bit rate image coding, α and β, i.e., thresholds in Section 5, are experimentally set to 20 and 50, respectively. The downsampling factors we used are $\mathbf{M} = [16, 10, 8, 7]^T$. For low bit rate and high bit rate cases, the prefilters are initialized by M_0 and M_1, respectively. Three images *Barbara* (512×512), *Birds* (768×512) and *Kodim20* (768×512) are used.

The R-D curves for test images are shown in Fig. 4. JPEG and IDID show worse performances than TDLTs. It is clear that the proposed method is better than LT-816 and LT-810. As expected, our method works well for the low to middle bit rates. Especially, it presents a significant improvement for *Barbara*.

In Fig. 5, (a) and (b) show selected prefilters for *Barbara* in low bit rate and high bit rate cases, respectively. For the low bit rate case, downsampling factors of the selected prefilters are larger than those for the high bit rate case. The prefiltered images are shown in Fig. 5 (c) and (d). Although the image sizes are globally reduced, complex textures are locally extended.

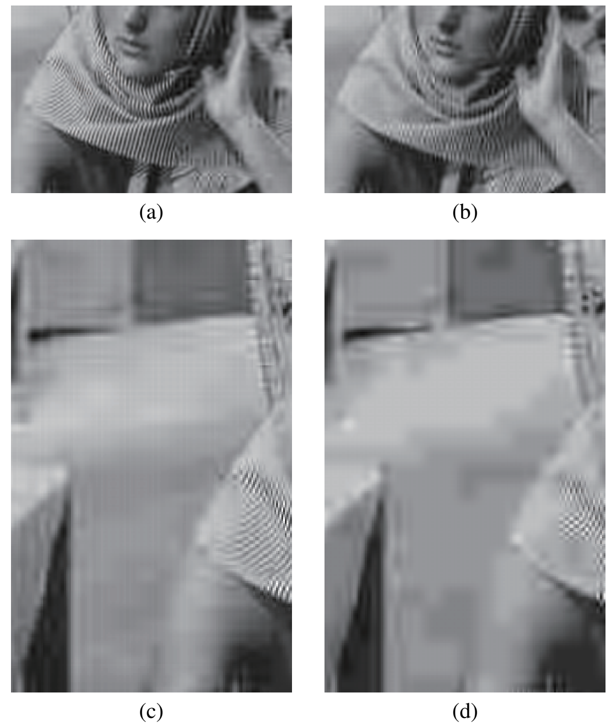

Fig. 6. Reconstructed image quality comparison of *Barbara*. All images are encoded at 0.15 bpp. (a), (c) proposed method. (b) LT-816. (d) LT-810.

Enlarged portions of the reconstructed *Barbara* are shown in Fig. 6. The proposed method shows clearer pattern in clothing than LT-816. It also shows smoother backgrounds, whereas blocking artifacts are visible in the reconstructed image by LT-810.

7. CONCLUSION

In this paper, an optimization method of prefilters for TDLTs with various downsampling factors was proposed. This approach selects the downsampling factor of the prefilter depending on the reconstruction error. Additionally, by employing the IDID algorithm, we obtain the optimized set of prefilters for a given postfilter set. Our optimal filter selection shows better R-D performances and visual image qualities than the conventional approaches.

8. REFERENCES

[1] K. R. Rao and P. Yip, *Descrete Cosine Transform: Algorithms, Advantages, Applications*, New York: Academic, 1990.

[2] W. B. Pennebaker and J. L. Mitchell, *JPEG: Still Image Compression Standard*, New York: Van Nostrand Reinhold, 1993.

[3] L. Chiariglione, "The development of an integrated audiovisual coding standard: MPEG," *Proc. IEEE*, vol. 83, pp.151-157 1995.

[4] T. D. Tran, J. Liang, and C. Tu, "Lapped transform via time-domain pre- and post-filtering," *IEEE Trans. Signal Process.*, vol. 51, no.6, pp. 1557-1571, Jun., 2003.

[5] H. S. Malvar and D. H. Staelin, "The LOT: transform coding without blocking effects," *IEEE Trans. Signal Process.*, vol. 37, no. 4, pp. 553-559, 1989.

[6] W. Lin, and L. Dong, "Adaptive downsampling to improve image compression at low bit rates", *IEEE Trans. Image Process.*, vol. 15, no. 9, pp. 2513-252, Sep., 2006.

[7] Y. Zhang, D. Zhao, J. Zhang, R. Xiong, and W. Gao, "Interpolation-dependent image downsampling", *IEEE Trans. Image Process.*, vol. 20, no. 11, pp. 3219-3296, Nov., 2011.

[8] L. Gan, C. Tu, J. Liang, T. D. Tran, and K.-K. Ma, "Undersampled boundary pre-/postfilters for low bit-rate dct-based block coders", *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 428-441, Feb., 2007.

[9] L. Gan and K.-K. Ma, "Time-domain oversampled lapped transforms: theory, structure, and application in image coding," *IEEE Trans. Signal Process.*, vol. 52, no. 10, pp. 2762-2775, Oct., 2004.

[10] Y. Tanaka, M. Hasegawa, and S. Kato, "Flexible combination of time-domain lapped transforms with various downsampling factors", *IEICE Trans. Fundamentals.*, vol.E95-A, no.11, pp. 2049-2058 Nov. 2012.