

An Edge-Guided Image Interpolation Algorithm via Directional Filtering and Data Fusion

Lei Zhang, *Member, IEEE*, and Xiaolin Wu, *Senior Member, IEEE*

Abstract—Preserving edge structures is a challenge to image interpolation algorithms that reconstruct a high-resolution image from a low-resolution counterpart. We propose a new edge-guided nonlinear interpolation technique through directional filtering and data fusion. For a pixel to be interpolated, two observation sets are defined in two orthogonal directions, and each set produces an estimate of the pixel value. These directional estimates, modeled as different noisy measurements of the missing pixel are fused by the linear minimum mean square-error estimation (LMMSE) technique into a more robust estimate, using the statistics of the two observation sets. We also present a simplified version of the LMMSE-based interpolation algorithm to reduce computational cost without sacrificing much the interpolation performance. Experiments show that the new interpolation techniques can preserve edge sharpness and reduce ringing artifacts.

Index Terms—Data fusion, edge preservation, image interpolation, linear minimum mean square-error estimation (LMMSE).

I. INTRODUCTION

MANY users of digital images desire to improve the native resolution offered by imaging hardware. Image interpolation aims to reconstruct a higher resolution (HR) image from the associated low-resolution (LR) capture. It has applications in medical imaging, remote sensing and digital photographs [3]–[5], etc. A number of image interpolation methods have been developed [1], [2], [5], [6], [8]–[16]. While the commonly used linear methods, such as pixel duplication, bilinear interpolation, and bicubic convolution interpolation, have advantages in simplicity and fast implementation [7], they suffer from some inherent defects, including block effects, blurred details and ringing artifacts around edges. With the prevalence of inexpensive and relatively LR digital imaging devices and the ever increasing computing power, interests in and demands for high-quality image interpolation algorithms have also increased.

The human visual systems are highly sensitive to edge structures, which convey much of the image semantics, so a key re-

quirement for image interpolation algorithms is to faithfully reconstruct the edges in the original scene. The traditional linear interpolation methods [1]–[3], [5], [6] do not work very well under the edge preserving criterion. Some nonlinear interpolation techniques [8]–[15] were proposed in recent years to maintain edge sharpness. The interpolation scheme of Jensen and Anastassiou [8] detects edges and fits them by some templates to improve the visual perception of enlarged images. Li and Orchard [9] used the covariance of the LR image to estimate the HR image covariance, which represents the edge direction information to some extent, and proposed a Wiener-filtering like interpolation scheme. Since this method needs a relatively large window to compute the covariance matrix for each missing sample, it may introduce some artifacts in local structures due to sample statistics change and, hence, the incorrect estimation of covariance. The image interpolator by Carrato and Tenze [10] first replicates the pixels and then corrects them by using some preset 3×3 edge patterns and optimizing the parameters in the operator. Muresan [15] detected the edge in diagonal and nondiagonal directions and then recovered the missing samples along the detected direction by using one-dimensional (1-D) polynomial interpolation.

Some nonlinear interpolation methods try to enlarge an image by predicting the fine structures in the HR image from its LR counterpart. To do so, a multiresolution representation of the image is needed. Takahashi and Taguchi [11] represented an image by Laplacian pyramid, and with two empirically determined parameters, they estimated the unknown high-frequency components from the LR Laplacian detail signal. In the past two decades, wavelet transform (WT) theory [17] has been well developed and it endows a good multiresolution framework for signal representation. WT decomposes a signal into several scales, along which the signal sharp edges have some correlation. Carey, *et al.* [12] exploited the Lipschitz property of sharp edges in wavelet scales. They used the modulus maxima information at coarse scales to predict the unknown wavelet coefficients at the finest scale. Then, the HR image is constructed by inverse WT. Muresan and Parks [14] extended this strategy by using the entire cone influence of a sharp edge in wavelet scale space, instead of only the modulus maxima, to estimate the finest scale coefficients through optimal recovery theory. The wavelet-based interpolation method by Zhu, *et al.* [13] uses a parametric discrete time model to characterize important edges. With this model, in wavelet domain the lost edge information at the finest scale is recovered through linear minimum mean square-error estimation (LMMSE). The

Manuscript received December 31, 2004; revised October 13, 2005. X. Wu was supported in part by the Natural Sciences and Engineering Research Council of Canada under Grants IRCPJ 283011-01 and RGP45978-2000. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Vicent Caselles.

L. Zhang is with the Department of Computing, The Hong Kong Polytechnic University, Kowloon, Hong Kong (e-mail: cszhang@comp.polyu.edu.hk).

X. Wu is with the Department of Electrical and Computer Engineering, McMaster University Hamilton, ON L8S 4K1 Canada (e-mail: xwu@mail.ece.mcmaster.ca).

Digital Object Identifier 10.1109/TIP.2006.877407

above schemes employ, implicitly or explicitly, an isolated sharp edge model, such as an ideal or smoothed step edge, in the algorithm development. For real images, however, the wavelet coefficients of a sharp edge may be interfered by the neighboring edges. Generally, nonlinear interpolation methods are better at edge preservation than linear methods. In [16], Malgouyres and Guichard analyzed some linear and nonlinear image enlargement methods theoretically and experimentally.

Compared with the discontinuities in 1-D signals, edges in two-dimensional (2-D) images have one more property: the direction. In the linear interpolation methods, 1-D filtering is done alternatively in horizontal and vertical directions without heeding the local edge structures. In the presence of a sharp edge if a missing sample is interpolated across instead of along the edge direction, large and visually disturbing artifacts will be introduced. A conservative strategy to avoid most severe artifacts is to use an isotropic 2-D filter. This, however, reduces the edge sharpness. A more “assertive” approach is to interpolate in an estimated edge direction. The problem with the latter is that the penalty to image quality is high if the estimated edge direction is wrong, which can happen due to the difficulty in determining the edge direction from insufficient data provided by the LR image.

This paper proposes a new balanced approach to the problem. A missing sample is interpolated in not one but two mutually orthogonal directions. The two interpolation results are treated as two estimates of the sample and adaptively fused using the statistics of a local window. Specifically, we partition the neighborhood of each missing sample into two oriented subsets in orthogonal directions. The hope is that the two observation sets will exhibit different statistics, since the missing sample has higher correlation with its neighbors in the edge direction. Each oriented subset yields an estimate of the missing pixel. The pixel is finally interpolated by combining the two directional estimates in the principle of LMMSE. This process can discriminate the two subsets in terms of their coherence to the missing sample, and make the subset perpendicular to the edge direction contribute less to the LMMSE estimate of the missing sample. The above new approach performs significantly better than linear interpolation methods in preserving edge sharpness while suppressing artifacts, by adapting interpolation to local image gradient. A drawback of the proposed interpolation approach is its relatively high computational complexity. We also develop a simplified interpolation algorithm of greatly reduced computation requirement but without significant degradation in performance.

The paper is organized as follows. Section II describes our edge-guided LMMSE-type image interpolation algorithm. Section III presents a simplified version of the algorithm, striving for fast, practical implementations. Section IV reports the experimental results. Section V concludes.

II. EDGE-GUIDED LMMSE-BASED INTERPOLATION

As in many previous papers, we assume an LR image I_l is directly downsampled from an associated HR image I_h through $I_l(n, m) = I_h(2n-1, 2m-1)$, $1 \leq n \leq N$, $1 \leq m \leq M$. Referring to Fig. 1, the black dots represent the available samples of I_l and the white dots represent the missing samples of I_h .

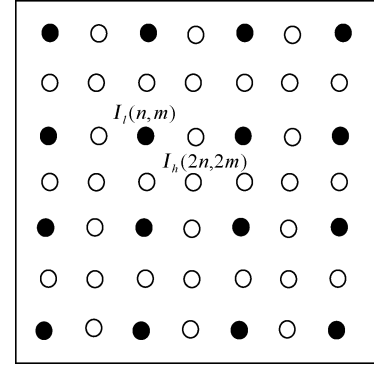


Fig. 1. Formation of an LR image from an HR image by directly down sampling. The black dots represent the LR image pixels and the white dots represent the missing HR samples.

The interpolation problem is to estimate the missing samples in HR image I_h , whose size is $2N \times 2M$, from the samples in LR image I_l , whose size is $N \times M$.

The central issue of image interpolation is how to infer and utilize the information on the missing sample that is hidden in the neighboring pixels. If the downsampled signal of the LR image exceeds the Nyquist sampling limit, the convolution-based interpolation methods will suffer from the aliasing problem in reconstructing the HR image. This is the cause of artifacts such as ringing effects in the interpolated images which are common to linear interpolation methods. Given that the human visual system is highly sensitive to the edges, especially in their spatial locations, it is crucial to suppress the interpolation artifacts while retaining the edge sharpness and geometry.

The edge direction is the most important information for the interpolation process. To extract and use this information, we partition the neighboring pixels of each missing sample into two directional subsets that are orthogonal to each other. From each subset, a directional interpolation is made, and then the two interpolated values are fused to arrive at an LMMSE estimate of the missing sample. We recover the HR image I_h in two steps. First, those missing samples $I_h(2n, 2m)$ at the center locations surrounded by four LR samples are interpolated. Second, the other missing samples $I_h(2n-1, 2m)$ and $I_h(2n, 2m-1)$ are interpolated with the help of the already recovered samples $I_h(2n, 2m)$.

A. Interpolation of Samples $I_h(2n, 2m)$

Referring to Fig. 2, we can interpolate the missing HR sample $I_h(2n, 2m)$ along two orthogonal directions: 45° diagonal and 135° diagonal. Denote by $\hat{I}_{45}(2n, 2m)$ and $\hat{I}_{135}(2n, 2m)$ the two directional interpolation results by some linear methods, such as bilinear interpolation, bicubic convolution, or spline interpolation [1]–[5]. Consider the directional interpolation outputs as the noisy measurements of the missing HR sample

$$\begin{cases} \hat{I}_{45}(2n, 2m) = I_h(2n, 2m) + v_{45}(2n, 2m) \\ \hat{I}_{135}(2n, 2m) = I_h(2n, 2m) + v_{135}(2n, 2m) \end{cases} \quad (2-1)$$

where the random noise variables v_{45} and v_{135} represent the interpolation errors in the corresponding direction.

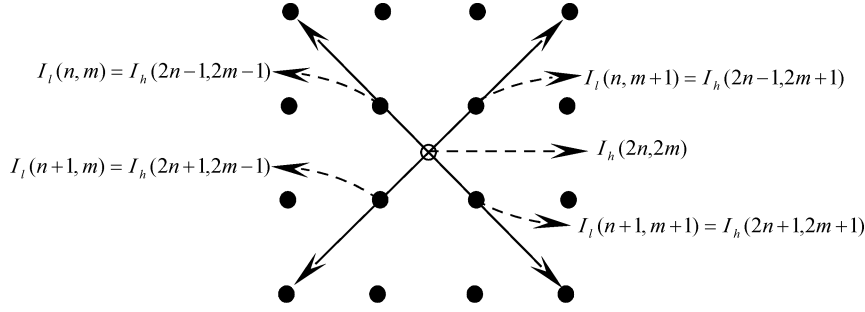


Fig. 2. Interpolation of the HR samples $I_h(2n, 2m)$. Two estimates of $I_h(2n, 2m)$ are made in the 45° and 135° directions as two noisy measurements of $I_h(2n, 2m)$.

To fuse the two directional measurements \hat{I}_{45} and \hat{I}_{135} into a more robust estimate, we rewrite (2-1) into matrix form

$$\mathbf{Y} = \mathbf{1} \cdot I_h + \mathbf{V} \quad (2-2)$$

where

$$\mathbf{Y} = \begin{bmatrix} \hat{I}_{45} \\ \hat{I}_{135} \end{bmatrix}, \quad \mathbf{1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \mathbf{V} = \begin{bmatrix} v_{45} \\ v_{135} \end{bmatrix}.$$

Now, the interpolation problem is to estimate the unknown sample I_h from the noisy observation \mathbf{Y} . This estimation can be optimized in minimum mean square-error sense. To obtain the minimum mean square-error estimation (MMSE) of I_h , i.e., $\hat{I}_h = E[I_h | \mathbf{Y}] = \int I_h p(I_h | \mathbf{Y}) dI_h$, we need to know the probability density function $p(I_h | \mathbf{Y})$. In practice, however, it is very hard to get this prior information or $p(I_h | \mathbf{Y})$ cannot be estimated at all. Thus, in real applications, LMMSE is often employed instead of MMSE. To implement LMMSE, only the first and second order statistics of I_h and \mathbf{Y} are needed, which may be estimated adaptively.

From (2-2), the LMMSE of I_h can be calculated as [18]

$$\hat{I}_h = \mu_h + \text{Cov}(I_h, \mathbf{Y}) (\text{Var}(\mathbf{Y}))^{-1} (\mathbf{Y} - E[\mathbf{Y}]) \quad (2-3)$$

where $\mu_h = E[I_h]$, $\text{Cov}(\mathbf{A}, \mathbf{B}) = E[(\mathbf{A} - E[\mathbf{A}])(\mathbf{B} - E[\mathbf{B}])^T]$ is the co-variance operator, and we abbreviate $\text{Cov}(\mathbf{A}, \mathbf{A})$ as $\text{Var}(\mathbf{A})$, the variance operator. Through the LMMSE operation, \hat{I}_h fuses the information provided by directional measurements \hat{I}_{45} and \hat{I}_{135} .

Let $\mu_{45}^v = E[v_{45}]$, $\mu_{135}^v = E[v_{135}]$. Through intensive experiments on 129 images, including outdoor and indoor images, portraits, MRI medical images, and SAR images, etc., we found that $\mu_{45}^v \approx 0$ and $\mu_{135}^v \approx 0$. Thus, noise vector \mathbf{V} can be considered to be zero mean. Denote by c_1 and c_2 the normalized correlation coefficients of v_{45} and v_{135} with I_h

$$c_1 = \frac{E[(v_{45} - \mu_{45}^v) \cdot (I_h - \mu_h)]}{\sqrt{E[(v_{45} - \mu_{45}^v)^2] E[(I_h - \mu_h)^2]}}$$

$$c_2 = \frac{E[(v_{135} - \mu_{135}^v) \cdot (I_h - \mu_h)]}{\sqrt{E[(v_{135} - \mu_{135}^v)^2] E[(I_h - \mu_h)^2]}}.$$

Our experiments also show that the values of c_1 and c_2 are very small. Thus, we consider v_{45} and v_{135} and, consequently, \mathbf{V} to be nearly uncorrelated with I_h . With the assumption that \mathbf{V} is

zero mean and uncorrelated with I_h , it can be derived from (2-3) that

$$\hat{I}_h = \mu_h + \sigma_h^2 \mathbf{1}^T (\mathbf{1} \cdot \sigma_h^2 + \mathbf{R}_V)^{-1} (\mathbf{Y} - \mathbf{1} \cdot \mu_h) \quad (2-4)$$

where $\sigma_h^2 = \text{Var}(I_h)$ and $\mathbf{R}_V = \text{Var}(\mathbf{V})$. To implement the above LMMSE scheme for I_h , parameters μ_h , σ_h^2 , and \mathbf{R}_V need to be estimated for each sample $I_h(2n, 2m)$ in a local window.

First, let us consider the estimation of μ_h and σ_h^2 . Again, referring to Fig. 2, the available LR samples around $I_h(2n, 2m)$ are used to estimate the mean and variance of $I_h(2n, 2m)$. Denote by \mathbf{W} a window that centers at $I_h(2n, 2m)$ and contains the LR samples in the neighborhood of $I_h(2n, 2m)$. For estimation accuracy, we should use a sufficiently large window as long as the statistics is stationary in \mathbf{W} . However, in a locality of edges, the image exhibits strong transient behavior. In this case, drawing samples from a large window will be counter-productive. To balance the conflicting requirements of sample size and sample consistency, we propose a Gaussian weighting in the sample window \mathbf{W} to account for the fact that the correlation between $I_h(2n, 2m)$ and its neighbors decays rapidly in the distance between them. The further an LR sample is from $I_h(2n, 2m)$, the less it should contribute to the mean value of $I_h(2n, 2m)$. We compute μ_h as

$$\mu_h = \sum_l \sum_k \mathbf{W}(l, k) \mathbf{G}(l, k) \quad (2-5)$$

where $\mathbf{G}(x, y) = (1/2\pi\zeta^2) \exp(-(x^2 + y^2)/2\zeta^2)$ is a 2-D Gaussian filter with scale ζ . The variance of $I_h(2n, 2m)$ is computed as

$$\sigma_h^2 = \sum_l \sum_k (\mathbf{W}(l, k) - \mu_h)^2 \mathbf{G}(l, k). \quad (2-6)$$

Next, we discuss the estimation of \mathbf{R}_V , the co-variance matrix of \mathbf{V} . Using (2-1) and the assumption that v_{45} and v_{135} are zero mean and uncorrelated with I_h , it can be easily derived that

$$\begin{cases} \text{Var}(v_{45}) = \text{Var}(\hat{I}_{45}) - \sigma_h^2 \\ \text{Var}(v_{135}) = \text{Var}(\hat{I}_{135}) - \sigma_h^2. \end{cases} \quad (2-7)$$

Since σ_h^2 has been estimated by (2-6), we need to estimate $\text{Var}(\hat{I}_{45})$ and $\text{Var}(\hat{I}_{135})$ in a local window to arrive at

$\text{Var}(v_{45})$ and $\text{Var}(v_{135})$. For this, we associate \hat{I}_{45} with a set of its neighbors in the 45° diagonal direction. Denote by \mathbf{Z}_{45} the vector that centers at $\hat{I}_{45}(2n, 2m)$

$$\mathbf{Z}_{45} = \left\{ \dots, \hat{I}_{45}(2n+2, 2m-2), I_l(n+1, m), \hat{I}_{45}(2n, 2m), I_l(n, m+1), \hat{I}_{45}(2n-2, 2m+2), \dots \right\}. \quad (2-8)$$

Set \mathbf{Z}_{45} encompasses \hat{I}_{45} and its neighbors, i.e., the original samples and the directional (45° diagonal) interpolated samples. Symmetrically, we define the sample set \mathbf{Z}_{135} for \hat{I}_{135} associated with interpolated results in the 135° diagonal

$$\mathbf{Z}_{135} = \left\{ \dots, \hat{I}_{135}(2n-2, 2m-2), I_l(n, m), \hat{I}_{135}(2n, 2m), I_l(n+1, m+1), \hat{I}_{135}(2n+2, 2m+2), \dots \right\}.$$

The estimates of $\text{Var}(\hat{I}_{45})$ and $\text{Var}(\hat{I}_{135})$ are computed as

$$\begin{aligned} \text{Var}(\hat{I}_{45}) &= \sum_k (\mathbf{Z}_{45}(k) - \mu_h)^2 g(k) \\ \text{and } \text{Var}(\hat{I}_{135}) &= \sum_k (\mathbf{Z}_{135}(k) - \mu_h)^2 g(k) \end{aligned} \quad (2-9)$$

where $g(x) = (1/\sqrt{2\pi}\xi) \exp(-x^2/2\xi^2)$ is a 1-D Gaussian filter with scale ξ .

Now, $\text{Var}(v_{45})$ and $\text{Var}(v_{135})$ can be computed by (2-8), and finally the co-variance matrix \mathbf{R}_V can be estimated as

$$\begin{aligned} \mathbf{R}_V &= \begin{bmatrix} \text{Var}(v_{45}) & c_3 \cdot \sqrt{\text{Var}(v_{45}) \cdot \text{Var}(v_{135})} \\ c_3 \cdot \sqrt{\text{Var}(v_{45}) \cdot \text{Var}(v_{135})} & \text{Var}(v_{135}) \end{bmatrix} \\ &\quad (2-10) \end{aligned}$$

where c_3 is the normalized correlation coefficient of v_{45} with v_{135}

$$c_3 = \frac{E[(v_{45} - \mu_{45}^v) \cdot (v_{135} - \mu_{135}^v)]}{\sqrt{E[(v_{45} - \mu_{45}^v)^2] E[(v_{135} - \mu_{135}^v)^2]}}.$$

Although v_{45} and v_{135} are nearly uncorrelated with I_h , they are somewhat correlated to each other because \hat{I}_{45} and \hat{I}_{135} have some similarities due to the high local correlation. We found that the values of c_3 are between 0.4 and 0.6 for most of the test images. The correlation between v_{45} and v_{135} varies, from relatively strong in smooth areas to weak in active areas. In the areas where sharp edges appear, which is the situation of our concern and interests, the values of c_3 are sufficiently low, and we can assume that v_{45} and v_{135} are uncorrelated with each other without materially affecting the performance of the proposed interpolation algorithm in practice. In practical implementation c_3 , the correlation coefficient between v_{45} and v_{135} , can be set as 0.5 or even 0 for most natural images. Our experiments reveal that the interpolation results are insensitive to c_3 . Varying c_3 from 0 to 0.6 hardly changes the PSNR value and visual quality of the interpolated image.

If a sharp edge presents in I_h in or near one of the two directions (the 45° diagonal or the 135° diagonals), the corresponding noise variances $\text{Var}(v_{45})$ and $\text{Var}(v_{135})$ will differ significantly from each other. By the adjustment of \mathbf{R}_V in (2-4),

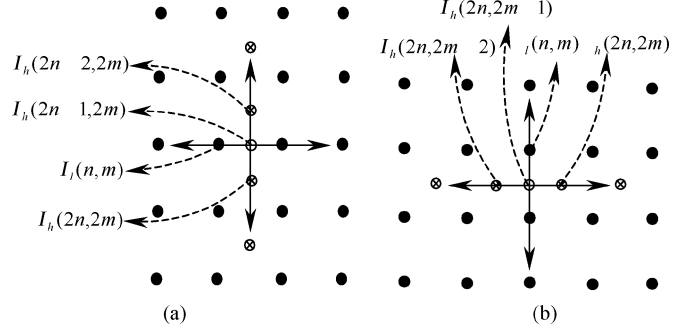


Fig. 3. Interpolations of the missing HR samples (a) $I_h(2n-1, 2m)$ and (b) $I_h(2n, 2m-1)$. The symbols “ \otimes ” represent the already recovered samples $I_h(2n, 2m)$. The two estimates of $I_h(2n-1, 2m)$ or $I_h(2n, 2m-1)$ are made in horizontal and vertical directions. $I_h(2n-1, 2m)$ and $I_h(2n, 2m-1)$ are estimated similarly to $I_h(2n, 2m)$.

the interpolation value \hat{I}_{45} or \hat{I}_{135} , whichever is in the direction perpendicular to the edge, will contribute far less to the final estimation result \hat{I}_h . The presented technique removes much of the ringing artifacts around the edges, which often appear in the interpolated images by cubic convolution and cubic spline interpolation methods.

B. Interpolation of Samples $I_h(2n-1, 2m)$ and $I_h(2n, 2m-1)$

After the missing HR samples $\hat{I}_h(2n, 2m)$ are estimated, the other missing samples $I_h(2n-1, 2m)$ and $I_h(2n, 2m-1)$ can be estimated similarly, but now with the aid of the just estimated HR samples. Referring to Fig. 3(a) and (b), the LR image pixels $I_l(n, m)$ are represented by black dots “ \bullet ,” while the estimated samples $\hat{I}_h(2n, 2m)$ by symbols “ \otimes .” The samples that are to be estimated are represented by white dots “ \circ .” As illustrated in Fig. 3, the missing sample $I_h(2n-1, 2m)$ or $I_h(2n, 2m-1)$ can be estimated in one direction by the original pixels of the LR image, and in the other direction by the already interpolated HR samples. Similar to (2-2), the two directional approximations of the missing sample are considered as the noisy measurements of $I_h(2n-1, 2m)$ or $I_h(2n, 2m-1)$, and then the LMMSE of the missing sample can be computed in a similar way as described in the previous section. Finally, the whole HR I_h is reconstructed by the proposed edge-guided LMMSE interpolation technique.

III. SIMPLIFIED LMMSE INTERPOLATION ALGORITHM

In interpolating the HR samples, the LMMSE technique of (2-4) needs to estimate μ_h , σ_h^2 , \mathbf{R}_V , and compute the inverse of a 2×2 matrix. This may amount to too heavy a computation burden for some applications that need high throughput. Specifically, if we set μ_h be the average of the four nearest LR neighbors of I_h to reduce computation, then computing μ_h needs three additions and one division and computing σ_h^2 needs seven additions, four multiplications, and one division. By setting the size of vector \mathbf{Z}_{45} and \mathbf{Z}_{135} as 5 and setting $\mathbf{R}_V = \text{diag}\{\text{Var}(v_{45}), \text{Var}(v_{135})\}$, i.e., $c_3 = 0$, in (2-10) to reduce the computational cost, we still need 20 additions and 20 multiplications to compute \mathbf{R}_V . The remaining operations in (2-4) include nine additions, eight multiplications, and one division. In total, the algorithm needs 39 additions, 32 multiplications, and three divisions to compute a \hat{I}_h with (2-4).

One way to reduce the computational complexity is to judiciously invoke the LMMSE algorithm only for pixels where high local activities are detected, and use a simple linear interpolation method in smooth regions. Since edge pixels represent minority in the total sample population, this will lead to significant savings in computations. In addition, we present a simplified version of the LMMSE-based interpolation algorithm while only slightly decreasing the performance.

We can see that the LMMSE estimate of HR sample I_h is actually a linear combination of \hat{I}_{45} , \hat{I}_{135} and μ_h . Referring to (2-4) and let $\Gamma = \sigma_h^2 \mathbf{1}^T (\mathbf{1} \cdot \sigma_h^2 \cdot \mathbf{1}^T + \mathbf{R}_V)^{-1}$, then Γ is a 2-D vector and we rewrite (2-4) as

$$\hat{I}_h = \Gamma_1 \cdot \hat{I}_{45} + \Gamma_2 \cdot \hat{I}_{135} + (1 - \Gamma \cdot \mathbf{1})\mu_h \quad (3-1)$$

where Γ_1 and Γ_2 are the first and second elements of Γ . We empirically observed that $(1 - \Gamma \cdot \mathbf{1})$ is close to zero, and, hence, μ_h has a light effect on \hat{I}_h . In this view, \hat{I}_h can be simplified to a weighted average of \hat{I}_{45} and \hat{I}_{135} , while the weights depend largely on the noise covariance matrix \mathbf{R}_V .

Instead of computing the LMMSE estimate of I_h , we determine an optimal pair of weights to make \hat{I}_h a good estimate of I_h . The strategy of weighted average leads to significant reduction in complexity over the exact LMMSE method. Let

$$\hat{I}_h = w_{45} \cdot \hat{I}_{45} + w_{135} \cdot \hat{I}_{135} \quad (3-2)$$

where $w_{45} + w_{135} = 1$. The weights w_{45} and w_{135} are determined to minimize the mean square-error (MSE) of \hat{I}_h : $\{w_{45}, w_{135}\} = \arg \min_{w_{45}+w_{135}=1} E[(\hat{I}_h - I_h)^2]$.

Although the measurement noises of \hat{I}_{45} and \hat{I}_{135} , i.e., v_{45} and v_{135} , are correlated to some extent, their correlation is sufficiently low to consider v_{45} and v_{135} as being approximately uncorrelated. This assumption holds better in the areas of edges that are critical to the human visual system and of interests to us. In fact, if v_{45} and v_{135} are highly correlated, that is to say, the two estimates \hat{I}_{45} and \hat{I}_{135} are close to each other, then \hat{I}_h varies little in w_{45} and w_{135} anyway. With the assumption that v_{45} and v_{135} are approximately uncorrelated, we can show that the optimal weights are

$$w_{45} = \frac{\text{Var}(v_{135})}{\text{Var}(v_{45}) + \text{Var}(v_{135})}, \quad w_{135} = 1 - w_{45}. \quad (3-3)$$

It is quite intuitive why the weighting scheme works. For instance, for an edge in or near the 135° diagonal direction, the variance $\text{Var}(v_{45})$ is higher than $\text{Var}(v_{135})$. From (3-3), we see that w_{45} will be less than w_{135} and, consequently, \hat{I}_{45} has less influence on \hat{I}_h than \hat{I}_{135} , and vice versa. To compute $\text{Var}(v_{45})$ and $\text{Var}(v_{135})$ as described in Section II, however, we still need 30 additions, 24 multiplications, and two divisions. In order to

TABLE I
OPERATIONS NEEDED FOR THE LMMSE ALGORITHM
AND THE SIMPLIFIED ALGORITHM

Operation	Addition	Multiplication	Division
LMMSE algorithm	39	32	3
Simplified algorithm	24	4	2

further simplify and speed up the computation of w_{45} and w_{135} , we use the following approximations:

$$\begin{aligned} \text{Var}(v_{45}) &\cong \left(\sum_{k=1}^5 |Z_{45}(k) - \mu_h| \right)^2 \\ \text{Var}(v_{135}) &\cong \left(\sum_{k=1}^5 |Z_{45}(k) - \mu_h| \right)^2 \end{aligned} \quad (3-4)$$

where “ \cong ” means nearly equivalent to. With the above simplification, we only need 23 additions, two multiplications, and two divisions to obtain w_{45} and w_{135} . Finally, with (3-2), we only need 24 additions, four multiplications, and two divisions to get \hat{I}_h . This yields a significant computational savings compared with (2-4), which needs 39 additions, 32 multiplications, and three divisions. Table I lists the operation counts of the simplified algorithm and the LMMSE algorithm.

As we will see in the next section, the simplified algorithm only has slightly lower PSNR result than the LMMSE algorithm, while its output images are virtually indistinguishable from those of the latter.

IV. EXPERIMENTAL RESULTS

The proposed image interpolation algorithms were implemented and tested, and their performance was compared with some existing methods. We downsampled some HR images to get the corresponding LR images, from which the original HR images were reconstructed by the proposed and competing methods. Since the original HR images are known in the simulation, we can compare the interpolated results with the true images, and measure the PSNR of those interpolated images.

The presented LMMSE-based interpolator was compared with the bicubic convolution interpolator, the bicubic spline interpolator, the subpixel edge detection-based interpolator of Jensen and Anastassiou [8], and the Wiener filtering-like interpolator of Li and Orchard [9]. To assess the sensitivity of the proposed interpolation algorithms to different initial directional estimates prior to fusing, they were tested when coupled with bicubic and bilinear convolution interpolators respectively. In the table and figure legends, the LMMSE method developed in Section II is labeled LMMSE_INTR_cubic or LMMSE_INTR_linear, depending on if bicubic or bilinear convolution interpolator is used to obtain initial directional estimates. Likewise, the simplified method in Section III is labeled OW_INTR_cubic (OW stands for optimal weighting) or OW_INTR_linear. In the experiments, we set the scale ζ of 2-D Gaussian filter \mathbf{G} [referring to (2-5)] around 1 and scale ξ of 1-D Gaussian filter \mathbf{g} [referring to (2-9)] around 1.5.

The PSNR results of the eight algorithms for the six test images are listed in Table II. The proposed LMMSE_INTR_cubic

TABLE II
PSNR (dB) RESULTS OF THE RECOVERED HR IMAGES BY THE EIGHT
ALGORITHMS, WHERE THE LR IMAGE IS DIRECTLY DOWN-SAMPLED
FROM THE HR IMAGE AS SHOWN IN FIG. 1

Image	Holes	Lena	Splash	Peppers	Cloth	Butterfly
Bicubic Convolution	23.26	28.95	33.22	28.84	28.92	27.28
Bicubic spline	22.91	28.78	33.17	28.63	28.71	27.23
Method in [8]	19.44	28.62	32.60	28.75	28.83	27.03
Method in [9]	22.17	29.39	33.31	29.21	29.24	27.60
LMMSE_INTR_cubic	25.18	29.28	33.90	29.28	29.30	28.11
LMMSE_INTR_linear	23.89	29.14	32.81	29.21	29.06	27.67
OW_INTR_cubic	24.55	29.19	33.70	29.15	29.13	27.85
OW_INTR_linear	23.37	29.10	33.17	29.15	28.93	27.35



Fig. 4. Interpolated image *Lena* by (a) LMMSE_INTR_cubic and (b) LMMSE_INTR_linear.

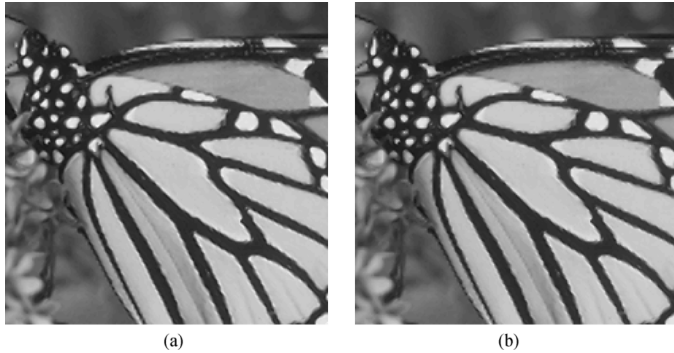


Fig. 5. Interpolated image *Butterfly* by (a) LMMSE_INTR_cubic and (b) LMMSE_INTR_linear.

algorithm achieves the highest PSNR results on all test images, except for image *Lena*, for which it has a slightly lower PSNR than the method in [9].

Our experimental results also call attention to a fact that the proposed methods are insensitive to their choice of the initial directional interpolators. Even with bilinear interpolation, which usually gets considerably worse result than bicubic interpolation, the final fused result is very close to that of bicubic interpolation, especially in terms of visual quality (the PSNR values are slightly lower as shown in Table II). Figs. 4 and 5 show the interpolated images *Lena* and *Butterfly* of the LMMSE_INTR_cubic and LMMSE_INTR_linear methods. In visual effects, the two methods are almost indistinguishable. This demonstrates the power of LMMSE-based data fusing strategy in correcting much of the interpolation errors of traditional linear methods.

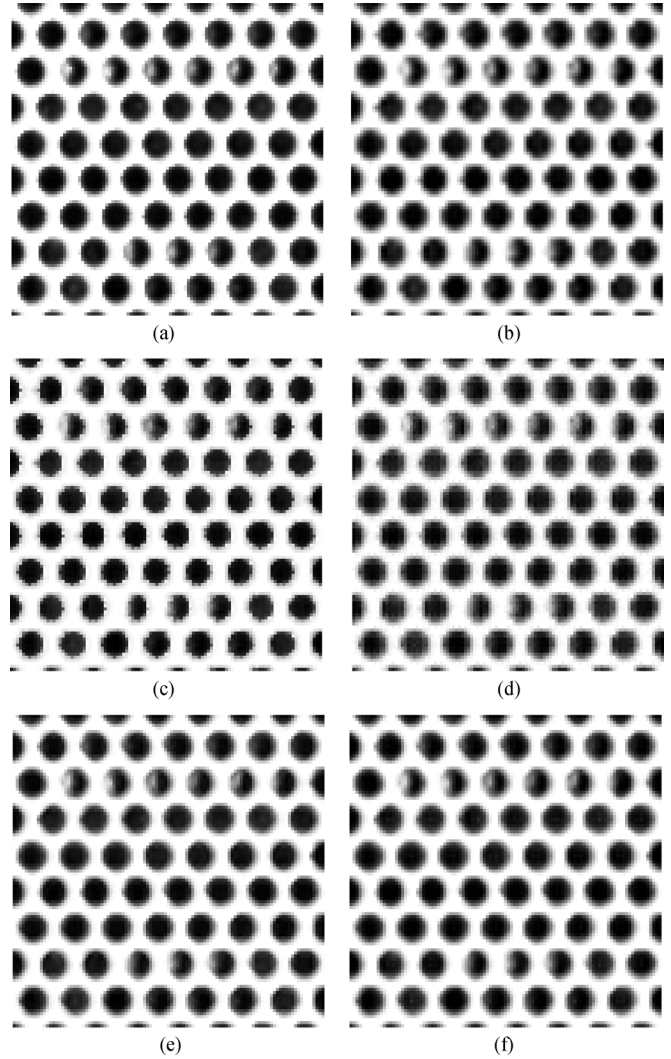


Fig. 6. Interpolation results of the image *Holes*. (a) Original image; interpolated image by (b) the cubic convolution; (c) the method in [8]; (d) the method in [9]; (e) the proposed LMMSE_INTR_cubic; (f) the proposed OW_INTR_cubic.

Fig. 6 presents the experimental results of a pattern image, *Holes*. Fig. 6(a) is the original image, and Fig. 6(b)–(f) is the interpolated images by the bicubic convolution, the method in [8], the method in [9], the LMMSE_INTR_cubic, and OW_INTR_cubic methods, respectively. The interpolated images of the bicubic spline interpolator are very similar to those of bicubic convolution interpolator and are not presented here to save space. The small holes in this image are good patterns to test the edge recovery ability of the interpolation algorithms. The LMMSE_INTR_cubic and OW_INTR_cubic methods reconstruct the original image very well. The fine round edge structures of the small holes are well preserved. The bicubic convolution method blurs the image and introduces many ringing effects around the perimeters of the holes. The artifacts introduced by the method in [8] are mainly due to the errors of edge detection carried out by the algorithm. The method in [9] also introduces artifacts around the hole's perimeters. This is because the algorithm uses a relatively large window to compute the covariance matrix for each missing sample,



Fig. 7. Interpolation results of the image *Lena*. (a) Original image, interpolated image by (b) the cubic convolution, (c) the method in [8], (d) the method in [9], (e) the proposed LMMSE_INTR_cubic, and (f) the proposed OW_INTR_cubic.

whereas the edge structure is small in scale in this test image, causing incorrect estimation of sample covariance.

In Figs. 7–11, we compare the visual quality of the test interpolation methods for natural images: *Lena*, *Splash*, *Pep-*

pers, *Cloth*, and *Butterfly*. The proposed methods remove many of the ringing and other visual artifacts of the other methods. The OW_INTR_cubic method is slightly inferior to the LMMSE_INTR_cubic method in reducing the ringing

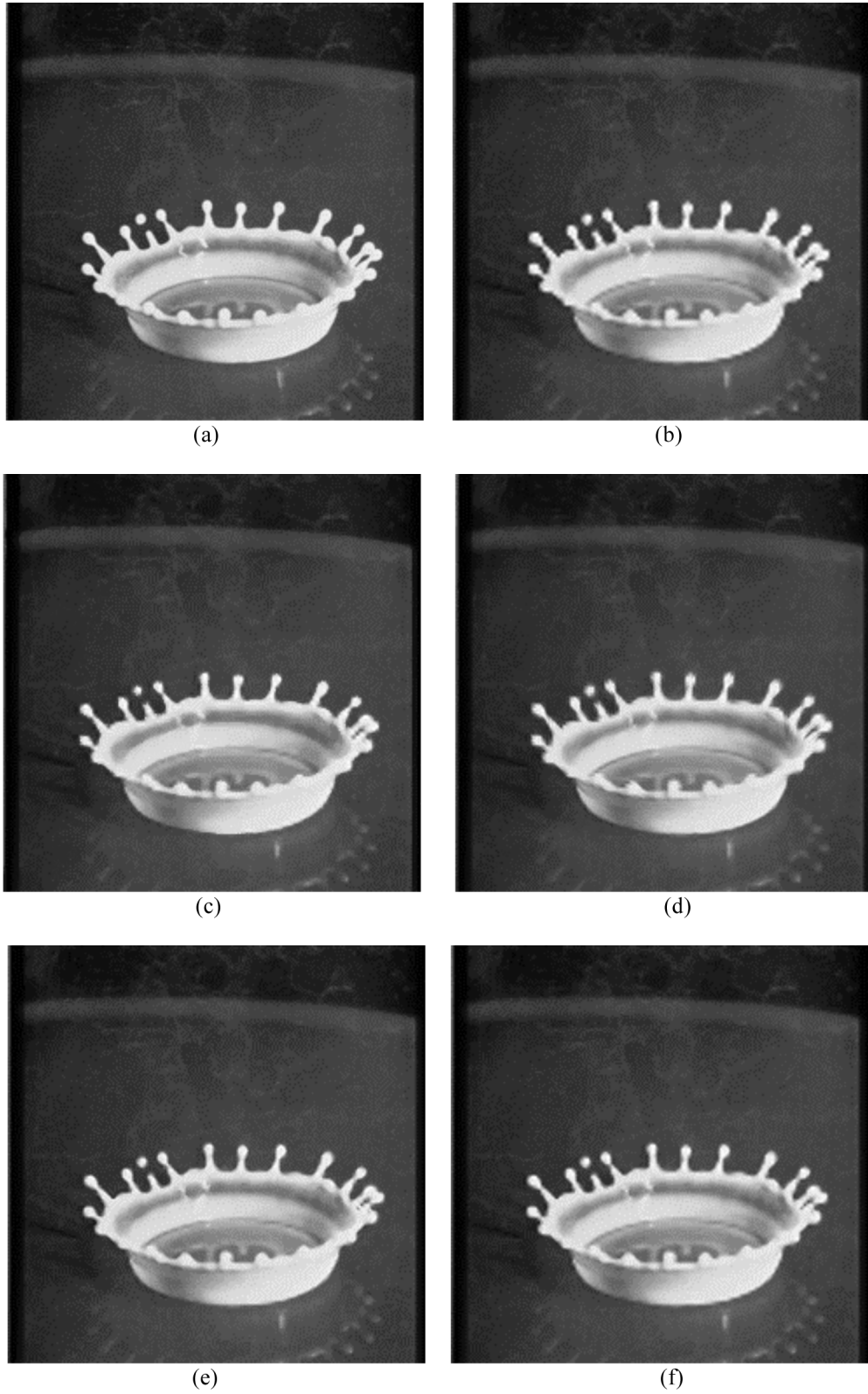


Fig. 8. Interpolation results of the image *Splash*. (a) Original image, interpolated image by (b) the cubic convolution, (c) the method in [8], (d) the method in [9], (e) the proposed LMMSE_INTR_cubic, and (f) the proposed OW_INTR_cubic.

effects, but this is a small price to pay for the computational savings of the former. The interpolator of Jensen and Anastassiou [8] can reproduce very thin edges in the object contour

because it contains a subpixel edge detection process, but it causes visible artifacts when the edge detector commits errors. This method leaves a considerable amount of ringing effects in

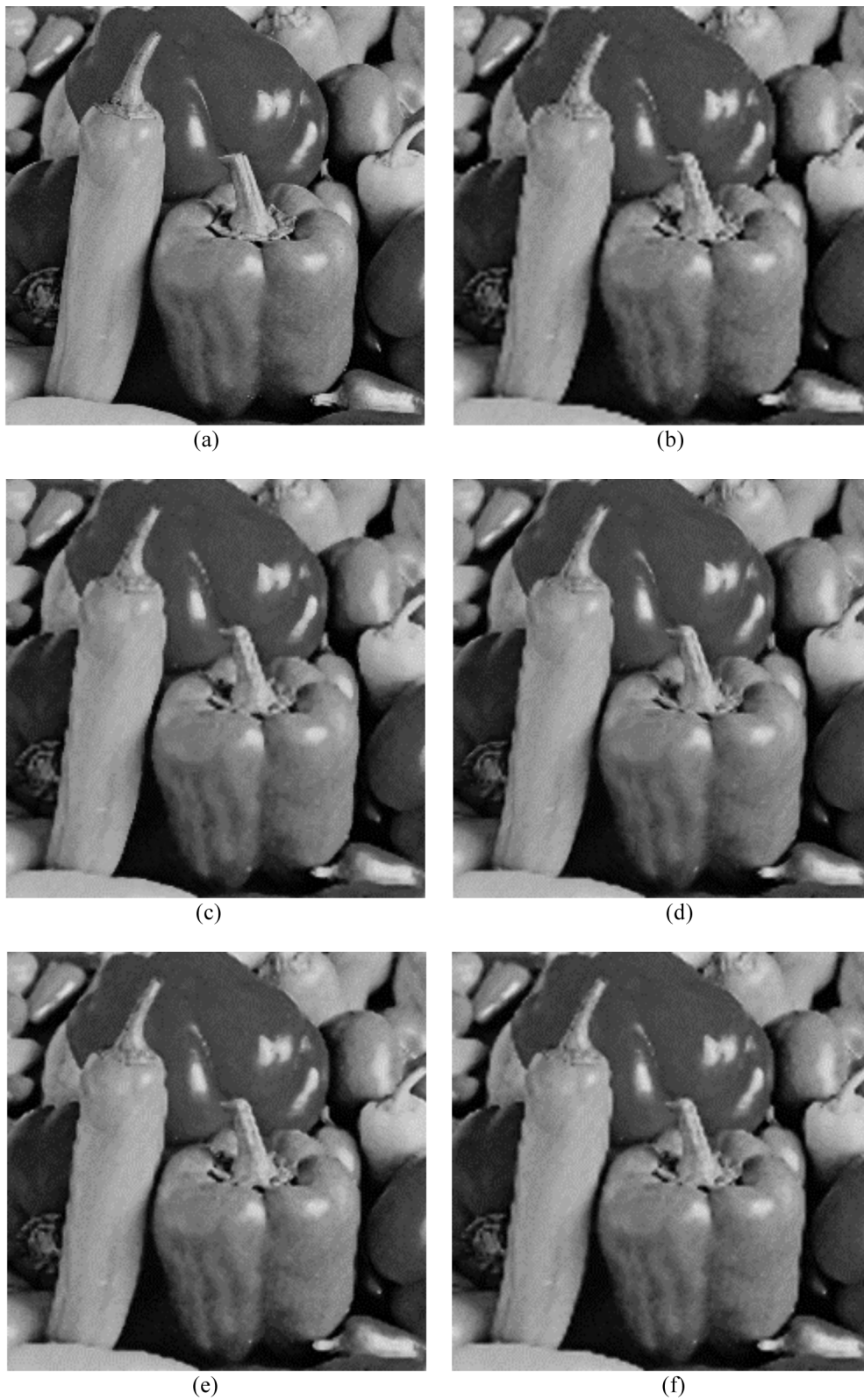


Fig. 9. Interpolation results of the image *Peppers*. (a) Original image, interpolated image by (b) the cubic convolution, (c) the method in [8], (d) the method in [9], (e) the proposed LMMSE_INTR_cubic, and (f) the proposed OW_INTR_cubic.

the hat of *Lena* and the wing of the *Butterfly*. The interpolator of Li and Orchard [9] can preserve large edge structures well, such as those in *Lena*; however, it introduces artifacts in the finer

edge structures, such as the drops of *Splash* and the head part of *Butterfly*. Another disadvantage of Li and Orchard's method is its high computational complexity. If an 8×8 window is

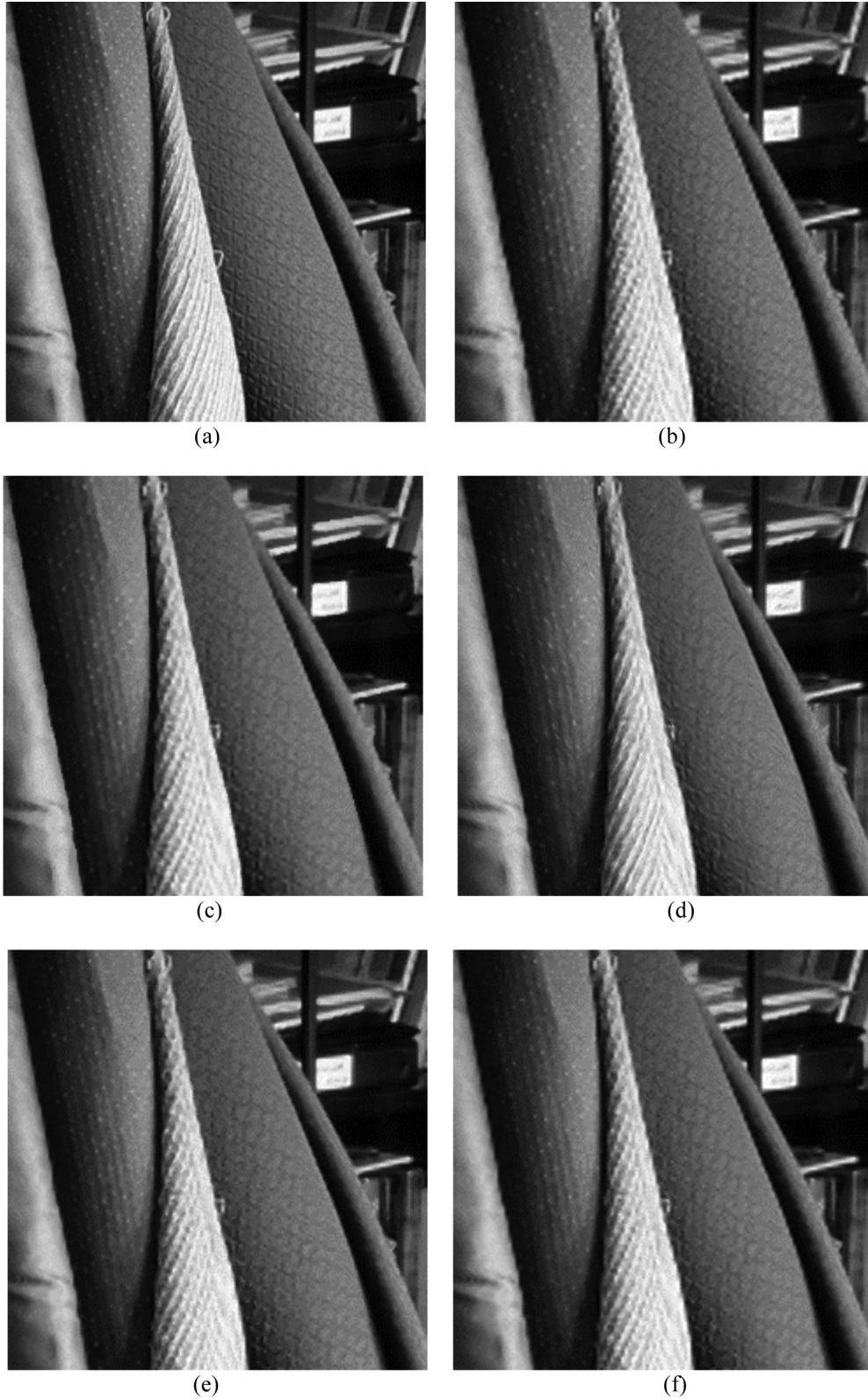


Fig. 10. Interpolation results of the image *Cloth*. (a) Original image, interpolated image by (b) the cubic convolution, (c) the method in [8], (d) the method in [9], (e) the proposed LMMSE_INTR_cubic, and (f) the proposed OW_INTR_cubic.

used to compute the covariance matrix, this algorithm requires about 1300 multiplications and thousands of additions. In comparison, the proposed LMMSE_INTR_cubic algorithm requires only tens of multiplications and divisions.

The downsampling process considered in this paper, through which an LR image is generated from the corresponding HR image, is ideal Dirac sampling. An alternative model of LR images is that of low-pass filtering followed by downsampling.

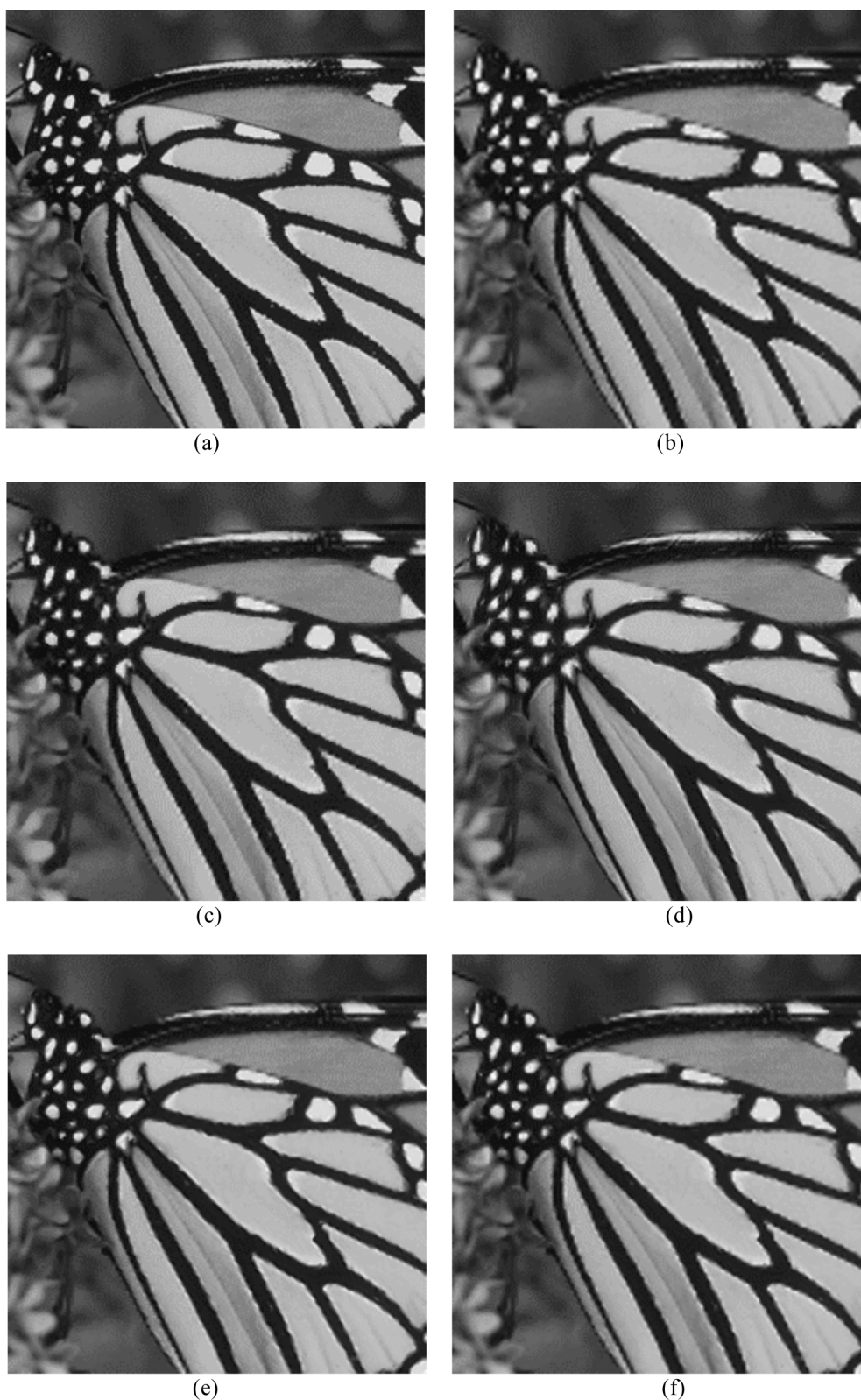


Fig. 11. Interpolation results of the image *Butterfly*. (a) Original image, interpolated image by (b) the cubic convolution, (c) the method in [8], (d) the method in [9], (e) the proposed LMMSE_INTR_cubic, and (f) the proposed OW_INTR_cubic.

In our experiments we also simulated the latter model and created LR test images by downsampling of low-pass filtered HR images (averaging of four pixels in a 2×2 window). The resulting LR images were processed by all the interpo-

lation methods of our comparison group. Table III lists the PSNR results. Fig. 12(a) shows the original *Butterfly* image. Fig. 12(b)–(f) shows the interpolation results of image *Butterfly* by different algorithms. Even though the new interpolation

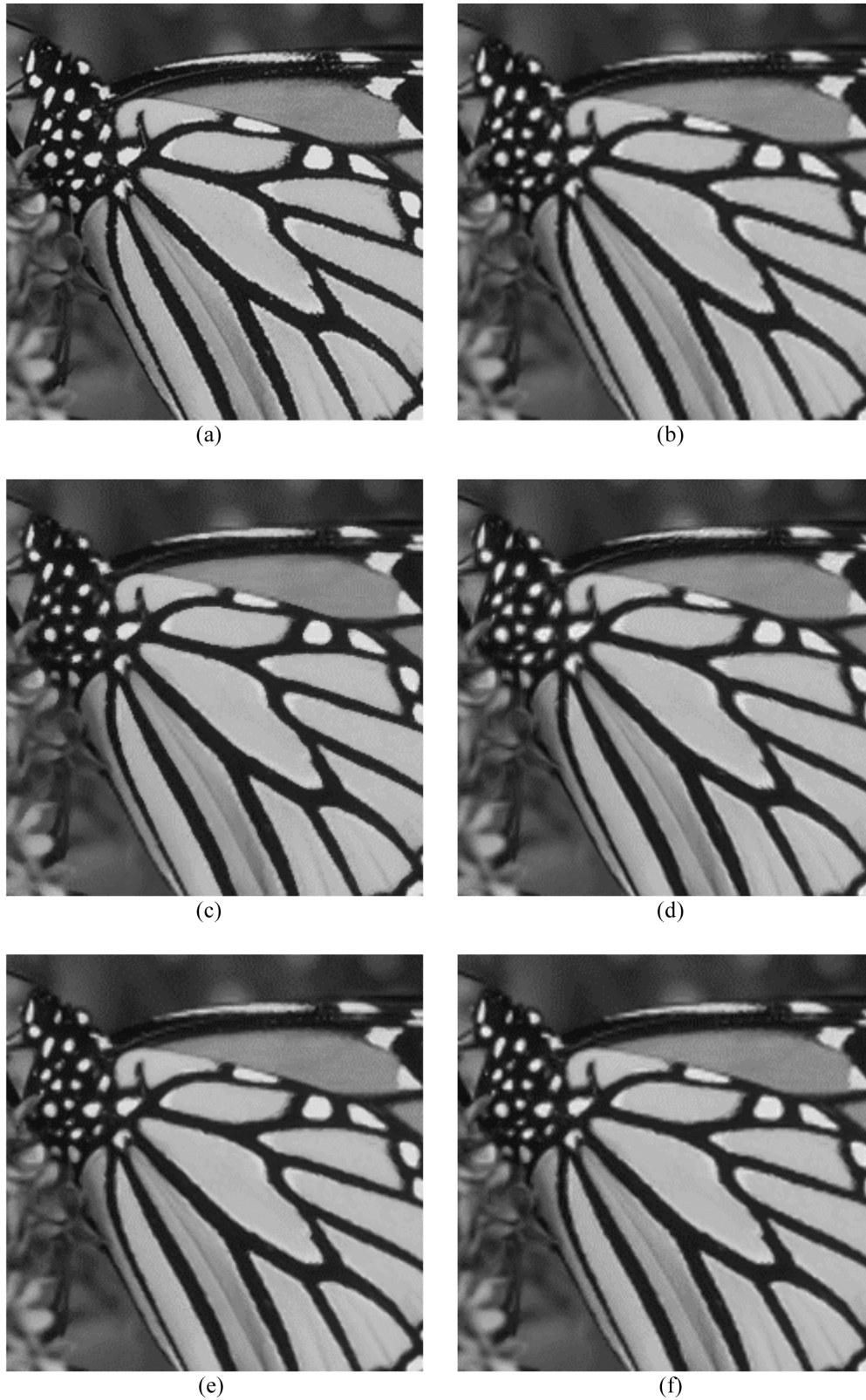


Fig. 12. Interpolation results of the image *Butterfly* when the LR image is generated by low-pass filtering and downsampling the HR image. (a) Original image, (b) the cubic convolution, (c) the method in [8], (d) the method in [9], (e) the proposed LMMSE_INTR_cubic, and (f) the proposed OW_INTR_cubic.

technique is not tailored to the alternative model of LR images, like others in the comparison group of this paper, it appears to perform well nonetheless. The new interpolation algorithms

reproduced much sharper edges than the bicubic convolution or bicubic spline methods, while being competitive against the methods of [8] and [9].

TABLE III
PSNR (dB) RESULTS OF THE RECOVERED HR IMAGES BY THE EIGHT
ALGORITHMS, WHERE THE LR IMAGE IS GENERATED BY LOW-PASS
FILTERING AND DOWNSAMPLING THE HR IMAGE

Image	Holes	Lena	Splash	Peppers	Cloth	Butterfly
Bicubic Convolution	17.79	26.38	29.88	26.89	26.57	24.59
Bicubic spline	17.81	26.41	29.97	26.93	26.63	24.73
Method in [8]	16.16	25.96	28.82	26.46	26.11	24.31
Method in [9]	17.54	26.21	29.52	26.76	26.44	24.57
LMMSE_INTR_cubic	18.01	26.35	29.92	26.87	26.57	24.82
LMMSE_INTR_linear	17.73	26.22	29.71	26.74	26.41	24.57
OW_INTR_cubic	17.94	26.30	29.92	26.85	26.56	24.72
OW_INTR_linear	17.64	26.21	29.67	26.72	26.36	24.42

V. CONCLUSION

We developed an edge-guided LMMSE-type image interpolation technique. For each pixel to be interpolated, we partitioned its neighborhood into two observation subsets in two orthogonal directions. Each observation subset was used to generate an estimate of the missing sample. These two directional estimates were processed as two noisy measurements of the missing sample. Using and combining the statistics of the two observation subsets, we fused the two noisy measurements into a more robust estimate via linear minimum mean square-error estimation. To reduce the computational complexity of the proposed method, we simplified it to an optimal weighting problem and determined the optimal weights. The simplified method had competitive performance with significant computational savings. Experimental results showed that the presented methods avoided interpolation against edge directions and, hence, achieved noticeable reduction in ringing and other visual artifacts.

REFERENCES

- [1] H. S. Hou, "Cubic splines for image interpolation and digital filtering," *IEEE Trans. Acoustic, Speech, Signal Process.*, vol. ASSP-26, no. 6, pp. 508–517, Dec. 1978.
- [2] R. G. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Trans. Acoustic, Speech, Signal Process.*, vol. ASSP-29, no. 6, pp. 1153–1160, Dec. 1981.
- [3] T. M. Lehmann, C. Gönner, and K. Spitzer, "Survey: Interpolation methods in medical image processing," *IEEE Trans. Med. Imag.*, vol. 18, no. 11, pp. 1049–1075, Nov. 1999.
- [4] E. Meijering, "A chronology of interpolation: From ancient astronomy to modern signal and image processing," *Proc. IEEE*, vol. 90, no. 3, pp. 319–342, Mar. 2002.
- [5] M. Unser, "Splines: A perfect fit for signal and image processing," *IEEE Signal Process. Mag.*, no. 11, pp. 22–38, Nov. 1999.
- [6] M. Unser, A. Aldroubi, and M. Eden, "Enlargement or reduction of digital images with minimum loss of information," *IEEE Trans. Image Process.*, vol. 4, no. 3, pp. 247–258, Mar. 1995.

- [7] B. Vrcelj and P. P. Vaidyanathan, "Efficient implementation of all-digital interpolation," *IEEE Trans. Image Process.*, vol. 10, no. 11, pp. 1639–1646, Nov. 2001.
- [8] K. Jensen and D. Anastassiou, "Subpixel edge localization and the interpolation of still images," *IEEE Trans. Image Process.*, vol. 4, no. 3, pp. 285–295, Mar. 1995.
- [9] X. Li and M. T. Orchard, "New edge-directed interpolation," *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1521–1527, Oct. 2001.
- [10] S. Carrato and L. Tenze, "A high quality $2 \times$ image interpolator," *IEEE Signal Process. Lett.*, vol. 7, no. 6, pp. 132–135, Jun. 2000.
- [11] Y. Takahashi and A. Taguchi, "An enlargement method of digital images with the prediction of high-frequency components," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing*, 2002, vol. 4, pp. 3700–3703.
- [12] W. K. Carey, D. B. Chuang, and S. S. Hemami, "Regularity-Preserving image interpolation," *IEEE Trans. Image Process.*, vol. 8, no. 9, pp. 1293–1297, Sep. 1999.
- [13] Y. Zhu, S. C. Schwartz, and M. T. Orchard, "Wavelet domain image interpolation via statistical estimation," in *Proc. Int. Conf. Image Processing*, 2001, vol. 3, pp. 840–843.
- [14] D. D. Muresan and T. W. Parks, "Prediction of image detail," in *Proc. Int. Conf. Image Processing*, 2000, vol. 2, pp. 323–326.
- [15] D. D. Muresan, "Fast edge directed polynomial interpolation," in *Proc. Int. Conf. Image Processing*, 2005, vol. 2, pp. 990–993.
- [16] F. Malgouyres and F. Guichard, "Edge direction preserving image zooming: A mathematical and numerical analysis," *SIAM J. Numer. Anal.*, vol. 39, pp. 1–37, 2001.
- [17] S. Mallat, *A Wavelet Tour of Signal Processing*. New York: Academic, 1999.
- [18] E. W. Karmen and J. K. Su, *Introduction to Optimal Estimation*. London, U.K.: Springer-Verlag, 1999.
- [19] W. K. Pratt, *Digital Image Processing*, 2nd ed. New York: Wiley, 1991.
- [20] I. Pitas, *Digital Image Processing Algorithms and Applications*. New York: Wiley, 2000.



Lei Zhang (M'04) was born in China in 1974. He received the B.S. degree from the Shenyang Institute of Aeronautical Engineering, Shenyang, China, in 1995, and the M.S. and Ph.D. degrees in electrical and engineering from Northwestern Polytechnical University, Xi'an, China, in 1998 and 2001, respectively.

From 2001 to 2002, he was a Research Associate with the Department of Computing, The Hong Kong Polytechnic University (HKPU), Kowloon. From 2003 to 2005, he was a Postdoctoral Fellow with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada. In 2006, he became an Assistant Professor with the Department of Computing, HKPU. His research interests include image and video processing, biometrics, pattern recognition, multisensor data fusion and optimal estimation theory, etc.



Xiaolin Wu (M'88–SM'96) received the B.Sc. degree in computer science from Wuhan University, Wuhan, China, and the Ph.D. degree in computer science from the University of Calgary, Calgary, AB, Canada, in 1982 and 1988, respectively.

He is currently a Professor in the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada, and a Research Professor of computer science at the Polytechnic University, Brooklyn, NY, and holds the NSERC-DALSA Research Chair in digital cinema. His research interests include image processing, multimedia coding and communications, data compression, and signal quantization, and he has published over 100 research papers in these fields.