

Discrete-cosine-transform-domain downsizing with windowing operation

Il Hong Shin, Jung Ju Yoo, and Jin Woo Hung
 Electronics and Telecommunications Research Institute,
 Broadcasting Media Research Group, Radio and
 Broadcasting Group, 161 Gajeong-dong, Yuseong-gu,
 Daejeon, South Korea
 E-mail: ssi@etri.re.kr

Abstract. A simple and effective method is presented for discrete-cosine-transform (DCT)-domain downsizing. Various methods employed in DCT-domain downsizing simply reuse the frequency component of DCT, which shows a severe aliasing effect. The proposed approach extends the downsizing method for alleviating or reducing the aliasing effect with a windowing operation, which adjusts the magnitude of the DCT coefficient. Visual inspection showed satisfactory results, with no complexity overhead and performance degradation regarding the peak-signal-to noise ratio (PSNR) after upsampling. © 2007 SPIE and IS&T. [DOI: 10.1117/1.2818176]

1 Introduction

Image resizing in the discrete-cosine-transform (DCT) domain is of interest for transcoding.¹⁻⁵ It allows fast implementation omitting inverse DCT, where implicit the downsizing operation is done by truncating the high-frequency component in the DCT domain. In general, the downsizing operation always needs an anti-aliasing filter prior to the downsampling. The method contains an anti-aliasing filter implicitly by truncating the high-frequency component, because the filter bank of DCT resolves from low to high frequency. The frequency response of this method looks well shaped where a narrow transition is shown.² However, most applications of transcoding are in small-size displays such as cell phones and mobile PCs, which need quarter common intermediate format (QCIF) or common intermediate format (CIF) resolution. Although the narrow band of the DCT-domain downsizing has a good shape, the visual appearance after downsizing shows a severe aliasing effect, as shown in Fig. 1(c). A suitable anti-aliasing filter is still questionable in the image processing, but we propose a simple method to improve the visual appearance with a windowing operation, which adjusts the DCT coefficient.

2 Proposed Method

One-dimensional (1-D) twofold downsizing in the spatial domain using DCT is expressed by combination of DCT and inverse discrete cosine transform (IDCT) as follows⁶:

$$D_{\frac{N}{2} \times \frac{N}{2}}^N = T_{\frac{N}{2} \times \frac{N}{2}}^t \times T_{\frac{N}{2} \times N}^u \times B_{N \times N}, \quad (1)$$

where T denotes the 1-D DCT kernel, and $B_{N \times N}$ and $D_{(N/2) \times (N/2)}$ are the original image and downsampled image,

respectively. T^u represents the upper kernels of the DCT from row 1 to $N/2$, and the superscript t represents the transpose of the matrix. Downsizing in the DCT domain and Eq. (1) using the DCT kernel are one and the same. Therefore, we present the proposed method in the spatial domain for simplicity and easy comprehension. Let P be the weighting matrix, which is diagonal. When P is identity, the downsizing matrix is identical to the previous method. However, a severe aliasing effect is shown after the downsizing of the image, since implicit anti-aliasing is not sufficient for friendly visual appearance. We propose a windowing operation in the DCT domain for reducing the aliasing effect, where windowing is simply scaling the DCT coefficient. The new downsizing matrix is written as follows:

$$\begin{aligned} D_{N/2 \times N/2}^P &= T_{N/2 \times N/2}^t \times P_{N/2 \times N/2} \times T_{N/2 \times N}^u \times B_{N \times N} \\ &= H_{DCT,N,P} \times B_{N \times N}, \end{aligned} \quad (2)$$

where $D_{N/2 \times N/2}^P$ and $H_{DCT,N,P}$ denote the proposed downsizing operation and the combined downsizing matrix with N data point and P windowing matrix.

Joint video team (JVT)⁷ recommends a two-fold downsizing filter with twelve taps, which has a phase shift in the downsizing.⁶ The frequency response of the JVT filter is shown in Fig. 2 (method 1), which shows strong anti-aliasing but sacrifices detail preservation. However, we adopt the JVT filter for visual appearance. The proposed method finds an optimal weighting parameter P having similar frequency response to the JVT filter. We used the least-square optimization method for determining the P matrix. In other words, we searched the optimal P matrix with the frequency response of the JVT filter. The frequency response of the DCT based downsizing is written as follows⁵:

$$\begin{aligned} D(z) &= \frac{1}{N} \sum_{k=0}^{N-1} B \left[\exp\left(-\frac{j \cdot 2 \cdot \pi \cdot k}{N}\right) \cdot \sqrt{z} \right] \\ &\cdot F_k \left[\exp\left(-\frac{j \cdot 2 \cdot \pi \cdot k}{N}\right) \cdot \sqrt{z} \right], \end{aligned} \quad (3)$$

where

$$F_k(z) = \sum_{i=0}^{\frac{N}{2}-1} z^{-2i} \cdot H_{DCT,N,P,i}(z) \cdot \exp\left(-\frac{j \cdot 4 \cdot \pi \cdot k}{N}\right), \quad (3)$$

where $H_{DCT,N,P,i}(z)$ is a z -transform of the n -tap filter, which is represented by the i th row of the $H_{DCT,N,P}$ matrix. As shown in Ref. 5, since the magnitude of $F_0(z)$ is dominant in comparison with the other component, we deal with only the frequency response of $F_0(z)$ for deriving proposed filter. The problem of finding the optimal P matrix is written as follows:

$$\arg \min_P (|H_{JVT}(z)| - |F_0(z)|)^2, \quad (4)$$

where $|\cdot|$ denotes the magnitude of the z -transformed result, and $H_{JVT}(z)$ is the z -transform of the JVT's downsampling filter. However, direct calculation of the P matrix is impos-

Paper 07048LR received Mar. 29, 2007; revised manuscript received Jun. 30, 2007; accepted for publication Sep. 27, 2007; published online Dec. 3, 2007.

1017-9909/2007/16(4)/040501-3/\$25.00 © 2007 SPIE and IS&T.

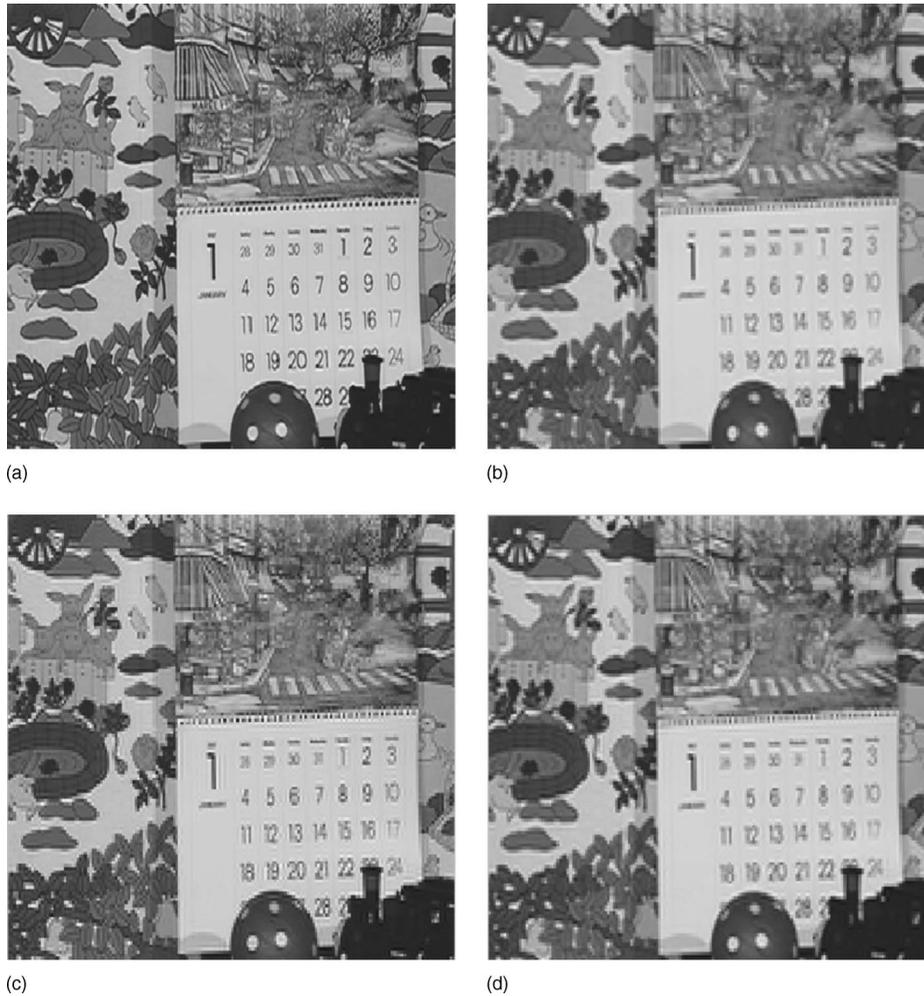


Fig. 1 Down-sized frame of the Mobile Calendar sequence: (a) original (CIF), (b) JVT filter (QCIF), (c) previous method (QCIF), and (d) proposed method (QCIF).

sible due to the nonlinear nature of the problem. We used the Levenverg Marquardt optimization method for finding the P matrix. The obtained P matrix is written as follows:

$$\text{diag}(P) = \{1, 1.0048, 1.0048, 1.0208, 1.0200, 0.8080, 0.6288, 0.0624\}, \quad (5)$$

where $\text{diag}(\cdot)$ denotes the diagonal elements of the matrix. The obtained weighting parameters decrease at the high-frequency index; hence, the index reduces the aliasing caused by the high-frequency data while lessening the detail of the image. The upsampling operation in the DCT domain is written as follows:

$$\begin{aligned} U_{N \times N}^P &= T_{N \times N/2}^{t,L} \times P_{N/2 \times N/2}^{-1} \times T_{N/2 \times N/2} \times D_{N/2 \times N/2}^P \\ &= T_{N \times N/2}^{t,L} \times (P_{N/2 \times N/2}^{-1} \times T_{N/2 \times N/2} \times T_{N/2 \times N/2}^t \\ &\quad \times P_{N/2 \times N/2}) \times T_{N/2 \times N}^u \times B_{N \times N} \\ &= U_{N \times N}, \end{aligned} \quad (6)$$

where $T^{t,L}$ and $U_{N \times N}$ represent the left kernel of the IDCT from column 1 to $N/2$ and the previous result after upsampling,² respectively. The inverse P matrix is inserted

in the DCT domain for restoring the adjusted DCT coefficient during downsampling with the proposed method. When the upsampling method in Eq. (6) is employed for image resizing after downsizing using the proposed method, the peak-signal-to noise ratio (PSNR) value is identical in comparison with the previous approach,² as shown in Eq. (6). When we applied the proposed method in the downsizing transcoder, no overhead is incurred in the computational aspect, where the P matrix is embedded in the DCT-domain down-upsizing matrix as a precalculated form such as the previous method.² Therefore, the proposed down-upsampling method in the DCT domain reduces aliasing in the downsized image, while it has no loss of PSNR after upsampling using the proposed method and no overhead in complexity during down-upsampling in the DCT domain.

3 Experimental Results

We used the two-fold downsizing matrix of Eq. (1) with $N=16$, where the DCT coefficient with N will be halved to make the downsized image. The visual appearance of “Mobile Calendar” is shown in Fig. 1. Fig. 1(b) shows good compromise with reduced aliasing and lessening details of

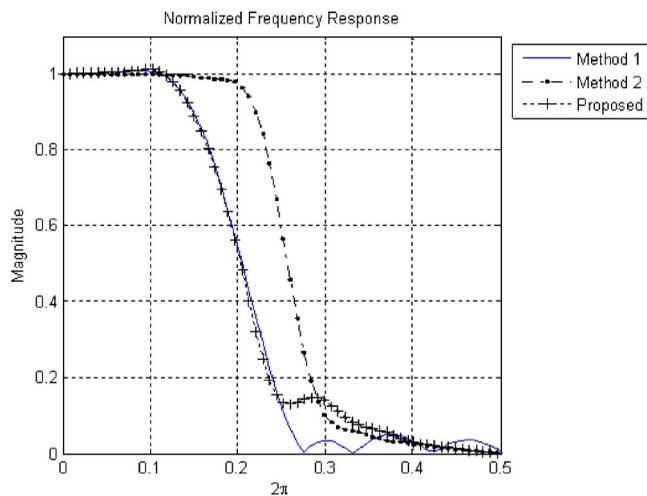


Fig. 2 Frequency response of JVT, previous, and proposed methods.

image. But Fig. 1(c) shows severe aliasing with the previous method. The visual appearance of the proposed method is similar to that of JVT.

Figure 2 shows the frequency response of the JVT filter (method 1), the previous method (method 2), and the proposed method. Method 1 shows strong anti-aliasing, whereas method 2 shows good preservation of high-frequency details. The frequency response of the proposed method shows similar shape to method 1. However, attenuation at the high-frequency band is shown, but the visual appearance shows a similar result. A large number of N may improve the frequency response with increased complexity. Moreover, adaptive determination of the weighting parameter will provide friendly visual quality. For example, when blocks containing a large high-frequency component will make the downsized block severely aliased, strong

anti-aliasing using the P matrix may make the block blurry for comfortable viewing, while low-frequency blocks perform weak anti-aliasing. However, we are searching for a method of selecting the proper P matrix through various images.

4 Conclusion

We proposed a simple and efficient windowing method for a downsizing transcoder. The experimental result shows that the proposed method improves visual quality with reducing the aliasing artifact. The windowing in the DCT domain shows a similar effect for conventional windowing of the frequency domain. The proposed method has the same computational complexity and PSNR performance after upsampling using the proposed approach in comparison with the previous DCT-domain downsizing method,² because the windowing operation in the DCT domain can be embedded in the down-upsizing operation. It can be expected that the transcoding application for downsizing will provide more friendly visual quality. Also, extension to arbitrary ratio downsizing for friendly visual quality is under way by the author.

References

1. J. Mukherjee and S. Mitra, "Image resizing in the compressed domain using subband DCT," *IEEE Trans. Circuits Syst. Video Technol.* **12**, 620–627 (2002).
2. H. W. Park, Y. S. Park, and S. K. Oh, "L/M fold image resizing in block-DCT domain using symmetric convolution," *IEEE Trans. Image Process.* **12**, 1016–1034 (2003).
3. C. L. Salazar and T. D. Tran, "On resizing images in the DCT domain," in *Proc. Int. Conf. Image Processing* (2004).
4. T. Frajka and K. Jegger, "Downsampling dependent upsampling of images," *Signal Process. Image Commun.* **19**, 257–265 (2004).
5. Y. S. Park and H. W. Park, "Design and analysis of image resizing filter in the block-DCT domain," *IEEE Trans. Circuits Syst. Video Technol.* **14**(2), 274–279 (2004).
6. I. H. Shin and H. W. Park, "Efficient down-up sampling using DCT kernel for MPEG-21 SVC," in *Proc. Int. Conf. Image Processing*, pp. 640–643 (2005).
7. Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, "Joint Scalable Video Model JSVM-5," JVT-S202, Geneva, Switzerland (2006).