

Spatio-Temporal Fuzzy Filtering for Coding Artifacts Reduction

Dung Trung Vo, Sehoon Yea, Anthony Vetro

TR2007-077 April 2008

Abstract

A 3D fuzzy-filtering scheme is proposed for reduction of compression artifacts such as blocking and ringing noises. The proposed scheme incorporates information from temporally-neighboring frames as well as from spatially-neighboring pixels by accounting for the spatio-temporal relationships in the definitions of spatial-rank orders and spread information for the fuzzy-filter. Extra information from a 3D set of pixels of the surrounding frames helps enhance the clustering characteristic of the fuzzy filter while preserving the frame edges. The proposed scheme also exploits the chroma components from neighboring frames to reconstruct the color of the current frame more faithfully. The experimental results show that both the subjective and the objective qualities of post-processed video are significantly improved.

SPIE Conf on Visual Communications and Image Processing, January 2008

This work may not be copied or reproduced in whole or in part for any commercial purpose. Permission to copy in whole or in part without payment of fee is granted for nonprofit educational and research purposes provided that all such whole or partial copies include the following: a notice that such copying is by permission of Mitsubishi Electric Research Laboratories, Inc.; an acknowledgment of the authors and individual contributions to the work; and all applicable portions of the copyright notice. Copying, reproduction, or republishing for any other purpose shall require a license with payment of fee to Mitsubishi Electric Research Laboratories, Inc. All rights reserved.

Spatio-Temporal Fuzzy Filtering for Coding Artifacts Reduction

Dũng Trung Võ^{*a}, Sehoon Yea^b, Anthony Vetro^b

^aDept. of Electrical & Computer Eng., University of California at San Diego, La Jolla, CA 92093¹;

^bMitsubishi Electric Research Laboratories, 201 Broadway, Cambridge, MA USA 02139

ABSTRACT

A 3D fuzzy-filtering scheme is proposed for reduction of compression artifacts such as blocking and ringing noises. The proposed scheme incorporates information from temporally-neighboring frames as well as from spatially-neighboring pixels by accounting for the spatio-temporal relationships in the definitions of spatial-rank orders and spread information for the fuzzy-filter. Extra information from a 3D set of pixels of the surrounding frames helps enhance the clustering characteristic of the fuzzy filter while preserving the frame edges. The proposed scheme also exploits the chroma components from neighboring frames to reconstruct the color of the current frame more faithfully. The experimental results show that both the subjective and the objective qualities of post-processed video are significantly improved.

Keywords: compression, blocking artifacts, ringing artifacts, fuzzy filter

1. INTRODUCTION

Video compression exploits inter- and intra-frame redundancies present in video to reduce the data rate while retaining as much information of the original sequence as possible. In the majority of standard-based video codecs the quality of the compressed sequence is often compromised with the presence of blocking and ringing artifacts, especially at low bit rates. Blocking artifacts occur when each macroblock in a frame is processed independently with coarse quantization of transform coefficients. Ringing artifacts, which are similar to Gibbs phenomenon [1] and visible around sharp edges, result from the loss of high frequency components of the signal during quantization.

Linear low-pass filters are proposed in pixel domain [2] and DCT domain [3][4] to remove these artifacts, but the enhanced image results in blur due to the low-pass filter. Other pixel-based methods were proposed such as projection on convex sets (POCS), constrained least squares (CLS) [5] and maximum a posteriori probability approach (MAP) [6]. However, they are often considered impractical as they require iteration processes with high computational load.

Non-linear filters, such as spatial-rank space filters in [7][8][9] and fuzzy filters in [10][11][12] reduced the ringing artifact effectively, especially for post-processing highly-compressed video sequences. Combined with a block-detection and an edge-map generation methods, the fuzzy filters were shown to be very effective in removing both the blocking and ringing noises while preserving the natural edges and details of the video. As will be briefly reviewed in the next section, the fuzzy filter possesses a very desirable property of ‘clustering’ which is ideal for removing noisy artifacts while retaining prominent edges of the given frames.

Most existing deblocking and deringing algorithms focus on using information from within the current frame. However, this often leads to compromised visual quality, especially in moving image sequences since temporal consistency is not maintained. In this paper, experimental results are provided to confirm that both the subjective and the objective qualities of post-processed video could be substantially improved by incorporating information from neighboring frames. A 3D fuzzy-filter is constructed by accounting for the spatio-temporal relationships in the definitions of spatial-rank orders and spread information. Extra information from a 3D set of pixels of the surrounding frames helps enhance the clustering characteristic of the fuzzy filter while preserving the frame edges. The proposed scheme also exploits the chroma components from neighboring frames to reconstruct the color of the current frame more faithfully.

The rest of this paper is organized as follows. A brief review of the fuzzy-filter is given in the next section. Section 3 describes the proposed 3D filtering method, followed by block classification in section 4 to adapt the filter parameter to the activity of each block. Experimental results are shown in section 5 and conclusions will be drawn in section 6.

¹ This work was performed while Dũng Trung Võ was with Mitsubishi Electric Research Laboratories.

2. FUZZY FILTER

The fuzzy filter [10] exploits the spatial correlation, rank order and diversity of the signal. It overcomes the limitation of the spatial-rank (SR) space filters such as weighted median [7], rank condition rank selection (RCRS) [8] and lower-upper-middle (LUM) [9] filters by adding the spread information to get better clustering performance: clustering the similarly valued samples to their mean value and keeping isolated values unchanged. This characteristic is very helpful in preserving the real edge of the image while reducing the ringing artifact in the flat area near these edges.

Assume that $x_i = [x_1, x_2, \dots, x_K]$ is the input vector in the spatial domain and its order statistic vector can be denoted as $x_L = [x_{(1)}, x_{(2)}, \dots, x_{(K)}]$ with $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(K)}$.

Then the linear transformations between x_i and x_L will be given as

$$x_i = x_L R^T \text{ and } x_L = x_i R, \quad (1)$$

where R is the crisp (binary) spatial-rank (SR) matrix defined by

$$R = \begin{bmatrix} R_{1,(1)} & \cdots & R_{1,(N)} \\ \vdots & \ddots & \vdots \\ R_{N,(1)} & \cdots & R_{N,(N)} \end{bmatrix} \quad (2)$$

with

$$R_{i,(j)} = \mu_C(x_i, x_{(j)}) = \begin{cases} 1, & \text{for } x_i \rightarrow x_{(j)} \\ 0, & \text{otherwise} \end{cases} . \quad (3)$$

In the fuzzy version of the above relationship, the binary function $\mu_C(x_i, x_{(j)})$ in (3) is replaced by a real-valued membership function $\mu_F(x_i, x_{(j)})$ to add the spread information into the SR relation. Here, $\mu_F(a, b)$ has to satisfy the following conditions:

$$1. \quad \lim_{|a-b| \rightarrow 0} \mu_F(a, b) = 1 \quad (4)$$

$$2. \quad \lim_{|a-b| \rightarrow \infty} \mu_F(a, b) = 0 \quad (5)$$

$$3. \quad |a_1 - b_1| \leq |a_2 - b_2| \Rightarrow \mu_F(a_1, b_1) \leq \mu_F(a_2, b_2) . \quad (6)$$

Widely used membership functions include uniform, Gaussian and triangular functions. The Gaussian membership function is defined as

$$\mu_F(a, b) = e^{-\frac{(a-b)^2}{2\sigma^2}}, \quad (7)$$

where σ is the spread parameter controlling the strength of the fuzzy filter. High σ values will average the signal while small σ values will keep the signal isolated.

The output of the fuzzy identity filter is just the fuzzy counterpart of the center sample in the observation (filtering) window, which can be obtained using the following simplified formula:

$$y = \tilde{x}_c = \frac{\sum_{j=1}^N x_{(j)} \mu_F(x_c, x_{(j)})}{\sum_{j=1}^N \mu_F(x_c, x_{(j)})}, \quad (8)$$

where x_c and \tilde{x}_c are crisp and fuzzy center samples, respectively.

For application to multi-dimension signals such as an image, the 2D signal has to be converted to a vector and the fuzzy filter in (8) is used to reduce the artifacts.

3. PROPOSED ALGORITHM

The diagram of the proposed algorithm is shown in Fig. 1. First, the current frame is subject to a deblocking process using the 1D fuzzy filter along the vertical and the horizontal block boundaries [11][12]. Then, the 3D fuzzy filter is applied to reduce the ringing artifacts of the current frame using information from neighboring frames as well as the current one. Note that the chroma components of frames are up-sampled to be the same size as that of the luma component before the 3D-filtering takes place for convenience in processing.

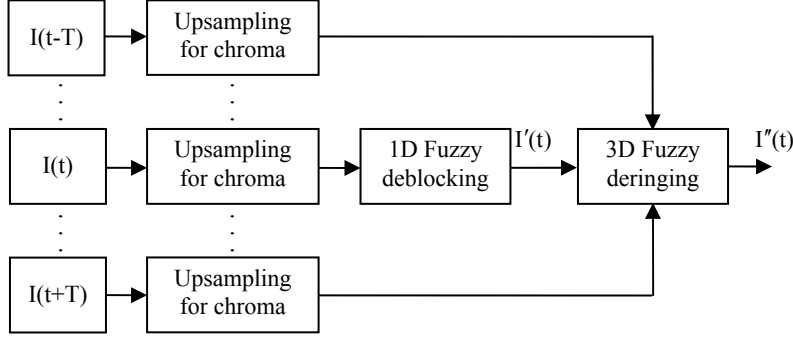


Fig. 1. Block diagram of the proposed method

3.1 Deblocking

A strong smoothing filter with Gaussian membership function will be applied only to pixels at the border of two neighboring blocks for deblocking. Let $I(t, m, n)$ denote the intensity value of a pixel located at (t, m, n) in the spatio-temporal coordinates. A pixel is classified either as an ‘edge pixel’ or a ‘non-edge pixel’ based upon the standard deviation calculated using 3x3 neighboring pixels as follows:

$$\sigma(I(t, m, n)) = \frac{1}{9} \sum_{k=m-1}^{m+1} \sum_{l=n-1}^{n+1} \left(I(t, k, l) - \frac{1}{9} \sum_{p=m-1}^{m+1} \sum_{q=n-1}^{n+1} I(t, p, q) \right)^2 \quad (9)$$

If $\sigma(I(t, m, n)) < \text{Threshold}$, a pixel will be considered a non-edge pixel. Otherwise, it will be considered an edge pixel.

In order to avoid destroying the real edges of image at the border, only the non-edge pixels at the borders will be filtered. For a non-edge pixel $I(t, m, n)$ around the block boundary, an N-tap fuzzy filter will be applied either along the horizontal or the vertical direction to form the filtered pixel $I'(t, m, n)$. For example, a 5-tap horizontal fuzzy filter will generate $I'(t, m, n)$ as follows:

$$I'(t, m, n) = \frac{\sum_{j=1}^5 I_{(j)} \mu_F(I(t, m, n), I_{(j)})}{\sum_{j=1}^5 \mu_F(I(t, m, n), I_{(j)})} = \frac{\sum_{j=1}^5 I_j \mu_F(I(t, m, n), I_j)}{\sum_{j=1}^5 \mu_F(I(t, m, n), I_j)}, \quad (10)$$

where the neighboring pixels form an input vector

$$\vec{I} = [I_1, I_2, I_3, I_4, I_5] = [I(t, m, n-2), I(t, m, n-1), I(t, m, n), I(t, m, n+1), I(t, m, n+2)] \quad (11)$$

and the equivalent rank vector is given as

$$\vec{I}_r = [I_{(1)}, I_{(2)}, I_{(3)}, I_{(4)}, I_{(5)}], \text{ where } I_{(1)} \leq I_{(2)} \leq I_{(3)} \leq I_{(4)} \leq I_{(5)}. \quad (12)$$

3.2 Deringing

In order to remove ringing artifacts, a spatio-temporal fuzzy filter will be applied on the pixels obtained as the outputs of the deblocking process described above. More specifically, a set of $(T^- + 1 + T^+)$ frames are used for the filtering, i.e., a set of T^- previous, the current and T^+ future frames. Note that T^- and T^+ could be different. For the current pixel $I'(t, m, n)$ of the current frame, the fuzzy filter is applied to a set of $(T^- + 1 + T^+) \times (2M + 1) \times (2N + 1)$ pixels spatially and temporally adjacent to it. In other words, these adjacent pixels form an input vector

$$\bar{I}' = [I'_{(0)}, I'_{(1)}, \dots, I'_{((T^- + T^+) \times (2M + 1) \times (2N + 1) - 1)}] = [I'(t - T^-, m - M, n - N), \dots, I'(t + T^+, m + M, n + N)] \quad (13)$$

and the equivalent rank vector is

$$\bar{I}'_r = [I'_{(0)}, I'_{(1)}, \dots, I'_{((T^- + T^+) \times (2M + 1) \times (2N + 1) - 1)}], \text{ where } I'_{(0)} \leq I'_{(1)} \leq \dots \leq I'_{((T^- + T^+) \times (2M + 1) \times (2N + 1) - 1)}. \quad (14)$$

The filtered pixel will then be calculated by

$$I''(t, m, n) = \frac{\sum_{l=0}^{(T^- + T^+) \times (2M + 1) \times (2N + 1) - 1} \{I'_{(l)} \times \mu_F(I'(t, m, n), I'_{(l)})\}}{\sum_{l=0}^{(T^- + T^+) \times (2M + 1) \times (2N + 1) - 1} \mu_F(I'(t, m, n), I'_{(l)})} \\ = \frac{\sum_{\Delta t = -T^-}^{T^+} \sum_{\Delta m = -M}^M \sum_{\Delta n = -N}^N I'(t + \Delta t, m + \Delta m, n + \Delta n) \mu_F(I'(t, m, n), I'(t + \Delta t, m + \Delta m, n + \Delta n))}{\sum_{\Delta t = -T^-}^{T^+} \sum_{\Delta m = -M}^M \sum_{\Delta n = -N}^N \mu_F(I'(t, m, n), I'(t + \Delta t, m + \Delta m, n + \Delta n))}. \quad (15)$$

4. BLOCK CLASSIFICATION

For deringing, the spread parameter σ in (7) should be adaptive to the ‘activity’ levels of the area to be filtered. In the proposed scheme, the ‘activity’ level is measured by the maximum of the standard deviations as calculated by eqn. (9) of the pixels in a given block and is used to differentiate among different types of areas such as edges, textures, and smooth areas. By judiciously adjusting the spread parameter, ringing artifacts could be effectively removed while avoiding overblurring of the textures. Fig.2 is the flow chart of classification adopted in the proposed scheme, where the MaxSTD is the aforementioned ‘activity’ level. According to the activity level, the blocks in a frame are classified into 5 categories and filters with varying spread parameters are applied as described below.

- **Strong-edge block:** each pixel is filtered using the strong smooth fuzzy filter with a spread parameter σ of 20.
- **Weak-edge block:** each pixel is filtered by fuzzy filter with a spread parameter σ of 15.
- **Strong-texture block:** each pixel is filtered using the smooth fuzzy filter with a spread parameter σ of 10.
- **Weak-texture block:** each pixel is filtered using the smooth fuzzy filter with a spread parameter σ of 10.
- **Smooth block:** a weak smooth filter with a spread parameter σ of 5 will be applied.

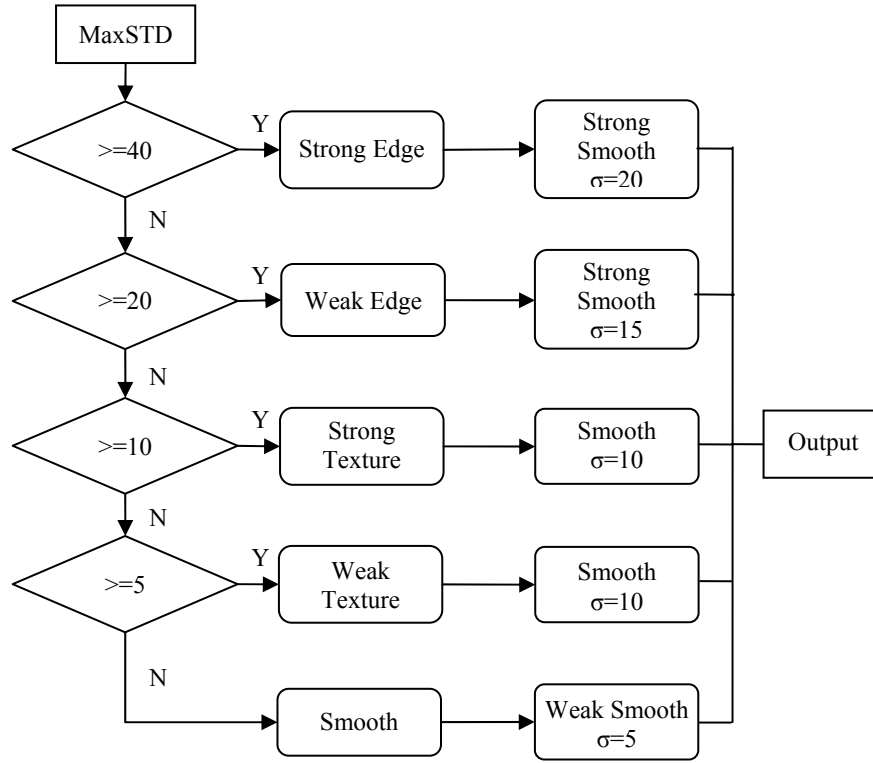


Fig. 2. Flow chart of block classification

5. SIMULATION RESULTS

In the following, the test conditions for our evaluation of the proposed filtering technique are first described. Experimental results are then provided and compared both objectively and subjectively. The impact of the temporal dependency on the filtering results is also studied.

5.1 Test Conditions

In the simulation, a Gaussian membership function as given in (7) was chosen and the spread parameter was set to $\sigma = 30$ for deblocking. For the spatio-temporal deringing filter, the spread parameter was set adaptively according to the classification in Fig.2. The window size was $5 \times 5 \times 5$ corresponding to $T^- = T^+ = 2$, $M = 2$, and $N = 2$ for the temporal, vertical and horizontal directions, respectively. For comparison, the methods using a low-pass filter for DCT coefficients of the shifted blocks proposed by Chen et al. [3] and by Liu and Bovik [4] as well as the 2D fuzzy filter previously proposed by the authors in [11][12] were simulated. The video sequences were compressed using a Motion JPEG codec.

5.2 Subjective and Objective Comparison of the Results

Fig. 3 compares the enhanced images obtained by Chen's method (Fig. 3c), Liu's method (Fig. 3d), and the 2D and 3D fuzzy filters (Figs. 3e and 3f, respectively). Although both Chen's and Liu's methods effectively remove blocking artifacts, neither of these techniques are capable of removing ringing artifacts and the resulting images tend to look blurry. On the other hand, the processed images using the fuzzy filtering techniques as shown in Figs. 3e and 3f reduce both ringing and blocking artifacts effectively while avoiding overblurring the images. Note that the enhanced image obtained by the proposed 3D fuzzy filter (Fig. 3f) shows significantly reduced ringing artifact and better color quality than the 2D counterpart (Fig. 3e). The drastic improvement in visual quality is more readily observable in the enlarged portion of the picture as shown in Fig. 4. As mentioned earlier, the deringing filter was applied only to the luminance component in the 2D fuzzy case as there were not enough chroma samples to gain any benefits from the clustering

property of the fuzzy filter for deringing. However, with the current 3D extension, more chroma samples are available from the neighboring frames and the use of deringing filter for chroma components helps improve the color quality significantly.

Next, Figs. 5a and 5b compare the PSNR's and the gains in PSNR of the tested methods with respect to the PSNR of the compressed images of the Mobile sequence, respectively. Note that all the PSNR values presented hereafter in this paper are calculated with respect to the uncompressed images. The plots clearly show the proposed method achieves quite consistent PSNR gains of about 0.67dB, 0.91dB, 0.72dB, and 0.40dB on average relative to the compressed images, those by Chen's method, Liu's method, and the 2D fuzzy filter, respectively. It should also be note that Chen's and Liu's methods incur a slight degradation in terms of PSNR as compared with non-post-filtered images even though they removed blocking-artifacts. Perhaps more importantly, the effectiveness of the proposed scheme was much more noticeable when played back as a sequence, as the 3D result produced a smoother video with significantly reduced mosquito artifacts.

Table 1 shows the PSNR results of similar simulations for Mobile, News, Foreman, Mother, Paris, and Silent sequences. These results show the proposed scheme consistently outperforms other methods throughout all the test sequences and demonstrates the effectiveness of the proposed scheme. The average gains in PSNR over that of the compressed images of the methods were 0.24 dB, 0.23 dB, 0.53 dB, and 0.78 dB for Chen's method, Liu's method, 2D fuzzy, and 3D fuzzy, respectively.

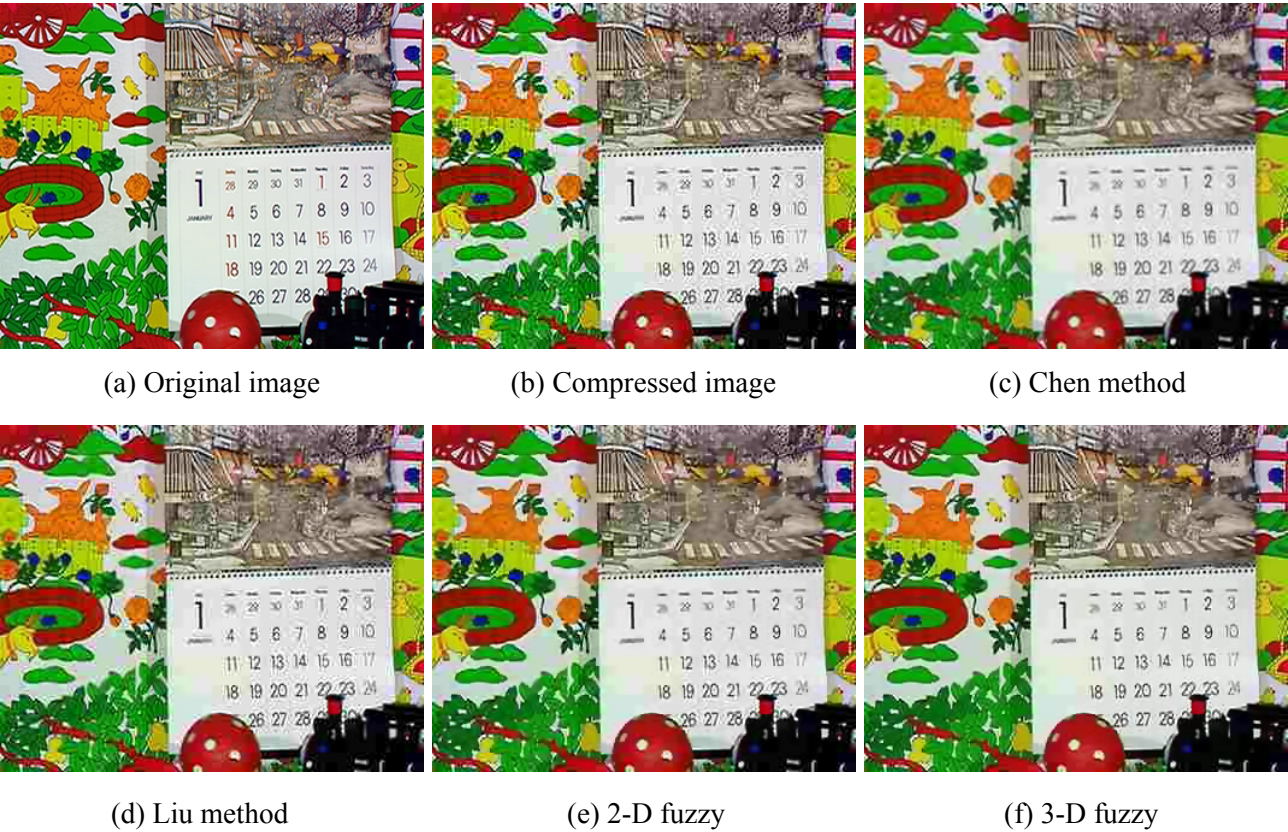


Fig. 3. Visual comparison of tested methods on Mobile sequence

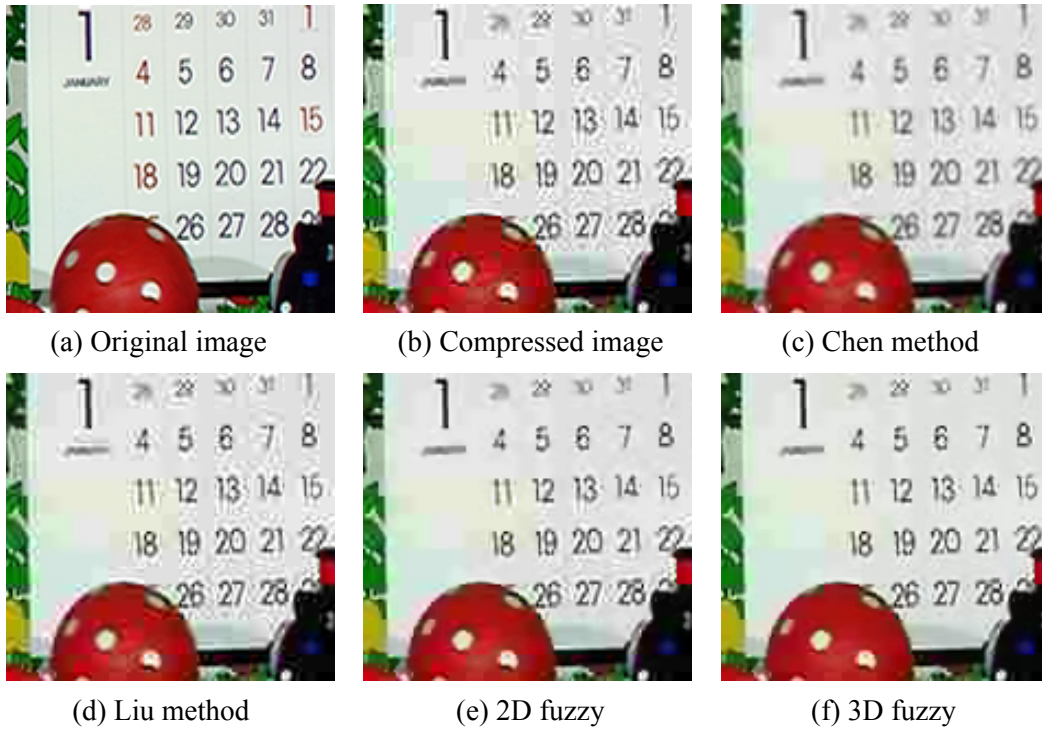


Fig. 4. Zoom-in result for color artifacts comparison

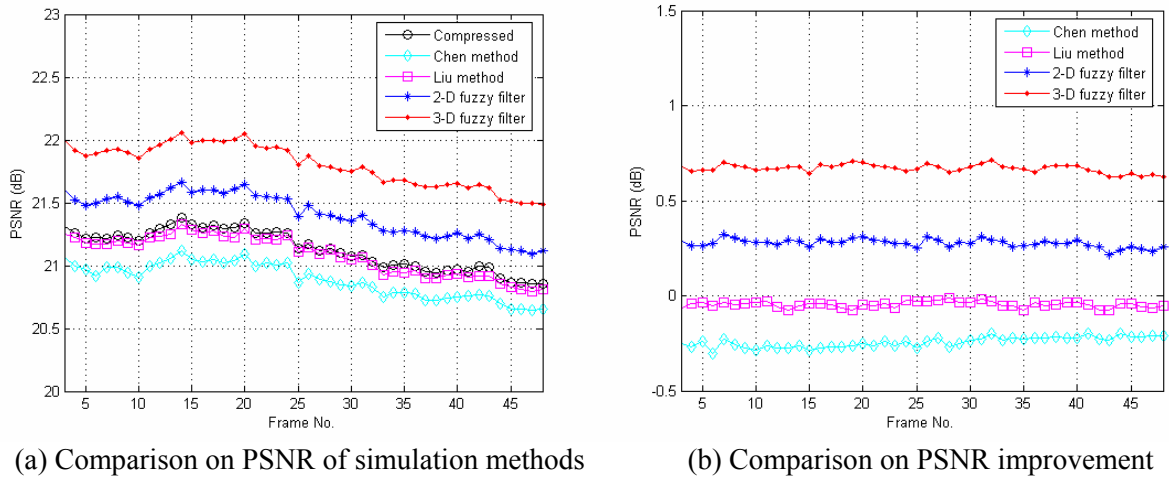


Fig. 5. PSNR for Mobile sequence

Table 1. PSNR Comparison of the methods with various sequences [dB]

Sequences	Compressed	Chen	Liu	2D fuzzy	3D fuzzy
Mobile	21.21	20.96	21.13	21.47	21.89
News	27.48	27.58	27.55	27.94	28.31
Foreman	28.06	28.46	28.41	28.78	28.92
Mother	31.02	31.83	31.62	31.77	31.99
Paris	23.38	23.25	23.31	23.80	24.04
Silent	27.84	28.37	28.33	28.38	28.51
Average gains	---	0.24	0.23	0.53	0.78

5.3 Choice of Neighboring Frames

As a practical implementation issue, the effect of using only the previous frames or different number of neighboring frames was investigated. The cases considered included: 2 previous and 2 future (Fig. 6a), 4 previous (Fig. 6b), and 6 previous frames (Fig. 6c). With the same number of temporal frames, using both previous and future frames (Fig. 6a) yielded higher quality than using only the previous frames (Fig. 6b). This is most likely due to the fact that the temporal correlation of the neighboring frames with the current one is higher in the former case. Next, it is observed that using more frames led to a slightly decreased PSNR as well as degraded visual quality (Fig. 6c). This is the result of using frames in the filtering process that have a large temporal distance with respect to the current frame that tend to be less correlated, and as a result, the effectiveness of the filter is decreased. These results reveal that additional factors need to be considered when applying temporal filters, in addition to the usual implementation issues such as memory and processing delay associated with the use of more frames.

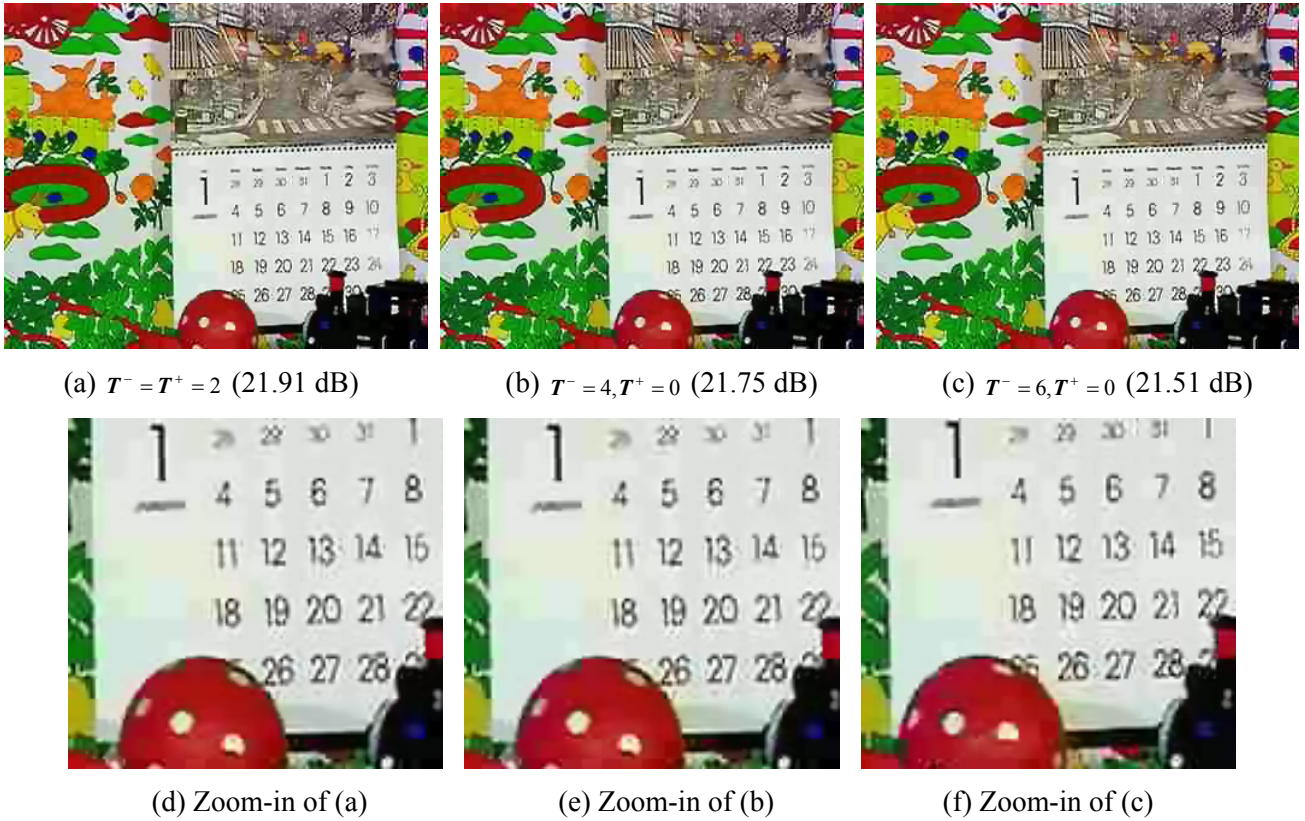


Fig. 6. Effect of different neighboring frame choices

6. CONCLUSION

A 3D filtering scheme is proposed for reduction of coding artifacts such as blocking and ringing noises. It is based upon a fuzzy-filtering algorithm incorporating spatio-temporal information. It is demonstrated that it achieves significantly improved qualities of post-processed video both objectively and subjectively. Future work will include the development of a more effective block classification algorithm as well as methods of adjusting the spread parameter of the fuzzy-filter by taking the spatio-temporal relationships of the pixels into account.

REFERENCES

- ¹ A.J. Jerri, "The Gibbs Phenomenon in Fourier Analysis, Splines and Wavelet Approximations", Kluwer Academic Publisher, Dordrecht, 1998.
- ² T. Jarske, P. Haavisto, I. Defee, Post Filtering Methods for Reducing Blocking Effects from Coded Images, IEEE Trans. Consumer Electronics, vol. 40, no. 3, pp. 521-526, Aug. 1994.
- ³ T. Chen, H.R. Wu, and B. Qiu, "Adaptive Postfiltering of Transform Coefficients for the Reduction of Blocking Artifacts," IEEE Trans. Circuits Syst. Video Technol., vol. 11, no. 5, pp. 594-602, May 2001.
- ⁴ S. Liu, and A.C. Bovik, "Efficient DCT-Domain Blind Measurement and Reduction of Blocking Artifacts," IEEE Trans. Circuits Syst. Video Technol., vol. 12, no. 12, pp. 1139-1149, Dec. 2002.
- ⁵ Y. Yang, N.P. Galatsanos, and A.K. Katsaggelos, "Regularized Reconstruction to Reduce Blocking Artifacts of Block Discrete Cosine Transform Compressed Images," IEEE Trans. Circuits Syst. Video Technol., vol. 3, no. 6, pp. 421-432, Dec. 1993.
- ⁶ R.L. Stevenson, "Reduction of Coding Artifacts in Transform Image Coding," Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP), vol. 5, pp. 401-404, Apr. 1993.
- ⁷ B. J. Justusson, "Median filtering: Statistical properties," in Two Dimensional Digital Signal Processing II, T. S. Huang, Ed. Berlin, Germany: Springer-Verlag, 1981.
- ⁸ K. E. Barner and R. C. Hardie, "Spatial-rank order selection filter," in Nonlinear Signal Processing, S. K. Mitra and G. Sicuranza, Eds. New York: Academic, 1999, pp. 69-110.
- ⁹ R. C. Hardie and C. G. Boncelet, "LUM filters: A class of rank-order-based filters for smoothing and sharpening," IEEE Trans. Signal Process., vol. 41, no. 3, pp. 1061-1076, Mar. 1993.
- ¹⁰ Y. Nie, K.E. Barner, "Fuzzy transformation and its applications Image Processing", Proc. of IEEE Int. Conf. Image Proc. (ICIP), 2003.
- ¹¹ H.S. Kong, A. Vetro, S. Huifang, "Edge Map Guided Adaptive Post-Filter for Blocking and Ringing Artifacts Removal", Proc. of IEEE Int. Sym. Circuits and Syst. (ISCAS), vol. 3, pp. 929-932, May 2004.
- ¹² H.S. Kong, Y. Nie, A. Vetro, H. Sun and K. Barner, "Adaptive Fuzzy Post-Filtering for Highly Compressed Video", Proc. of IEEE Int. Conf. Image Proc. (ICIP), pp. 1802-1806, 2004.