MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

Adaptive Fuzzy Filtering for Artifact Reduction in Compressed Images and Videos

Dung T. Vo, Truong Q. Nguyen, Sehoon Yea, Anthony Vetro

TR2009-022 June 2009

Abstract

A fuzzy filter adaptive to both sample's activity and the relative position between samples is proposed to reduce the artifacts in compressed multidimensional signals. For JPEG images, the fuzzy spatial filter is based on the directional characteristics of ringing artifacts along the strong edges. For compressed video sequences, the motion compensated spatiotemporal filter (MCSTF)is applied to intraframe and interframe pixels to deal with both spatial and temporal artifacts. A new metric which considers the tracking characteristic of human eyes is proposed to evaluate the flickering artifacts. Simulations on compressed images and videos show improvement in artifact reduction of the proposed adaptive fuzzy filter over other conventional spatial or temporal filtering approaches.

IEEE Transactions on Image Processing

Adaptive Fuzzy Filtering for Artifact Reduction in Compressed Images and Videos

Dũng T. Võ, *Student Member, IEEE*, Truong Q. Nguyen, *Fellow, IEEE*, Sehoon Yea, *Member, IEEE*, and Anthony Vetro, *Senior Member, IEEE*

***Abstract*—A fuzzy filter adaptive to both sample's activity and the relative position between samples is proposed to reduce the artifacts in compressed multidimensional signals. For JPEG images, the fuzzy spatial filter is based on the directional characteristics of ringing artifacts along the strong edges. For compressed video sequences, the motion compensated spatiotemporal filter (MCSTF) is applied to intraframe and interframe pixels to deal with both spatial and temporal artifacts. A new metric which considers the tracking characteristic of human eyes is proposed to evaluate the flickering artifacts. Simulations on compressed images and videos show improvement in artifact reduction of the proposed adaptive fuzzy filter over other conventional spatial or temporal filtering approaches.**

***Index Terms*—Artifact reduction, flickering metric, fuzzy filter, motion compensated spatio-temporal filter.**

I. Introduction

BLOCK-BASED compressed signals suffer from blocking, ringing, mosquito, and flickering artifacts, especially at low-bit-rate coding. Separately compressing each block breaks the correlation between pixels at the border of neighboring blocks and causes blocking artifacts. Ringing artifacts occur due to the loss of high frequencies when quantizing the DCT coefficients with a coarse quantization step. Ringing artifacts are similar to the Gibbs phenomenon [1] and are most prevalent along the strong edges. On the order hand, mosquito artifacts come from ringing artifacts of many single compressed frames when displayed in a sequence. For intercoded frames, mosquito artifacts become more annoying for blocks on the boundary of moving object and background which have significant interframe prediction errors in the residual signal [2]. Flickering artifacts [3], [4] happen due to the inconsistency in quality over frames at the same spatial position. This inconsistency is from the temporal distortion over compressed frames caused by quantizing the residual signal. These flickering artifacts, which are perceived more in the flat areas, also come from different quantization levels for rate-distortion optimization.

Many filter-based denoising methods have been proposed to reduce these artifacts, most of which are frame-based enhancement. For blocking artifact reduction, a linear low-pass filter was used in [5] to remove the high frequencies caused by blocky edges at borders, but excessive blur was introduced since the high frequencies components of the image were also removed. In [6]–[8], low-pass filters were applied to the DCT coefficients of shifted blocks. In particular, the adaptive linear filters in [7] and [8] were proposed to overcome the problem of over-blurring the images, but these methods require high computational complexity. In [9], a projections onto convex set-based method was proposed with multiframe constraint sets to reduce the blocking artifacts. This method required to extract the motion between frames and quantization information from the video bit-stream.

To reduce ringing artifacts, the methods in [10] and [11] utilized the linear or nonlinear isotropic filters to the ringing areas. As an encoder-based approach, [12] proposed a noise shaping algorithm to find the optimal DCT coefficients which adapts to the noise variances in different areas. All of these methods can only reduce ringing artifacts in each frame. To deal with the temporal characteristic of mosquito artifacts, [13] applied the spatiotemporal median filter in transform domain for surrounding 8×8 blocks. The improvement in this case is limited by the small correlation between DCT coefficients of the spatial neighboring 8×8 blocks as well as the lack of motion compensation in the scheme.

For flickering artifact removal, most of the current methods focused on reducing flickering artifacts in all intraframe coding. In [3], the quantization error is considered to obtain the optimal intra prediction mode and to help reducing the flickering artifact. Also for intraframe coding, [4] included the flickering artifact term in the cost function to find the optimal prediction and block-size mode. A similar scheme is implemented in [14] for flickering reduction in Motion JPEG 2000. Note that all of these approaches are encoder-based.

In order to reduce the temporal artifacts such as mosquito and flickering artifacts more efficiently, not only the spatial correlation among pixels but also the temporal one need to be incorporated. Fig. 1 shows the correlation between the 5th frame of compressed Mobile sequence and its surrounding frames. Compared to the auto-correlation of the current frame, the cross-correlation in the plot between the center frame and its surrounding frames is still rather large when the frame distance is small.

Manuscript received June 12, 2008; revised January 13, 2009. First published May 02, 2009; current version published May 13, 2009. Part of this work was done while D. T. Võ was with Mitsubishi Electronic Research Laboratories. This work was supported in part by Texas Instruments, Inc. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Zhigang (Zeke) Fan.

D. T. Võ and T. Q. Nguyen are with the Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA 92037-0407 USA (e-mail: d3vo@ucsd.edu; nguyent@ece.ucsd.edu).

S. Yea and A. Vetro are with Mitsubishi Electric Research Laboratories, Cambridge, MA 02139 USA (e-mail: yea@merl.com; avetro@merl.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2009.2017341

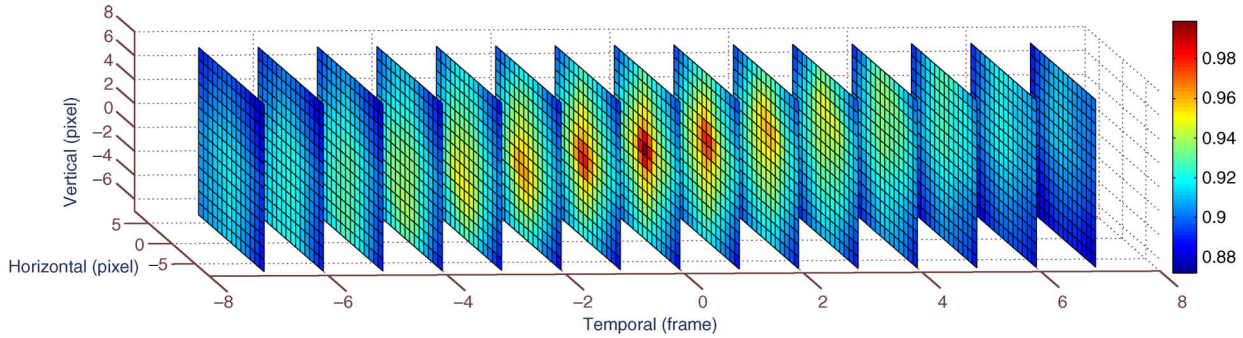

Fig. 1. Correlation between the current frame of compressed mobile sequence and its surrounding frames.

Using extra information from temporally neighboring samples, such as pixels of surrounding frames in video sequences, can further enhance the quality of compressed video sequences.

In this paper, spatial neighboring pixels are used to deal with blocking and ringing artifacts while temporal neighboring pixels are utilized to remove mosquito and flickering artifacts. To avoid the blurring effect of linear filters, a fuzzy filter is implemented. Fuzzy filter is a specific case of bilateral filters [15], [16]. Fuzzy filters help denoising the artifacts while retaining the sharpness of real edges. One drawback of fuzzy filters for multidimensional signals is that the signal is converted to a vector that ignores the relative position of the pixels. In image and video compression, the artifacts such as blocking or ringing artifacts are spatially directional and flickering artifacts are temporally directional. In this work, a novel fuzzy filter is proposed to adapt to the pixel's activity and directions between the pixel of interest and its surrounding pixels. This adaptive fuzzy filter is considered for both cases of compressed images and video sequences. To assess the filter performance in reducing the flickering artifact, a novel flickering metric based on the metric in [17] is proposed with the extension of flickering consideration for motion areas. The spatial adaptation and directional adaptation make the proposed adaptive fuzzy filter different from the conventional bilateral filters, which adapt to the distance between pixels. Another adaptation of bilateral filters in the offset and the width of the range filter was discussed in [18]–[20]. These locally adaptive methods require complicated training-based approach and are only used for image enhancement.

This paper is organized as follows. Section II provides background on fuzzy filtering. Section III introduces the concept of directional fuzzy filter for artifact removal in JPEG images and describes an edge-based scheme that realizes the directional fuzzy filtering concept. Adaptive fuzzy motion compensated spatiotemporal filter (MCSTF) for quality enhancement in compressed video sequences is discussed in Section IV. The new flickering metric is proposed in Section V. Simulation results comparing the proposed filter to existing approaches are presented in Section VI. Finally, Section VII gives concluding remarks.

II. Fuzzy Filter

Fuzzy filters, such as those described in [11] and [21], improve on median filters [22] or rank condition rank selection filters [23] by replacing the binary spatial-rank relation by a real-valued relation. The conventional way to define the fuzzy filters is by generalizing the binary spatial-rank relation. In this paper, the fuzzy filter is introduced from the artifact reduction aspect. Assume that a filter h is applied to a set Ω of neighboring samples $x[m+m', n+n']$ around the input $x[m,n]$ to form the output

$$y[m,n] = \sum_{[m',n']\in\Omega} h\left(x[m+m', n+n'], x[m,n]\right) \times x[m+m', n+n'] \quad (1)$$

and its unbiased form with normalization

$$y[m,n] = \frac{\sum_{[m',n']\in\Omega} h(x[m+m', n+n'], x[m,n])\, x[m+m', n+n']}{\sum_{[m',n']\in\Omega} h\left(x[m+m', n+n'], x[m,n]\right)}. \quad (2)$$

In (1), $h(x[m+m', n+n'], x[m,n])$ controls the contribution of the input $x[m+m', n+n']$ to the output. For a linear filter, h is fixed and input-independent. In the case of a nonlinear filter, h is a function of the input, such as for median filter

$$h\left(x[m+m', n+n'], x[m,n]\right) = \begin{cases} 1, & \text{if } \operatorname{rank}\left(x[m+m', n+n']\right) = \operatorname{round}\left(\frac{\operatorname{size}(\Omega)}{2}\right) \\ 0, & \text{otherwise} \end{cases}$$

where $round(u)$ is the nearest integer of u.

Due to the input independence of the filter coefficients, a low-pass filter which is designed to perform effectively in the flat areas may introduce blurring artifacts in detail areas. In artifact reduction, especially for low bit-rate compression, it is desirable to preserve the details while removing the artifacts. This can be achieved by imposing the constraint such that if $x[m+m', n+n']$ is far from $x[m,n]$, its contribution to the output is small. In that case, the filter coefficients $h[k,l]$ must follow the constraints

$$\lim_{|x[m+m',n+n']-x[m,n]|\to 0} h(x[m+m', n+n'], x[m,n]) = 1 \quad (3)$$

$$\lim_{|x[m+m',n+n']-x[m,n]|\to \infty} h(x[m+m', n+n'], x[m,n]) = 0 \quad (4)$$

and

$$h\left(x\left[m+m_1', n+n_1'\right], x[m,n]\right) \geq h\left(x\left[m+m_2', n+n_2'\right], x[m,n]\right) \text{ if } \left|x\left[m+m_1', n+n_1'\right] - x[m,n]\right| \leq \left|x\left[m+m_2', n+n_2'\right] - x[m,n]\right|. \quad (5)$$

The function $h(x[m+m', n+n'], x[m,n])$ is referred to as the membership function and there are many functions which fulfill these requirements. For a Gaussian membership function

$$h\left(x[m+m', n+n'], x[m,n]\right) = exp\left(-\frac{\left(x[m+m', n+n'] - x[m,n]\right)^2}{2\sigma^2}\right) \quad (6)$$

where σ represents the spread parameter of the input and controls the strength of the fuzzy filter. Note that the contribution of the input $x[m,n]$ to the output is always highest compared to the contribution of other samples

$$h\left(x[m,n], x[m,n]\right) = 1 \geq h\left(x[m+m', n+n'], x[m,n]\right) \quad \forall k. \quad (7)$$

For the same $|x[m+m', n+n'] - x[m,n]|$, the higher the σ value, the higher the contribution of $x[m+m', n+n']$ relatively compared to the contribution of $x[m,n]$ to the output. This implies that $x[m,n]$ will be more averaged to $x[m+m', n+n']$. Smaller σ values will keep the signal $x[m,n]$ more isolated from its neighboring samples. This spread parameter should be adaptive to different areas which have different activity levels such as smooth or detail areas. For multidimensional signals, the conventional fuzzy filter assigns a fixed spread parameter for every surrounding sample and ignores the relative position between them. In image and video compression, artifacts such as blocking, ringing or flickering artifacts are directional, and, thus, the fuzzy filter should consider the directions between $x[n]$ and its surrounding samples $x[m+m', n+n']$. This can be achieved by an adaptive spread parameter

$$\sigma\left(x[m+m', n+n'], x[m,n]\right) = K[m+m', n+n'] \times \sigma_m[m,n] \quad (8)$$

where σ_m is a position-dependent amplitude of the spread parameter σ and K is the scaling function controlled by the direction of $x[m+m', n+n']$ to $x[m,n]$. The extensions of membership function σ in (8) will be discussed in Section III for compressed images and Section IV for compressed video sequences.

III. Directional Fuzzy Spatial Filter

A. Directional Spread Parameter

When highly compressed, the ringing artifacts in JPEG images are prevalent along strong edges and the filter strength should adapt to the edge direction. For example, in Fig. 2(b), the filter should ideally apply stronger smoothing in the horizontal direction, where the ringing artifacts are likely to have no relation with the original value, and a weaker filtering in the vertical direction, which is the edge direction of the image. One general form of cosine-based spread parameter which satisfies this requirement is

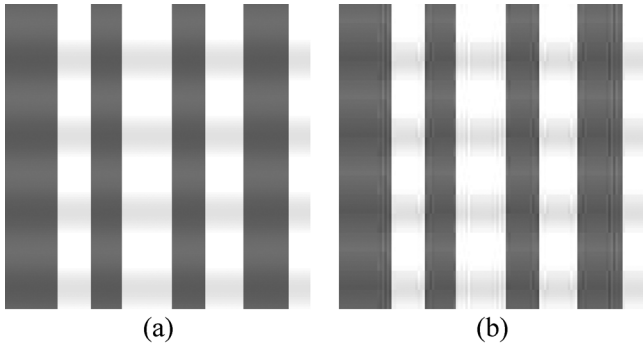

Fig. 2. Example of directional JPEG artifacts with scaling factor of 4 for the quantization step matrix. (a) Original image; (b) compressed.

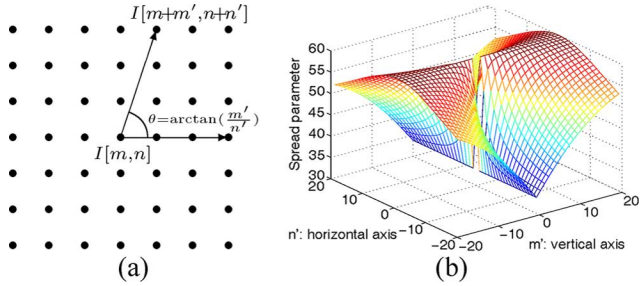

Fig. 3. Angle and spread parameter for directional fuzzy filter. (a) Angle θ; (b) spread parameter.

$$\sigma(\theta) = \sigma_m\left(\alpha + \beta\cos^2(\theta)\right) \quad (9)$$

where θ is the direction between the pixel of interest $I[m,n]$ and its surrounding pixels $I[m+m', n+n']$ as shown in Fig. 3(a), σ_m is the amplitude of the spread parameter, α and β are positive scaling factors which control the maximum and minimum strength of the directional filter. In (9), $\sigma(\theta)$ attains the minimum value $\sigma_{\min} = \alpha\sigma_m$ in the vertical direction and the maximum value $\sigma_{\max} = (\alpha+\beta)\sigma_m$ in the horizontal direction. An example of the directional spread parameter is plotted in Fig. 3(b) with $\sigma_m = 15$, $\alpha = 0.5$ and $\beta = 3.5$.

B. Edge-Based Directional Fuzzy Filter

For real images with more complicated edges, the strongest filtering is applied to the direction perpendicular to the edge. Based on the Sobel operator with horizontal and vertical derivative approximation of the gradient $G_x = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \times I$ and $G_y = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix} \times I$, the edges are detected by using the gradient magnitude $G = \sqrt{G_y^2 + G_x^2}$. Its corresponding direction is determined by $\theta_0 = atan(G_y/G_x)$. The spread function in this case is determined by the angle $(\theta - \theta_0)$ instead of θ in (9), where the angles θ and θ_0 are defined as in

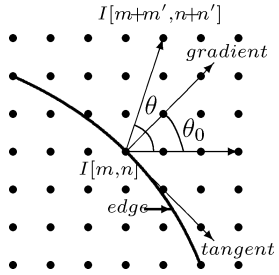

Fig. 4. Angles θ and θ_0 of the edge-based directional fuzzy filter.

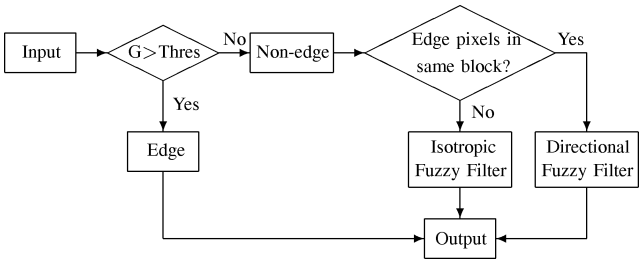

Fig. 5. Flow chart of the directional fuzzy filter.

Fig. 4. To be adaptive for different areas having different activity levels, the standard deviation $STD(I[m,n])$ of pixels in the window W centered on $I[m,n]$ is used to control the amplitude of the spread parameter σ_m in (9) as

$$\sigma_m[m,n]=\sigma_0\left((1-\gamma)\left(\frac{STD\,(I[m,n])-STD_{\min}}{STD_{\max}-STD_{\min}}\right)+\gamma\right) \tag{10}$$

where $STD_{\max}$ and $STD_{\min}$ are, respectively, the maximum and minimum value of all $STD(I[m,n])$ values in the current frame, σ_0 is the maximum spread parameter value and γ is the scaling factor in [0,1]. σ_m is scaled to $[\gamma\sigma_0\ \sigma_0]$ so that the fuzzy filter is still applied with $\sigma_m=\gamma\sigma_0$ to the lowest activity areas. By adjusting σ_0 and γ, the balance between edge preservation and artifact removal can be achieved. The proposed algorithm for edge-based directional fuzzy filtering is shown in Fig. 5. The pixels are first classified into edge pixels and nonedge pixels by comparing the gradient magnitude to an empirically determined threshold. Edge pixels are not be filtered because they are not ringing pixels. For nonedge pixels, if there are no edge pixels in the same block, the ringing artifacts in this block are not considered to be oriented in any particular direction and are filtered with an isotropic fuzzy filter. For the remaining nonedge pixels, the tangent angle of their nearest edge pixel is used to control the directional spread parameter.

IV. Adaptive Fuzzy Compensated Spatiotemporal Filter

In this section, the directional fuzzy filter is extended for artifact reduction in compressed video sequences I. To increase the correlation between pixels, the surrounding frames are motion compensated before applying the MCSTF as shown in Fig. 6. The chroma components are first upsampled to the same size of the luma component. To obtain more accurate motion vectors, each frame is enhanced by an isotropic spatial fuzzy filter before the motion estimation phase. Next, the adaptive fuzzy filter

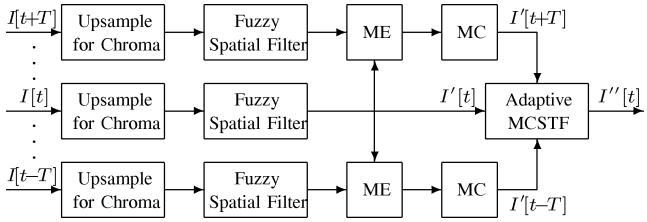

Fig. 6. Block diagram of the adaptive fuzzy MCSTF.

is applied to the set Ω of spatiotemporal surrounding pixels centered by the pixel of interest $I'[t,m,n]$

$$I''[t,m,n]=\frac{\sum\limits_{[t',m',n']\in\Omega}h[t',m',n']I'[t+t',m+m',n+n']}{\sum\limits_{[t',m',n']\in\Omega}h[t',m',n']} \tag{11}$$

where

$$h[t',m',n']=exp\left(-\frac{(I'[t+t',m+m',n+n']-I'[t,m,n])^2}{2\sigma^2[t',m',n']}\right) \tag{12}$$

is the fuzzy filter coefficient for the surrounding pixel at the location $[t',m',n']$ from the pixel of interest $I'[t,m,n]$ and

$$\sigma[t',m',n']=K[t',m',n']\times\sigma_m[m,n] \tag{13}$$

is the spread parameter with the amplitude σ_m and scaling factor K as mentioned in (8).

Similar to Section III-B, the standard deviation of pixels in spatiotemporal cubic C centered on $I'(t,m,n)$ is used to adaptively control the amplitude of the spread parameter σ_m as in (10). Furthermore, the fuzzy MCSTF filter should apply strongest filtering to the pixels in surrounding motion compensated frames at the same spatial position due to their strongest correlation to $I'(t,m,n)$ and weaker in other positions. Based on the cross-correlation value of pixels in the windows of the current frame and its surrounding frames, the scaling factor of the spread parameter is determined by

$$K[t',m',n']=\frac{\sigma_{12}}{\sigma_1\sigma_2} \tag{14}$$

where

$$\sigma_{12}=\sum_{[m_0,n_0]\in V}I'[t+t',m+m'+m_0,n+n'+n_0]\times I'[t,m+m_0,n+n_0] \tag{15}$$

$$\sigma_1=\sqrt{\sum_{[m_0,n_0]\in V}I'[t+t',m+m'+m_0,n+n'+n_0]^2} \tag{16}$$

and

$$\sigma_2=\sqrt{\sum_{[m_0,n_0]\in V}I'[t,m+m_0,n+n_0]^2} \tag{17}$$

where V is a spatial window centered on the pixel of interest $I'(t,m,n)$ of the current frame or $I'[t+t',m+m',n+n']$ of the

surrounding frames. Higher correlation between the pixels in the 2 windows leads to more contribution to the output of the pixel at $[t', m', n']$. This scaling factor also follows the constraint

$$K[t', m', n'] \leq K[0, 0, 0] = 1 \tag{18}$$

which makes sure that the input $I'[t, m, n]$ always has highest contribution to the output.

V. Motion Compensated Metric for Flickering Artifact Evaluation

This section discusses a new metric to evaluate the flickering artifacts. Previous flickering metrics focused on flickering artifacts of intra frame coding in H.264 [3] and Motion JPEG2000 [14]. In [3], the flickering of the $[i, j]^{th}$ block was calculated by the sum of square difference (SSD) between the temporal flickering in the original frames O and compressed frames I

$$SSD_{\text{dif}}[t, i, j] = \sum_{[m,n] \in [i,j]^{\text{th}}\text{block}} (D[t, m, n] - D[t-1, m, n])^2 \tag{19}$$

where

$$D[u, m, n] = O[u, m, n] - I[u, m, n]. \tag{20}$$

The metric S for the whole frame only took into account the blocks with small temporal SSD value in the original sequence

$$S[t] = \frac{1}{L} \sum_{SSD_{\text{org}}[t,i,j] \leq \epsilon} SSD_{\text{dif}}[t, i, j] \tag{21}$$

where L was the number of blocks in frame t which satisfy $SSD_{\text{org}}[t, i, j] \leq \epsilon$. SSD_{org} was defined as the SSD over temporal direction between original frames at times t and $t-1$

$$SSD_{\text{org}}[t, i, j] = \sum_{[m,n] \in [i,j]^{\text{th}}\text{block}} (O[t, m, n] - O[t-1, m, n])^2. \tag{22}$$

In (19), the metric does not consider SSD_{org}, but the same SSD_{dif} makes the flickering artifact less perceptible with high SSD_{org} than small SSD_{org}. A normalized metric should be considered to make it comparable for different blocks, different frames or different sequences.

In [14], the metric applied the sum of square difference (SSD) operator to the metric proposed by [17]. These SSDs between the original and compressed blocks were calculated separately for the current and the previous frames. For $[i, j]^{th}$ block, the final metric was a fraction of the difference and the sum of these two SSDs

$$SSD_{\text{dif}}[t, i, j] = \frac{\sum_{[m,n] \in [i,j]^{\text{th}}\text{block}} \left(D[t, m, n]^2 - D[t-1, m, n]^2\right)}{\sum_{(m,n) \in [i,j]^{\text{th}}\text{block}} (D[t, m, n]^2 + D[t-1, m, n]^2)}. \tag{23}$$

Because of the square function in (23), this metric ignores the signs of the differences between the collocated blocks before and after compression. As shown in the example in Fig. 7 with the case of $D[t-1, m, n] = 0.375$ and $D[t, m, n] = -0.375$,

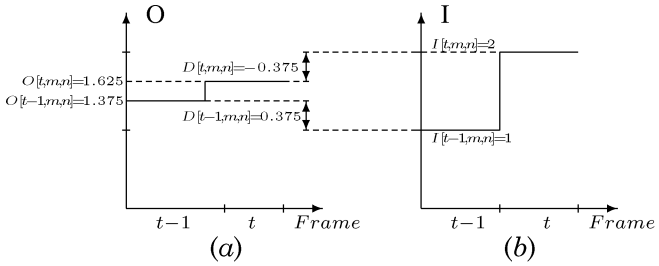

Fig. 7. Example where the flickering metric in [14] has problem. (a) Original sequence; (b) compressed sequence.

TABLE I
Comparison of PSNR in Units of dB for Different Methods

Sequences	4Q	Chen	Liu	Conventional Fuzzy	Adaptive Fuzzy
News	27.48	27.58	27.55	27.94	**28.05**
Silent	27.84	28.37	28.33	28.33	**28.58**
Foreman	28.06	28.46	28.41	28.78	**28.87**
Mobile	21.22	20.96	21.13	21.50	**21.55**
Mother	31.02	31.83	31.62	31.77	**32.00**
Paris	23.38	23.25	23.31	23.80	**23.84**
Average gain		0.2433	0.2267	0.5200	**0.6483**

the metric results in no flickering although there is flickering at pixel $[t, m, n]$.

For interframe coding, the flickering also happens due to the coarse quantization or varied bit allocation for residual signals. Because of the tracking effect of human eyes [24], the motion compensation should be implemented before applying the metric. Therefore, the proposed normalized metric considers the motion of the moving object as well as the signs of these differences

$$SSD_{\text{dif}}[t, i, j] = \frac{\sum_{[m,n] \in [i,j]^{\text{th}}\text{block}} (D[t, m, n] - D[t-1, m+\Delta m, n+\Delta n])^2}{SSD_{\text{Org}}[t, i, j]} \tag{24}$$

where $[\Delta m, \Delta n]$ is the motion vector of block $[i, j]^{th}$ which is estimated based on the original frames. The metric for the whole frame is determined similarly as in (21). The smaller the SSD_{dif} value, the smaller the flickering artifacts.

VI. Simulation Results

A. Enhancement for Compressed Images

Simulations are performed to demonstrate the effectiveness of the directional fuzzy filtering scheme. The qualities of the different approaches are compared in terms of visual quality and PSNR. For comparison, the denoising methods proposed by Chen [7], Liu [8], and Kong [11] are implemented. In the experiments, a 1-D fuzzy deblocking filter as in [11] is applied prior to the proposed directional fuzzy deringing-filter to reduce the blocking artifacts. Only the nonedge pixels that have $G > 210$ are filtered to avoid destroying the real edges of the image. All parameters in Section III are chosen experimentally over a wide range of sequences to achieve the best visual quality.

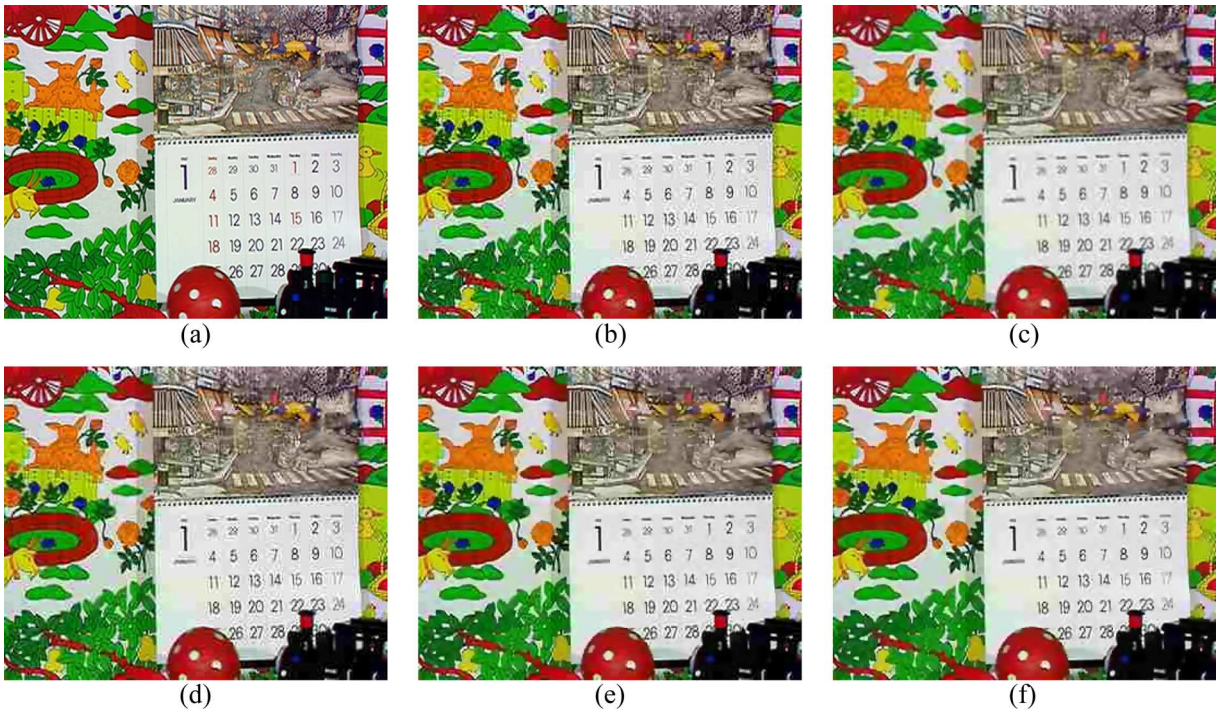

Fig. 8. Comparison of filtered results. (a) Original frame; (b) compressed; (c) Chen's method; (d) Liu's method; (e) conventional fuzzy filter; (f) directional fuzzy filter.

σ_0 is chosen to effectively remove the overall artifacts. γ controls the balance between removing the artifacts in flat areas and keeping the details in high activity areas. α and β are used to adjust relative filtering strength between the gradient and tangent directions of edges. These parameters are experimentally chosen with $\sigma_0 = 15$, $\alpha = 0.5$, $\beta = 3.5$ and $\gamma = 0.5$. The set Ω of neighboring pixels and the spatial window W size are set to 5×5. Several CIF resolution video sequences are compressed using motion JPEG with a scaling factor of 4 for the quantization step matrix. The test images are the frames taken from Silent, Foreman, Mobile, Paris, News, and Mother sequences.

In the case of the JPEG image in Fig. 2(b) with only vertical edges, Fig. 10 shows the enhanced images using the isotropic fuzzy filter and the directional fuzzy filter. For this simulation, the spread parameter of the isotropic fuzzy filter is fixed with $\sigma = 15$. Compared to the compressed image in Fig. 2(b) (39.77 dB), the enhanced image using the isotropic fuzzy filter in Fig. 10(a) (45.53 dB) and the enhanced image using the directional fuzzy filter in Fig. 10(b) (47.82 dB) achieve significant improvement in visual quality and PSNR. This shows the effectiveness of fuzzy filter in reducing both blocking and ringing artifacts. It also demonstrates the basic merit of the directional fuzzy filter to more substantially reduce the ringing artifacts compared to isotropic fuzzy filtering.

For images with more complicated edges, the simulation is performed on the 4th frame of the Mobile sequence. The Sobel operator as described in Section III-B is used to estimate the gradient of the edges. As shown in Fig. 11(a) for one part of the deblocked image and in Fig. 11(b) for its gradient, the Sobel operation is robust in estimating the gradient of the edges having ringing artifacts. Fig. 12 shows the deblocked image and its classification map for directional deringing. In this map, the cyan pixels are edge-pixels, magenta pixels are nonedge pixels which are directionally filtered and blue pixels are nonedge pixels which are isotropic filtered as with edge-pixels. Table I summarizes the PSNR results for one frame of all sequences when different enhancement techniques are applied. Each row shows the PSNR in dB of one frame of each video sequence when using different methods for quality enhancement. The last row indicates the average gain in PSNR of the enhanced image over its compressed image. These numerical results show that the directional fuzzy filter provides higher PSNR improvement over existing techniques including Chen's method, Liu's method and the conventional fuzzy filtering method that employs isotropic fuzzy spatial filtering. The average gains for Chen's method, Liu's method, the conventional fuzzy filtering method and the proposed method are 0.2433, 0.2267, 0.5200, 0.6483 dB, respectively.

To evaluate the visual quality, results with different denoising techniques on compressed 4th frame of the Mobile sequence are shown in Fig. 8 for full frame views and Fig. 9 for zoomed views. The results show that the DCT-based low-pass filtering techniques proposed by Chen is able to suppress some of the ringing artifacts, but introduces a substantial amount of blur in the processed image. Liu's method is able to retain some of the sharpness, but is not able to reduce the ringing artifacts. The conventional fuzzy filter shows much less ringing around the edges, especially within the calendar area. It is clear from these visual results that the directional fuzzy filter has the best quality as it is able to further reduce ringing over the conventional fuzzy filtering approach and outperforms other existing denoising techniques.

To see the individual contributions of the spatial and directional adaptations respectively, another simulation was

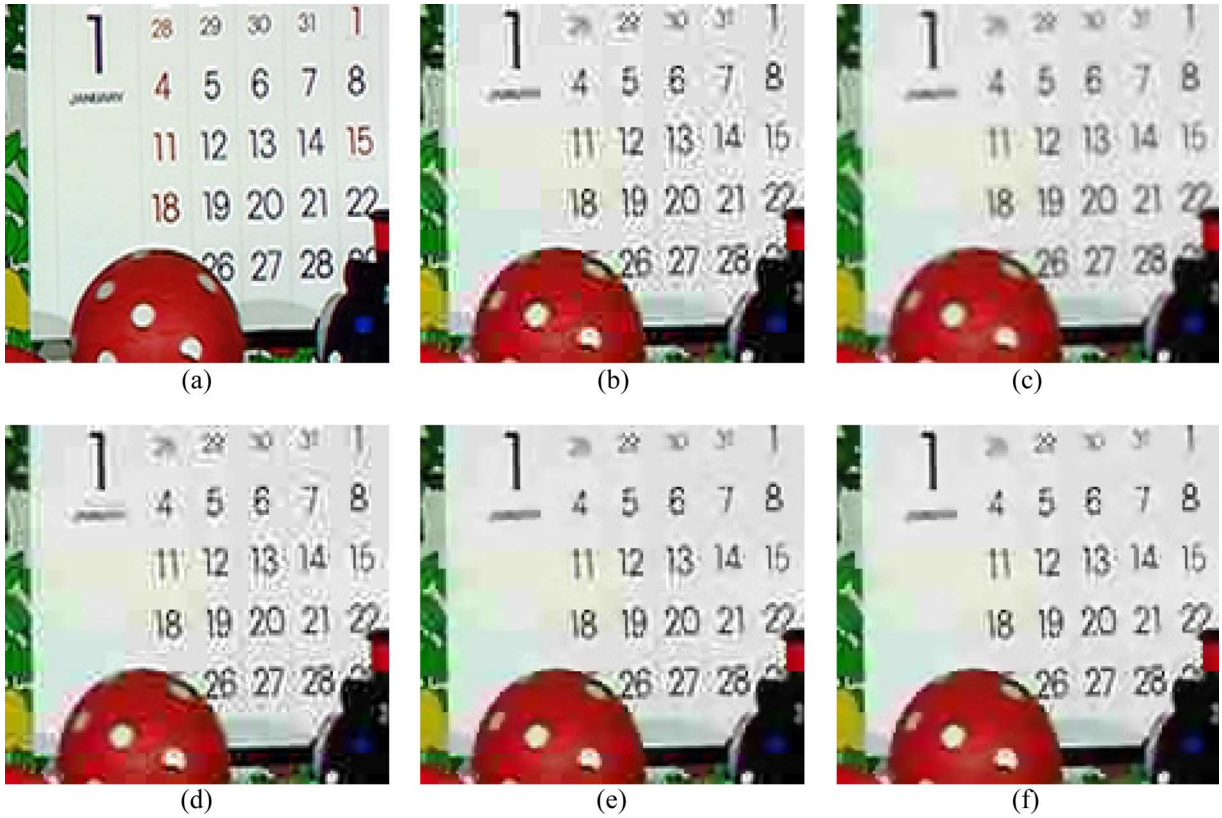

Fig. 9. Zoomed images for comparison of filtered results. (a) Original frame; (b) compressed; (c) Chen's method; (d) Liu's method; (e) conventional fuzzy filter; (f) directional fuzzy filter.

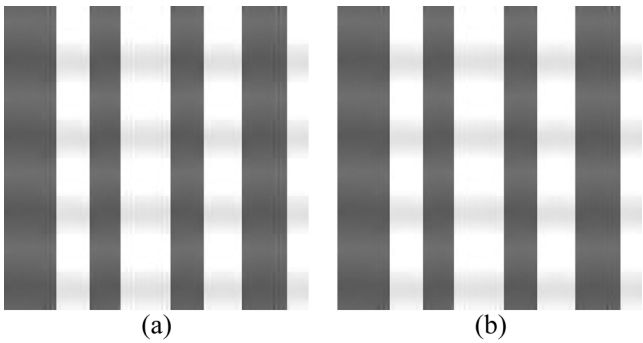

Fig. 10. Result of using a fuzzy filter. (a) Isotropic (45.53 dB); (d) directional (47.82 dB).

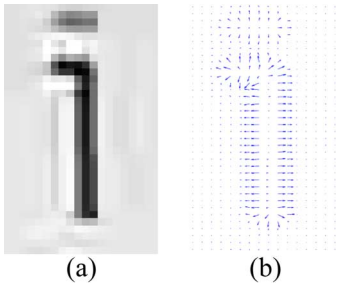

Fig. 11. Gradient of the deblocked image. (a) One part of the deblocked image; (b) gradient of (a).

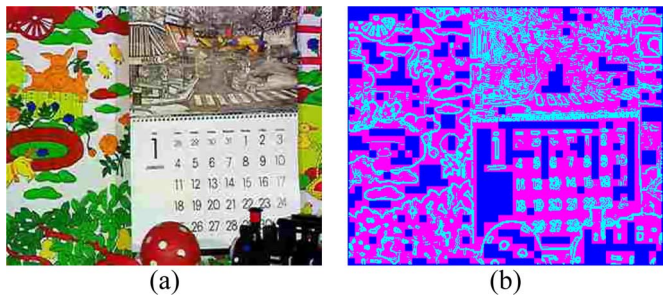

Fig. 12. Pixel classification for directional filtering. (a) Deblocked image; (b) pixel classification of (a).

performed for the cases of using only the spatial adaptation (without directional adaptation), using only the directional adaptation (without spatial adaptation) and using both the spatial and directional adaptations. The results are shown in Fig. 13 for the whole filtered frames and in Fig. 14 for one zoomed in part. Figs. 13(b) and 14(b) show that only using the directional adaptation reduces effectively the ringing artifacts but blurs the filtered frame. The blurriness is caused by using the fixed amplitude of the spread parameter σ_m for all pixels. Using only the spatial adaptation preserves the details but cannot effectively reduce the ringing artifact, as shown in Figs. 13(c) and 14(c). Combining spatial and directional adaptation can both reduce the ringing artifacts and still keep the details of the enhanced frames, as shown in Figs. 13(a) and 14(a). Spatial adaptation helps removing the overall ringing artifacts and avoiding blurring the frame while directional adaptation helps further removing the ringing artifacts around

Fig. 13. Comparison on the contribution of spatial and directional adaptations. (a) Spatial-directionally adaptive; (b) directionally adaptive; (c) spatially adaptive.

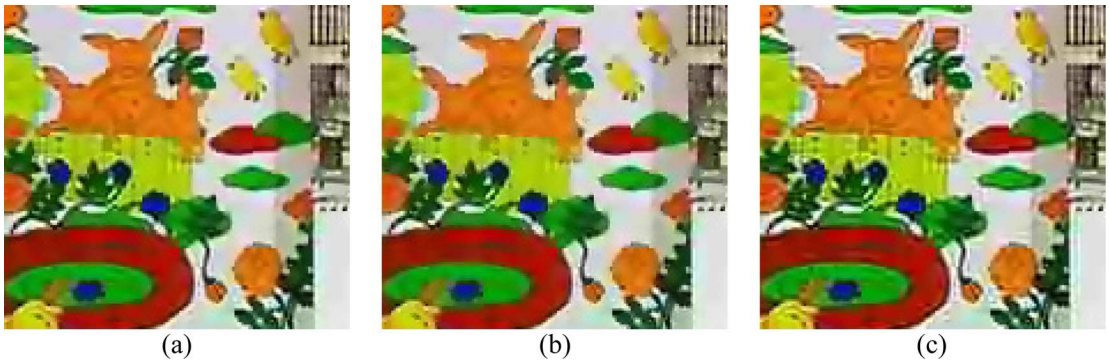

Fig. 14. Zoomed images for comparison on the contribution of spatial and directional adaptations. (a) Spatial-directionally adaptive; (b) directionally adaptive; (c) spatially adaptive.

TABLE II
COMPARISON OF PSNR IN UNITS OF dB OF DIFFERENT CLASSIFIED PIXELS AND OF SPATIAL AND DIRECTIONAL ADAPTATIONS

Sequences	Percentage of Classified Pixels (%)			Non-Edge Pixels				Total			
				Directional Filtering		Isotropic Filtering					
	Edge	Non-edge		JPEG	Proposed	JPEG	Proposed	JPEG	Proposed	Directionally	Spatially
		Directional	Isotropic	(dB)	(dB)	(dB)	(dB)	(dB)	(dB)	Adaptive (dB)	Adaptive (dB)
News	7.11	21.45	71.44	24.92	25.78	30.39	31.08	27.48	28.05	27.58	28.08
Silent	2.72	14.19	83.09	25.61	26.75	28.56	29.24	27.84	28.58	28.45	28.59
Foreman	6.40	22.78	70.82	25.86	27.25	30.72	31.53	28.06	28.87	28.71	28.95
Mobile	25.73	53.02	21.25	21.54	22.16	25.73	26.08	21.22	21.55	21.20	21.50
Mother	1.49	8.06	90.45	27.52	29.41	31.64	32.48	31.02	32.00	31.65	32.03
Paris	18.84	38.37	42.79	22.59	23.43	28.27	28.96	23.38	23.84	23.45	23.75
Ave. Gain					1.1234		0.6767		0.6483	0.3400	0.6500

the edges to achieve better visual quality. The PSNR values of the enhanced frames are listed in the last four columns of Table II for all sequences. The average PSNR improvement of using spatial-directional adaptation, using only directional adaptation and using only spatial adaptation are 0.6483, 0.3400, and 0.6500 dB, respectively. Although having slightly smaller PSNR improvement than using only spatial adaptation, the combined spatial-directional adaptation has the best visual quality.

An additional simulation was also performed to justify the contribution of directional filtering and isotropic filtering of the nonedge pixels in Fig. 5. The percentage and PSNR of the classified pixels for all sequences are shown in Table II. The average PSNR improvements of directional filtering and isotropic filtering of the nonedge pixels are 1.1234 and 0.6767 dB, respectively. These results validate the effectiveness of the edge-based directional fuzzy filter discussed in Section III-B. The overall average PSNR improvement of the proposed adaptive fuzzy filter is 0.6483 dB. This improvement is smaller than the averaged PSNR improvement of filtering the nonedge pixels with directional filtering or isotropic filtering. That is because the edge pixels are not filtered and there is no improvement from these pixels.

Compared to Chen's method and Liu's method which are DCT-based methods, the proposed directional fuzzy filtering method is performed in the pixel domain and has less computational complexity. On the other hand, the proposed filter requires an edge detection phase, which increases the complexity of the proposed method slightly compared to the conventional fuzzy filter. However, with the merit of the di-

(a) (b) (c)

Fig. 15. Comparison of filter results for MJPEG sequences. (a) Compressed (21.22 dB); (b) fuzzy spatial filter (21.50 dB); (c) proposed fuzzy spatiotemporal filter (21.89 dB).

(a) (b) (c)

Fig. 16. Zoomed views for images in Fig. 15. (a) Compressed; (b) fuzzy spatial filter; (c) proposed fuzzy spatiotemporal filter.

rectional fuzzy filter in further removing the ringing artifacts around the edges, this extra complexity seems well-justifiable in many applications.

B. Enhancement For Compressed Video Sequences

1) Enhancement For MJPEG Video Sequences: To demonstrate the advantage of using temporal correlation, the simulation in this section is performed on MJPEG sequences. In this codec, each frame is compressed separately using the JPEG standard and the temporal redundancies between frames are not utilized for coding as in other codecs. Therefore, it is expected that the use of such temporal redundancies (i.e., correlation among frames) for postfiltering could lead to more pronounced quality improvement in this case. For the purposes of practical implementation and focusing on demonstrating the advantage of using extra information from surrounding frames, the motion compensation stage in Fig. 6 is omitted to reduce running time and the scaling factor of the spread parameter is chosen $K = 1$. To be consistent with the method of using fuzzy spatial filter in [11], the same 1-D deblocking fuzzy filter and the same algorithm for choosing amplitude of spread parameter σ_m are used in the adaptive fuzzy spatiotemporal filter. The sizes of the set Ω and the spatiotemporal cubic C are $5 \times 5 \times 5$ pixels while that of the spatial window V is 5×5 pixels.

Fig. 15 compares the enhanced images obtained by the fuzzy spatial and proposed fuzzy spatiotemporal filters. The enhanced image obtained by the proposed fuzzy spatiotemporal filter [Fig. 15(c)] shows significantly reduced ringing artifacts and better color quality than the spatial counterpart [Fig. 15(b)]. The drastic improvement in visual quality is more readily observable in the enlarged portion of the picture as shown in Fig. 16. In the conventional fuzzy case, the deringing filter was applied only to the luminance component 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 spatiotemporal extension, more chroma samples are available from the neighboring frames and the use of deringing filter for chroma components helps to improve the color quality significantly, as shown in Figs. 15 and 16.

Next, Fig. 18 compares the PSNRs of all the tested methods for the Mobile sequence. The plots clearly demonstrate that the proposed fuzzy spatiotemporal filter achieves consistent PSNR gains of about 0.67, 0.91, 0.72, and 0.40 dB on average relative to the compressed images, those by Chen's method, Liu's method, and the conventional fuzzy spatial filter, respectively. The effectiveness of the proposed scheme was much more noticeable when the processed frames were played back as a sequence, as the proposed spatiotemporal result produces a smoother video with significantly reduced mosquito and flickering artifacts.

To validate the effectiveness of the proposed flickering metric, the metrics in [3] and [14] and the proposed metric is applied to the compressed and enhance Mobile sequences. These results are shown in Fig. 17. The subjective tests show that the compressed sequence has the most flickering artifacts

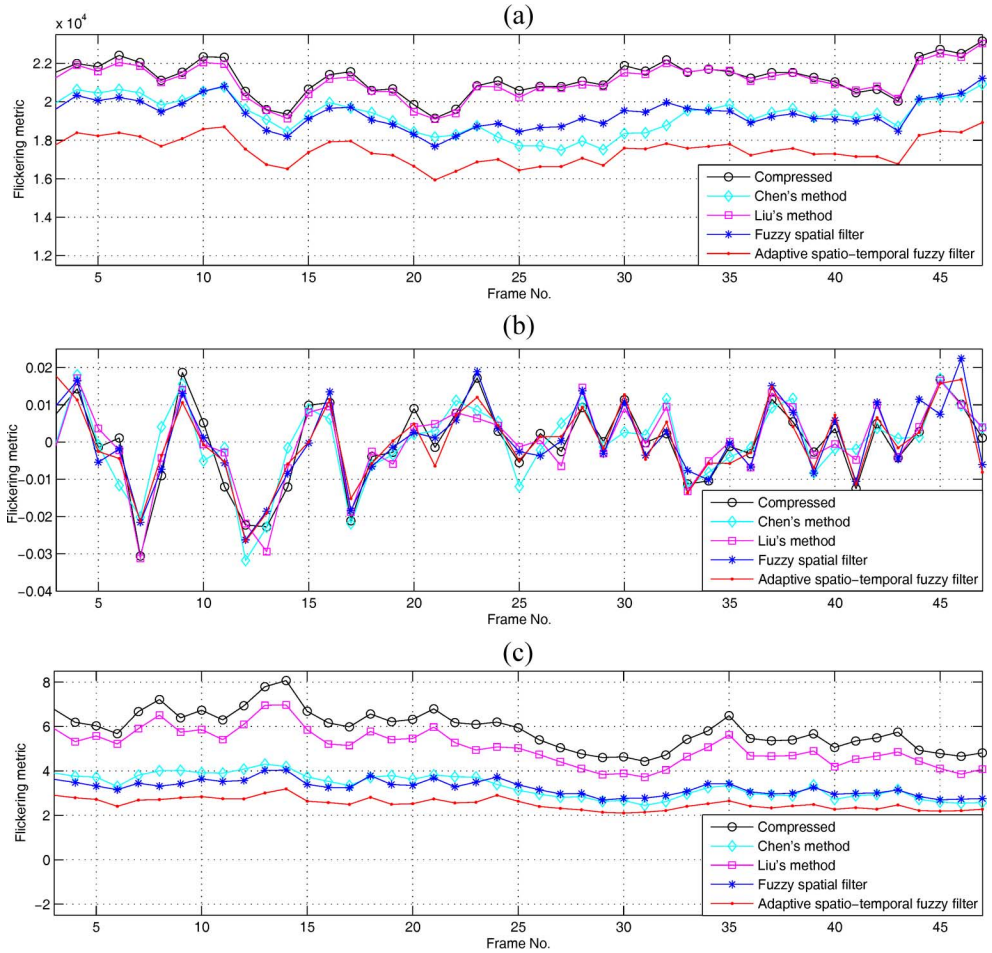

Fig. 17. Comparison on flickering artifacts of simulated methods for mobile sequence.

while the enhanced sequence using the proposed spatiotemporal has the least flickering artifacts. The metric in [14] is not correlated to the flickering artifacts as shown in Fig. 17(b). The metric in [3] gives the similar flickering evaluation for both the compressed sequence and enhanced sequence of Liu's method as shown in Fig. 17(a), but the enhanced sequences of Liu's method has less flickering than the compressed sequence when played back. In contrast, the proposed metric was well-correlated with the subjective flickering evaluation of the Mobile as shown in Fig. 17(c) as well as with those of other video sequences consistently. Furthermore, the resulted metric is comparable when the flickering evaluation for different frames or different sequences is needed.

2) Enhancement For H.264 Video Sequences: In order to demonstrate that the proposed fuzzy filter is beneficial even for interframe-coded videos, which tend to have less flickering artifacts compared with the intracoded MJPEG sequences in the previous subsection, further experiments are performed with H.264-coded videos. The Foreman sequence was compressed with the prediction structure of IBBBPBBBP at bit-rate 132 Kbps. The in-loop deblocking filter was enabled. The spread parameter was set to $\sigma_0 = 20$ and the offset γ in (10) was set to 0.5. These parameters were chosen experimentally to get the best visual quality for a wide range of enhanced sequences. The motion vectors were estimated by full-search

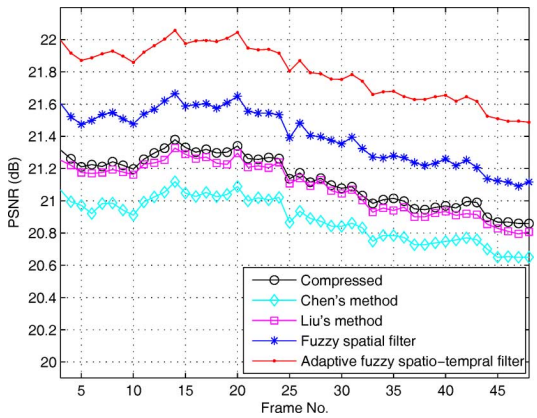

Fig. 18. Comparison on PSNR of simulated methods for mobile sequence.

motion estimation with a search range 24×24. Fig. 19 shows the compressed image using the in-loop deblocking filter [Fig. 19(b)], enhanced images obtained by Chen's method [Fig. 19(c)], Liu's method [Fig. 19(d)], fuzzy spatial filter [Fig. 19(e)] and adaptive fuzzy MCSTF [Fig. 19(f)]. Chen's method effectively removes artifacts but the resulting images tend to look blurry. Both Liu's method and fuzzy spatial filter only slightly remove the blocking artifact. Recall from (5) that the difference in the pixel intensity values determines the

Fig. 19. Comparison of filter results for H.264 sequences. (a) Original frame; (b) compressed (30.77 dB); (c) Chen's method (30.37 dB); (d) Liu's (30.51 dB); (e) fuzzy spatial filter (30.90 dB); (f) adaptive fuzzy MCSTF (31.09 dB).

relative contribution of each input sample to the filter output. Because of the large difference in the intensity values of the pixel of interest and its spatiotemporally neighboring pixels having blocking artifacts, blocking artifacts in the surrounding motion-compensated frames have small contribution to the output when using the proposed adaptive fuzzy MCSTF. The proposed method significantly reduces the artifacts and yields better color quality than other methods. It also has the highest PSNR improvement (+0.32 dB), comparing to the PSNR improvement of Chen's method (−0.40 dB), Liu's method (−0.26 dB) and fuzzy spatial filter (+0.13 dB). This improvement is consistent for the Foreman sequence which is verified by the PSNR curves in Fig. 20. Additional simulation shows that the spatial adaptation and directional adaptation contribute +0.26 dB and +0.06 dB to the total PSNR improvement, respectively. Spatial adaptation removes the overall artifacts and preserves the details while directional adaptation further reduces the ringing and flickering artifacts. Further simulation using the proposed directional fuzzy spatial filter in Section III shows that the single frame-based enhanced image (30.97 dB) has less blocking and ringing artifacts than the compressed image [Fig. 19(b)] but has more flickering artifacts than the enhanced image using the directional fuzzy spatiotemporal filter [Fig. 19(f)].

The flickering artifacts are evaluated by the proposed flickering metric and these results are shown in Fig. 21. This metric shows that the flickering artifacts are reduced when using the directional MCSTF. The subjective tests also validate this conclusion. The enhanced video sequence using the proposed method has less blocking, mosquito and flickering than the compressed sequence using the in-loop deblocking filter and other enhanced sequences using Chen's method, Liu's method and fuzzy spatial filter. The PSNR improvement of the proposed method for different bit-rates of the Foreman sequence is shown in Fig. 22. The adaptive fuzzy MCSTF yields more than +0.2 dB PSNR improvement for bit-rates from 70 Kbps to 170 Kbps. The enhanced sequences using the adaptive fuzzy MCSTF also have better visual quality with less artifacts than other methods for the bit-rates in this range. Please note that the proposed method requires the motion compensation stage and the spatiotemporal filter, so its computational complexity is higher compared to those of Chen's method, Liu's method or fuzzy spatial filtering method. However, the motion compensation step is necessary to better align matching pixels and increase the correlation of the surrounding pixels to the pixel of interest. Running time to enhance one frame of the fuzzy spatial filtering method is comparable to the running time of Chen's method and Liu's method. The running time of the spatiotemporal filtering method using five frames is 1.5 and 8 times longer than the fuzzy spatial filtering method for the versions without and with motion compen-

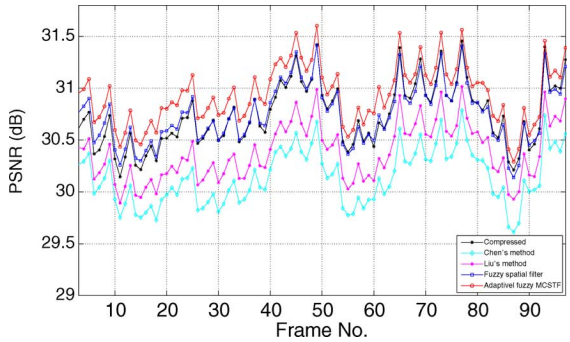

Fig. 20. Comparison of PSNR for all frames in the Foreman sequence.

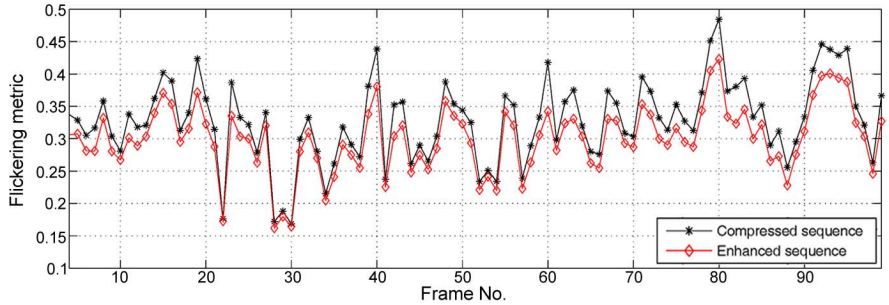

Fig. 21. Comparison of flickering metric for all frames in the Foreman sequence.

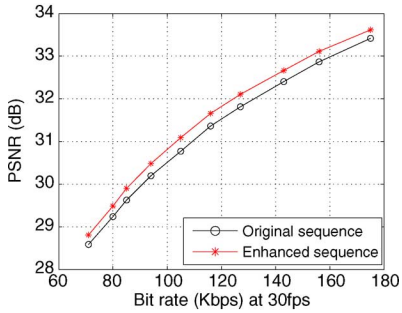

Fig. 22. Comparison of PSNR with different bit-rates of the Foreman sequence.

sation, respectively. All image and video results can be found at http://www.videoprocessing.ucsd.edu/~dungvo/MCSTF.html.

VII. Conclusion

An effective algorithm for image and video denoising using an adaptive fuzzy filter is proposed. This novel method overcomes the limitations of conventional nonlinear filters by accounting for pixel's activity and the direction between pixels. It is shown that the proposed adaptive fuzzy filter improves both visual quality and PSNR of compressed images and videos compared to existing approaches. The flickering artifact reduction is evaluated by the proposed flickering metric. The proposed adaptive scheme can be applied to bilateral filters which do not use the directional information between pixels. A future adaptive MCSTF can be considered for segmented moving objects over frames. Human visual system (HVS) should be incorporated to evaluate the flicking artifacts based on artifact perception for different areas.

Acknowledgment

The authors would like to thank to the authors of [11] for the fuzzy spatial filter code which is referenced in the simulation. They would also like to thank the reviewers for their useful comments.

References

[1] A. Jerri, *The Gibbs Phenomenon in Fourier Analysis, Splines and Wavelet Approximations*. Dordrecht, The Netherlands: Kluwer, 1998.

[2] M. Kaneko, Y. Hatori, and A. Koike, "Improvements of transform coding algorithm for motion-compensated interframe prediction errors-DCT/SQ coding," *IEEE J. Sel Areas Commun.*, vol. 5, no. 8, pp. 1068–1078, Aug. 1987.

[3] X. Fan, W. Gao, Y. Lu, and D. Zhao, "Flicking reduction in all intra frame coding," Joint Video Team of ISO/IEC MPEG and ITU-T VCEG, JVT-E070, Oct. 2002.

[4] S. Sakaida, K. Iguchi, S. Gohshi, and Y. Fujita, "Adaptive quantization control for reducing flicker of AVC/H.264 intra frames," presented at the Picture Coding Symp., Dec. 2004.

[5] T. Jarske, P. Haavisto, and I. Defee, "Post-filtering methods for reducing blocking effects from coded images," *IEEE Trans. Cosum. Electron.*, vol. 40, no. 8, pp. 521–526, Aug. 1994.

[6] A. Nosratinia, "Embedded post-processing for enhancement of compressed images," in *Proc. IEEE Data Compression Conf.*, 1999, pp. 62–71.

[7] T. Chen, H. 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.

[8] S. Liu and A. 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.

[9] B. Gunturk, Y. Altunbasak, and R. M. Mersereau, "Multiframe blocking-artifact reduction for transform-coded video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 4, pp. 276–282, Apr. 2002.

[10] S. Oguz, Y. Hu, and T. Nguyen, "Image coding ringing artifact reduction using morphological post-filtering," in *Proc. IEEE Int. Work. Multimedia Signal Processing*, 1998, pp. 628–633.

[11] H. Kong, Y. Nie, A. Vetro, H. Sun, and K. Barner, "Adaptive Fuzzy Post-Filtering for Highly Compressed Video," in *Prof. IEEE Int. Conf. Image Proc.*, 2004, pp. 1802–1806.

[12] S. Westen, R. Lagendijk, and J. Biemond, "Adaptive spatial noise shaping for DCT based image compression," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, May 1996, vol. 4, pp. 2124–2127.

[13] S. DelCorso, C. Miro, and J. Jung, "MNR: A novel approach to correct MPEG temporal distortions," *IEEE Trans. Consum. Electron.*, vol. 49, no. 2, pp. 229–236, Feb. 2003.

[14] A. Leontaris, Y. Tonomura, T. Nakachi, and P. Cosman, "Flicker suppression in JPEG2000 using segmentation-based adjustment of block truncation lengths," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, Apr. 2007, vol. 1, pp. 1117–1120.

[15] S. M. Smith and J. M. Brady, "Susan-a new approach to low level image processing," *Int. J. Comput. Vis.*, vol. 23, pp. 45–78, 1997.

[16] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. Int. Conf. Comput. Vis.*, 1998, p. 839846.

[17] E. D. Gelasca and T. Ebrahimi, "On evaluating metrics for video segmentation algorithms," presented at the 2nd Int. Workshop on Video Processing and Quality Metrics for Consumer Applications, Jan. 2006.

[18] B. Zhang and J. P. Allebach, "Adaptive bilateral filter for sharpness enhancement and noise removal," *IEEE Trans. Image Process.*, vol. 17, no. 5, pp. 664–678, May 2008.

[19] S. Kim and J. P. Allebach, "Optimal unsharp mask for image sharpening and noise removal," *J. Electron. Imag.*, vol. 15, p. 0230071, 2005.

[20] H. Hu and G. de Haan, "Trained bilateral filters and applications to coding artifacts reduction," in *Proc. IEEE Int. Conf. Image Processing*, 2007, vol. 1, p. 325328.

[21] Y. Nie and K. Barner, "The fuzzy transformation and its application in image processing," *IEEE Trans. Image Process.*, vol. 15, no. 4, pp. 910–927, Apr. 2006.

[22] B. Justusson, "Median filtering: Statistical properties," in *Two Dimensional Digital Signal Processing II*, T. S. Huang, Ed. Berlin, Germany: Springer-Verlag, 1981.

[23] K. Barner and R. Hardie, "Spatial-rank order selection filter," in *Nonlinear Signal Processing*, S. K. Mitra and G. Sicuranza, Eds. New York: Academic, Apr. 2006, vol. 15, pp. 910–927.
[24] D. A. Robinson, "The mechanics of human smooth pursuit eye movement," *J. Physiol.*, vol. 180, no. 3, pp. 569–591, Oct. 1965.

Dũng T. Võ (S'06) received the B.S. degree in electrical and electronics engineering and the M.S. degree in electronics engineering from the Ho Chi Minh City University of Technology, HCMC, Vietnam, in 2002 and 2004, respectively. He is currently pursuing the Ph.D. degree at University of California at San Diego, La Jolla, as a Fellow of Vietnam Education Foundation (VEF).

He is a teaching staff of the Ho Chi Minh City University of Technology since 2002. He interned at Mitsubishi Electric Research Laboratories (MERL) in summer 2007 and at Thomson Corporate Research in summer 2008. His research interests are interpolation and super-resolution, data pruning-based compression, and algorithms and applications in image and video enhancement.

Mr. Võ received second prize in the Student Scientific Research Competition in 2002, hosted by the Ministry of Education and Training of VietNam, and second prize in the VIFOTEC Contest for Students of the Vietnam Foundation Of Technological Creativity.

Truong Q. Nguyen (F'05) is currently a Professor at the ECE Department, University of California, San Diego, La Jolla. His research interests are video processing algorithms and their efficient implementation. He is the coauthor (with Prof. G. Strang) of the popular textbook *Wavelets and Filter Banks* (Wellesley-Cambridge, 1997) and the author of several Matlab-based toolboxes on image compression, electrocardiogram compression and filter bank design. He has authored over 200 publications.

Prof. Nguyen received the IEEE TRANSACTIONS ON SIGNAL PROCESSING Paper Award (Image and Multidimensional Processing area) for the paper he co-wrote with Prof. P. P. Vaidyanathan on linear-phase perfect-reconstruction filter banks (1992). He received the National Science Foundation Career Award in 1995 and is currently the Series Editor (Digital Signal Processing) for Academic Press. He served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING (1994–1996), the IEEE SIGNAL PROCESSING LETTERS (2001–2003), the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS (1996–1997, 2001–2004), and the IEEE TRANSACTIONS ON IMAGE PROCESSING (2004–2005).

Sehoon Yea (M'06) received the B.S. degree in control and instrumentation engineering in 1993 from Seoul National University, Korea, the M.S. degree in electrical engineering from Nagoya University, Japan, in 1996, and the Ph.D. degree in electrical engineering from Rensselaer Polytechnic Institute, Troy, NY, in 2006.

He has been with Mitsubishi Electric Research Laboratories (MERL) since January 2006, where he is currently a Principal Member of Technical Staff. From 1996 to 2001, he was a Research Engineer at the Institute for Advanced Engineering, Korea, working in control system development. In the summer of 2004, he was an Intern with Sarnoff Corporation, Princeton, NJ. His research interests include image and video processing and multimedia communication.

Anthony Vetro (S'92–M'96–SM'04) received the B.S., M.S., and Ph.D. degrees in electrical engineering from Polytechnic University, Brooklyn, NY.

He joined Mitsubishi Electric Research Labs, Cambridge, MA, in 1996, where he is currently a Group Manager responsibile for research and development in the area of multimedia and information coding. He has published more than 130 papers on video coding, display processing, biometric security, and multimedia adaptation. He has also been an active member of the MPEG and JVT standardization committees for many years, where he has served as an editor for several specifications. Most recently, he has been a key contributor to the Multiview Video Coding amendment of the H.264/AVC standard. He also serves as Vice-Chair of the U.S. delegation to MPEG.

Dr. Vetro is also active in various IEEE conferences, technical committees, and editorial boards. He is currently Chair of the Technical Committee on Multimedia Signal Processing of the IEEE Signal Processing Society and serves on the steering committees for ICME and IEEE TRANSACTIONS ON MULTIMEDIA. He served as an Associate Editor for IEEE SIGNAL PROCESSING MAGAZINE (2006–2007), Conference Chair for ICCE 2006, Tutorials Chair for ICME 2006, and on the Publications Committee of the IEEE TRANSACTIONS ON CONSUMER ELECTRONICS (2002–2008). He is a member of the Technical Committees on Visual Signal Processing and Communications and Multimedia Systems and Applications of the IEEE Circuits and Systems Society. He has also received several awards for his work on transcoding, including the 2003 IEEE Circuits and Systems CSVT Transactions Best Paper Award.