Image Interpolation by Adaptive 2-D Autoregressive Modeling and Soft-Decision Estimation

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

***Abstract*—The challenge of image interpolation is to preserve spatial details. We propose a soft-decision interpolation technique that estimates missing pixels in groups rather than one at a time. The new technique learns and adapts to varying scene structures using a 2-D piecewise autoregressive model. The model parameters are estimated in a moving window in the input low-resolution image. The pixel structure dictated by the learnt model is enforced by the soft-decision estimation process onto a block of pixels, including both observed and estimated. The result is equivalent to that of a high-order adaptive nonseparable 2-D interpolation filter. This new image interpolation approach preserves spatial coherence of interpolated images better than the existing methods, and it produces the best results so far over a wide range of scenes in both PSNR measure and subjective visual quality. Edges and textures are well preserved, and common interpolation artifacts (blurring, ringing, jaggies, zippering, etc.) are greatly reduced.**

***Index Terms*—Autoregressive process, image interpolation, image modeling, optimization, soft decision estimation.**

I. Introduction

ONE of the most important quality metrics of digital images is spatial resolution. Despite steady increase of native sensor resolutions of digital cameras and scanners, new applications will always emerge that demand even higher spatial resolution. Image interpolation is an algorithmic means to increase the native resolution of an input image. An obvious application of image interpolation is the reproduction of images captured by digital cameras for high quality prints in magazines, catalogs, wall posters, or even home use. Another important application is upconversion of standard-definition video frames for playback on high-definition television receivers and computer monitors. Besides consumer electronics, image interpolation is beneficial and in some cases even necessary in computer vision, surveillance, medical imaging, remote sensing, and other fields.

Many image interpolation techniques of different tradeoffs between computational complexity and reproduction quality were developed. Popular methods, as commonly used in image/video software and hardware products, are bilinear interpolation, cubic convolution interpolation [1] and cubic spline interpolation [2]. The main advantage of these methods is their relatively low complexity. Their common drawback is the inability to adapt to varying pixel structures in a scene, due to the use of scene-independent interpolators. As a result, they are all susceptible to defects such as jaggies, blurring, and ringing.

With ever-increasing computation power in image and video processing, more sophisticated adaptive image interpolation methods were proposed in recent years. Many researchers advocated the approach of edge-guided interpolation. Jensen and Anastassiou published a scheme that detects edges and fits them with some templates to improve the visual perception of interpolated images [3]. Carrato and Tenze used some predetermined edge patterns to optimize the parameters in the interpolation operator [4]. To preserve edge structures in interpolation, Li and Orchard proposed to estimate the covariance of high-resolution (HR) image from the covariance of the low-resolution (LR) image, and then interpolate the missing pixels based on the estimated covariance [5]. This edge-directed interpolation work was cast by Muresan and Parks into the framework of adaptive optimal recovery [6]. Alternatively, Zhang and Wu proposed to interpolate a missing pixel in multiple directions, and then fuse the directional interpolation results by minimum mean square-error estimation [7].

Wavelets were also used in image interpolation. The idea is to exploit the statistical similarity between different scales of a wavelet-decomposed image. The interpolation is done by predicting the HR details from the LR observation [8]–[10]. Wu and Zhang studied image interpolation in a framework of pattern classification [11]. In [12], Malgouyres and Guichard analyzed some linear and nonlinear image enlargement methods theoretically and experimentally.

The reproduction quality of any image interpolation algorithm primarily depends on its adaptability to varying pixel structures across an image. In fact, modeling of nonstationarity of image signals is a common challenge facing many image processing tasks, such as compression, restoration, denoising, and enhancement. We had a measured success in this regard in a research on predictive lossless image compression [13]. In that work, a natural image is modeled as a piecewise 2-D autoregressive process. The model parameters are estimated on the fly for each pixel using sample statistics of a local window, assuming that the image is piecewise stationary. In this paper, we extend this approach to image interpolation. An obvious difference is in that the sample set for parameter estimation has to be causal to the current pixel for predictive coding, but does not need to be so for interpolation, which is to the advantage of the latter task. On the other hand, for image interpolation

Manuscript received September 1, 2007; revised March 13, 2008. This work was supported by the Natural Sciences and Engineering Research Council of Canada. This paper was presented in part at the 2007 IEEE International Conference on Image Processing. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Vicent Caselles.

The authors are with the Department of Electrical and Computer Engineering McMaster University Hamilton, Ontario, ON L8S 4K1 Canada (e-mail: zhangxj@grads.ece.mcmaster.ca; zhangx26@mcmaster.ca; xwu@ece.mcmaster.ca).

Digital Object Identifier 10.1109/TIP.2008.924279

the fit of the model to the true HR image signal is made more difficult by the fact that only a LR version of the original can be observed.

All image interpolation methods involve fitting missing pixels to some sample structure learnt from the LR image. This is best accomplished by estimating a block of missing pixels in relation to the nearby known pixels rather than estimating the individual missing pixels in isolation as done up to now. The main contribution of this paper is a new image interpolation technique, called soft-decision adaptive interpolation (SAI). The SAI technique, via a natural integration of piecewise 2-D autoregressive modeling and block estimation, achieves superior image interpolation results to those reported in the literature. It is shown that the new SAI technique is equivalent to interpolation using an adaptive nonseparable 2-D filter of high order.

The rest of the paper is structured as follows. Section II defines a 2-D piecewise autoregressive (PAR) image model to facilitate the subsequent development. Section III presents the most important result of this paper: a soft-decision estimation technique for adaptive image interpolation (SAI). Section IV discusses how to estimate the PAR model parameters in the LR image. Implementation details of the SAI algorithm are presented in Section V, where the reader can also gain an insight into the inner work of the SAI technique. Experimental results and a comparison study with some existing popular image interpolation techniques are presented in Section VI. Section VII concludes.

II. Piecewise Stationary Autoregressive Model

For the purpose of adaptive image interpolation, we model the image as a piecewise autoregressive (PAR) process

$$X(i,j) = \sum_{(m,n)\in T} \alpha(m,n) X(i+m, j+n) + \upsilon_{i,j} \qquad (1)$$

where T is a spatial template for the regression operation. The term $\upsilon_{i,j}$ is a random perturbation independent of spatial location (i,j) and the image signal, and it accounts for both fractal-like fine details of image signal and measurement noise. The validity of the PAR model hinges on a mechanism that adjusts the model parameters $\alpha(m,n)$ to local pixel structures. The fact that semantically meaningful image constructs, such as edges and surface textures, are formed by spatially coherent contiguous pixels, suggests piecewise statistical stationarity of the image signal. In other words, in the setting of the PAR model, the parameters $\alpha(m,n)$ remain constant or near constant in a small locality, although they may and often do vary significantly in different segments of a scene. The piecewise stationarity makes it possible to learn pixel structures such as edges and textures by fitting samples of a local window to the PAR model.

The validity of the PAR model with locally adaptive parameters is corroborated by the success of this modeling technique in lossless image compression. Among all known lossless image coding methods, including CALIC [14], TMW [15], and invertible integer wavelets [16], those that employ the PAR model with adjusted parameters on a pixel-by-pixel basis have delivered the lowest lossless bit rates [13], [17]. In the principle of Kolmogorov complexity, the true model of a stochastic process is the one that yields the minimum description length. Thus, we have strong empirical evidence to support the appropriateness and usefulness of the PAR model for natural images.

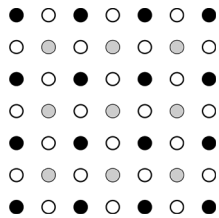


Fig. 1. Formation of a LR image from a HR image by down-sampling. The solid dots are the LR image pixels and the circles are the missing HR pixels. Interpolation is done in two passes. The first pass interpolates the missing pixels marked by shaded circles, and the second pass interpolates the remaining missing pixels marked by empty circles.

In Section III, we will integrate the PAR model into a soft-decision estimation framework for the purpose of image interpolation, and develop the SAI algorithm.

III. Adaptive Interpolation With Soft Decision

First we introduce some notations that are necessary for the description of the SAI algorithm. Let I_h be the HR image to be estimated by interpolating the LR image I_l observed. The LR image I_l is a down sampled version of the HR image I_h by a factor of two as illustrated by Fig. 1. Let $x_i \in I_l$ and $y_i \in I_h$ be the pixels of images I_l and I_h, respectively. We write the neighbors of pixel location i in the HR image as $y_{i\diamond t}$, $t = 1, 2, \ldots$. Since $x_i \in I_l$ implies $x_i \in I_h$, we also write an HR pixel $y_i \in I_h$ as x_i (likewise, $y_{i\diamond t}$ as $x_{i\diamond t}$) when it is in the LR image, $y_i \in I_l$, as well.

The SAI algorithm interpolates the missing pixels in I_h in two passes in a coarse to fine progression. The work of the two passes is shown by Fig. 1, in which the solid dots are known LR pixels, the shaded dots are those missing pixels to be interpolated in the first pass, and the empty dots are the remaining missing pixels to be interpolated in the second pass. The pixels generated by the first pass and the known LR pixels form a quincunx sublattice of the HR image (the union of solid and shaded dots). The second pass completes the reconstruction of the HR image by interpolating the other quincunx sublattice of empty dots.

Fig. 2(a) illustrates the spatial configuration of known and missing pixels involved in the first pass. To avoid intricate notations, from now on we use a single position index to denote a pixel location instead of 2-D coordinates. For a missing pixel $y_i \in I_h$, its four 8-connected neighbors are available LR pixels, denoted by $x^{(8)}_{i\diamond t} \in I_l$, and its four 4-connected neighbors are missing HR pixels, denoted by $y^{(4)}_{i\diamond t} \in I_h, t = 1, 2, 3, 4$. Here the relative subscript t is a generic notation to index a 2-D neighbor with respect to position i. Similarly, for a pixel $x_i \in I_l$, its four 8-connected neighbors are missing HR pixels, written as $y^{(8)}_{i\diamond t} \in I_h$, and its four 4-connected neighbors are available LR pixels written as $x^{(4)}_{i\diamond t} \in I_l$, $t = 1, 3, 4$.

To interpolate a missing pixel $y_i \in I_h$ in the first pass, we use a PAR model of parameters $\boldsymbol{a} = (a_1, a_2, a_3, a_4)$ to characterize the diagonal correlations of the image signal in a local

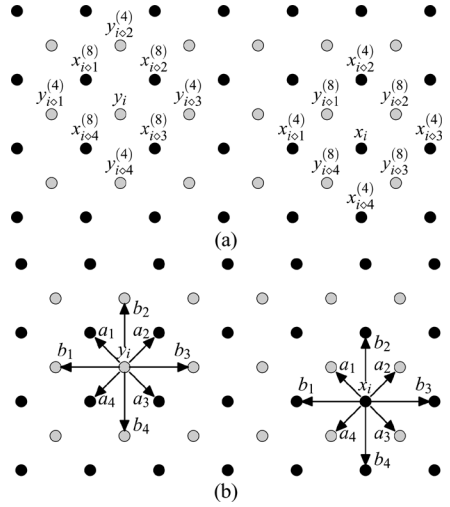


Fig. 2. (a) Spatial configuration in the first pass. (b) PAR model parameters $\boldsymbol{a} = (a_1, a_2, a_3, a_4)$ and $\boldsymbol{b} = (b_1, b_2, b_3, b_4)$ in relationship to spatial correlations of pixels.

window W [see Fig. 2(b)]. Using the simplified notations to replace $\alpha(\cdot,\cdot)$, we rewrite (1) as

$$y_i = \sum_{1\le t\le 4} a_t x_{i\diamond t}^{(8)} + \upsilon_i. \tag{2}$$

With the PAR model, we interpolate n missing pixels $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$ in window W by a least-squares block estimation

$$\hat{\boldsymbol{y}} = \arg\min_{\boldsymbol{y}} \left\{ \sum_{i\in W} \left\| y_i - \sum_{1\le t\le 4} a_t x_{i\diamond t}^{(8)} \right\| + \sum_{i\in W} \left\| x_i - \sum_{1\le t\le 4} a_t y_{i\diamond t}^{(8)} \right\| \right\}. \tag{3}$$

The above image interpolation approach has an important distinction from its predecessors (e.g., [2] and [5]). Existing image interpolation methods estimate each missing pixel independently from others, which we characterize as hard-decision estimation. In contrast, we adopt a strategy of soft-decision estimation in resemblance to block decoding of error correction codes. Rather than estimating one sample at a time in isolation, the objective function of (3) requires all missing pixels in a local window W to be estimated jointly. Moreover, the soft-decision estimation approach brings in a new feedback mechanism that is the second term in (3). This additional term requires the estimates of the missing HR pixels $\boldsymbol{y} \in I_h$ to fit the known LR pixels $\boldsymbol{x} \in I_l$ with the very same PAR model that fits $\boldsymbol{x} \in I_l$ to $\boldsymbol{y} \in I_h$. Aided by the feedback mechanism that accounts for mutual influences between the estimates of the missing pixels in a local window W, the SAI algorithm can mitigate errors of hard-decision estimation by preventing the PAR model, when applied to estimated HR pixels $y \in I_h$, from being violated on neighboring known LR pixels $x \in I_l$.

To include horizontal and vertical correlations into the SAI algorithm, we introduce four more parameters $\boldsymbol{b} = (b_1, b_2, b_3, b_4)$ whose geometric meanings are shown in Fig. 2(b). These parameters $\boldsymbol{b}$ are used to impose the same directional correlation between LR pixels x_i and $x_{i\diamond t}^{(4)}$ on between HR pixels y_i and $y_{i\diamond t}^{(4)}$, namely

$$y_i = \sum_{1\le t\le 4} \left\| b_t y_{i\diamond t}^{(4)} \right\| + \upsilon_i. \tag{4}$$

The soft-decision estimation technique can incorporate (4) into (3). However, one should practice caution since the pixels y_i and $y_{i\diamond t}^{(4)}$ in (4) are all unknown. By using a Lagrangian multiplier λ to regulate the contribution of (4), we extend (3) to the following constrained optimal block estimation problem [see (5), shown at the bottom of the page]. In minimizing $J(\lambda)$, the value of λ is chosen such that $\sum_W \left\| y_i - \sum_{1\le t\le 4} b_t y_{i\diamond t}^{(4)} \right\| \approx \sum_W \left\| x_i - \sum_{1\le t\le 4} b_t x_{i\diamond t}^{(4)} \right\|$. The SAI algorithm iterates on λ until the constraint is satisfactorily met, by decreasing λ if the left side of the constraint is less than the right side and vise versa. This constraint holds if the sample statistics is shift invariant in the window W. We observe that the value of λ is in the range of $0.2 \sim 0.7$ when meeting the constraint. For most natural images, one can simply choose $\lambda = 0.5$ with no material loss of performance compared with the iteratively computed λ.

Compared to existing autoregressive methods that use parameters $\boldsymbol{a}$ only [5], [7], the SAI algorithm expands the model parameter space by using two sets of parameters $\boldsymbol{b}$ and $\boldsymbol{a}$. The expanded PAR model has the potential of representing the HR image more accurately than in [5], [7]. However, to circumvent the risk of data overfitting, we do not directly use an autoregressive model of order 8, but rather split model parameters $\boldsymbol{a}$ and $\boldsymbol{b}$ in two separate terms of the objective function (5). In fact, in separation from $\boldsymbol{a}$, the parameters $\boldsymbol{b}$ can be better estimated than parameters $\boldsymbol{a}$ using samples in the LR image, as we will see in Section IV.

$$J(\lambda) = \min_{\boldsymbol{y}} \left\{ \sum_{i\in W} \left\| y_i - \sum_{1\le t\le 4} a_t x_{i\diamond t}^{(8)} \right\| + \sum_{i\in W} \left\| x_i - \sum_{1\le t\le 4} a_t y_{i\diamond t}^{(8)} \right\| + \lambda \sum_{i\in W} \left\| y_i - \sum_{1\le t\le 4} b_t y_{i\diamond t}^{(4)} \right\| \right\}$$
$$\text{subject to } \sum_{i\in W} \left\| y_i - \sum_{1\le t\le 4} b_t y_{i\diamond t}^{(4)} \right\| \approx \sum_{i\in W} \left\| x_i - \sum_{1\le t\le 4} b_t x_{i\diamond t}^{(4)} \right\| \tag{5}$$

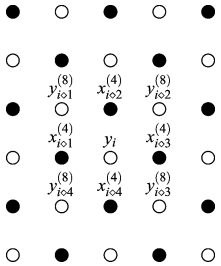


Fig. 3. Spatial configuration in the second pass of interpolation.

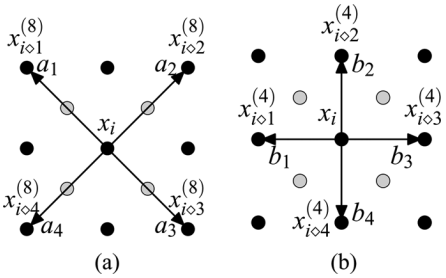


Fig. 4. Sample relations in estimating model parameters. (a) Parameter $\boldsymbol{a}$, (b) parameter $\boldsymbol{b}$.

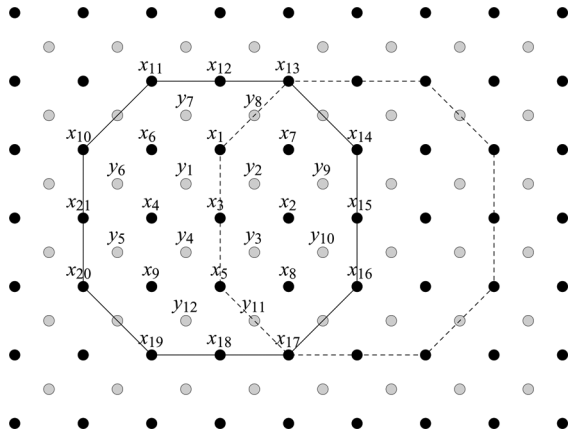


Fig. 5. Possible configuration used in the soft-decision interpolation algorithm.

With the block-based soft-decision estimation and the increased order in piecewise autoregressive modeling, the new SAI algorithm achieves unprecedented interpolation accuracy. More importantly, it performs consistently well over a wide range of images, and the performance is far less sensitive to feature scales than existing techniques. We will return to these points in Section VI when the experimental results are presented and discussed.

Up to now, we have only described the interpolation process of the first pass. Once the missing HR pixels in the first pass are interpolated as described above, half of the HR pixels are obtained. The remaining half of the missing HR pixels are to be interpolated in the second pass. The interpolation problem in the second pass is essentially the same as in the first pass. The only difference is that the SAI algorithm now interpolates the missing HR pixels $y_i \in I_h$ using their four 4-connected neighbors, which are either known in I_l or estimated in the first pass. The problem has the same formulation as in (5), if we simply rotate the spatial configuration of Fig. 2(a) by 45° (see Fig. 3).

IV. Model Parameter Estimation

A key to the success of the SAI algorithm is how well the model parameters $\boldsymbol{a}$ and $\boldsymbol{b}$ in (5) can be estimated using LR image samples. Referring to the spatial relation between the samples in Fig. 4(b), one gets a linear least-square estimator of the model parameter vector $\boldsymbol{b}$

$$\hat{\boldsymbol{b}} = \arg\min_{\boldsymbol{b}} \sum_{i \in W} \left(x_i - \sum_{1 \le t \le 4} b_t x_{i\diamond t}^{(4)} \right)^2 \tag{6}$$

where $x_{i\diamond t}^{(4)}$ are the four 4-connected neighbors of the location i in I_l as labeled in Fig. 4(b). Note that the estimates of $\boldsymbol{b}$ in (6) are made using the LR pixels $x_i \in I_l$ that have the same spatial orientation and the same scale as the way the HR pixels $y_i \in I_h$ are related by $\boldsymbol{b}$ in (5) [this is also clear in Fig. 2(b)]. Hence, the resulting estimates $\hat{\boldsymbol{b}}$ are optimal in the least-square sense under the assumption that the sample covariances do not change in the local window W, which is generally true for natural images.

However, the estimation of the model parameters $\boldsymbol{a}$ is more problematic. One can simply, as proposed by Li and Orchard [5], compute $\boldsymbol{a}$ via the following linear least-square estimation

$$\hat{\boldsymbol{a}} = \arg\min_{\boldsymbol{a}} \sum_{i \in W} \left(x_i - \sum_{1 \le t \le 4} a_t x_{i\diamond t}^{(8)} \right)^2 \tag{7}$$

where $x_{i\diamond t}^{(8)}$ are the four 8-connected neighbors of the location i in I_l as labeled in Fig. 4(a). The accuracy of (7) relies on a stronger assumption that the correlation between pixels is unchanged in different scales. This is because the distance between x_i and $x_{i\diamond t}^{(8)}$ in (7) is twice the distance between y_i and $x_{i\diamond t}^{(8)}$ in (5).

As argued in [5], the above assumption holds if the window in question has edge(s) of a fixed orientation and of sufficiently large scale. However, experiments show (see results in Section VI) that previous edge-based interpolation methods are prone to artifacts on small-scale spatial features of high curvature, for which the second order statistics may differ from LR to HR images. In such cases, the soft-decision estimation strategy of (5) can moderate the effects of estimation errors of (7), making the proposed SAI approach considerably more robust.

V. Algorithm Details

To perform soft-decision estimation, the SAI algorithm needs to operate on blocks of pixels. The neighboring blocks should have some overlaps to prevent possible block visual artifacts. Many spatial configurations of the overlapped blocks can be used. To be concrete let us consider a particular configuration as illustrated in Fig. 5. As shown in the figure, a block of 12 unknown pixels $y_1, y_2, \ldots, y_{12}$, arranged in an octagonal window (bounded by the solid line in Fig. 5), are jointly estimated, constrained by the 21 available LR pixels $x_1, x_2, \ldots, x_{21}$. Solving the least-squares problem of (5) in the octagonal window will yield a group of 12 estimated missing pixels. However, the SAI

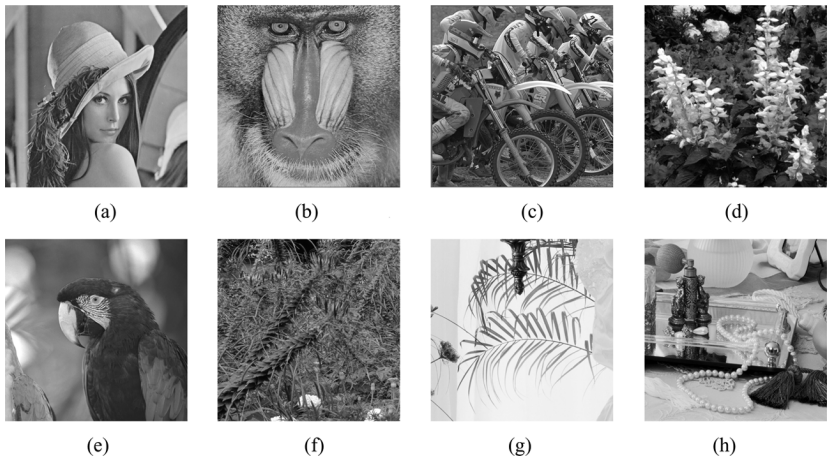


Fig. 6. Eight sample images in the test set. (a) Lena. (b) Baboon. (c) Bike. (d) Flower. (e) Parrot. (f) Bush. (g) Leaves. (h) Necklace.

algorithm will only output the interpolated values of the 4 innermost unknown pixels y_1, y_2, y_3, y_4. In other words, the estimation is done in a moving octagonal window with one layer of perimeter pixels being overlapped with neighboring windows.

The choice of window size and the degree of spatial overlap are design details related to implementation complexity and viewer preference on image appearance. Since the algorithm interpolates one block of missing pixels at a time by solving (5), the larger the block and the smaller the overlap, the faster the algorithm runs. However, the large block size may reduce the adaptability of the PAR model if there are varying features of small scale in a locality. Although higher degree of spatial overlap of neighboring windows means less likelihood of block artifacts, particularly near the boundaries of different features, it may cause some blurring of sharp edges.

If L_2 norm is used, the SAI algorithm involves solving three least-squares problems, namely the determination of model parameters $\boldsymbol{b}$, the determination of model parameters $\boldsymbol{a}$, and the minimization problem of (5). For each block, we can compute $\boldsymbol{b}$ and $\boldsymbol{a}$ in the closed form solutions of (6) and (7). Namely

$$\hat{\boldsymbol{b}} = (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{v} \tag{8}$$

where the column vector $\mathbf{v}$ is composed of all LR pixels x_i inside the block. The ith row of matrix $\mathbf{B}$ consists of the four 4-connected neighbors $x_{i\diamond t}^{(4)}$ of x_i, $t = 1, 2, 3, 4$, and

$$\hat{\boldsymbol{a}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{v} \tag{9}$$

where the ith row of matrix $\mathbf{A}$ consists of the four 8-connected neighbors $x_{i\diamond t}^{(8)}$ of x_i, $t = 1, 2, 3, 4$.

We rewrite (5) in matrix form

$$\min_{\boldsymbol{y}} \{||\boldsymbol{Cy} - \boldsymbol{Dx}||\} \tag{10}$$

where $\boldsymbol{y} = (y_1, y_2, \ldots, y_{12})$ is the vector of the 12 unknown pixels in the current octagon window as labeled in Fig. 5, $\boldsymbol{x} = (x_1, x_2, \ldots, x_{21})$ is the vector of the 21 available LR pixels inside and on the boundary of the octagon window in Fig. 5, and

$$\boldsymbol{C} = \begin{bmatrix} \mathbf{I}_{12} \\ \mathbf{C}_1 \\ \lambda\mathbf{C}_2 \end{bmatrix} \quad \boldsymbol{D} = \begin{bmatrix} \mathbf{D}_1 \\ \mathbf{D}_2 \\ \mathbf{0}_{4\times 21} \end{bmatrix} \tag{11}$$

where $\mathbf{I}$ is the identity matrix with the subscript being its dimension, $\mathbf{0}$ is the zero matrix whose dimension is indicated by the subscript, λ is the Lagrangian factor in (5), and

$$C_1^T = \begin{bmatrix} \hat{a}_4 & 0 & \hat{a}_1 & \hat{a}_2 & 0 \\ \hat{a}_3 & \hat{a}_1 & \hat{a}_2 & 0 & 0 \\ 0 & \hat{a}_4 & \hat{a}_3 & 0 & \hat{a}_2 \\ 0 & 0 & \hat{a}_4 & \hat{a}_3 & \hat{a}_1 \\ 0 & 0 & 0 & \hat{a}_4 & 0 \\ 0 & 0 & 0 & \hat{a}_1 & 0 \\ \hat{a}_1 & 0 & 0 & 0 & 0 \\ \hat{a}_2 & 0 & 0 & 0 & 0 \\ 0 & \hat{a}_2 & 0 & 0 & 0 \\ 0 & \hat{a}_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \hat{a}_3 \\ 0 & 0 & 0 & 0 & \hat{a}_4 \end{bmatrix}$$

$$C_2^T = \begin{bmatrix} 1 & -\hat{b}_4 & 0 & -\hat{b}_1 \\ -\hat{b}_2 & 1 & -\hat{b}_1 & 0 \\ 0 & -\hat{b}_3 & 1 & -\hat{b}_2 \\ -\hat{b}_3 & 0 & -\hat{b}_4 & 1 \\ 0 & 0 & 0 & -\hat{b}_4 \\ -\hat{b}_4 & 0 & 0 & 0 \\ -\hat{b}_1 & 0 & 0 & 0 \\ 0 & -\hat{b}_1 & 0 & 0 \\ 0 & -\hat{b}_2 & 0 & 0 \\ 0 & 0 & -\hat{b}_2 & 0 \\ 0 & 0 & -\hat{b}_3 & 0 \\ 0 & 0 & 0 & -\hat{b}_3 \end{bmatrix} \tag{12}$$

$$D_2 = [\mathbf{I}_5 \quad \mathbf{0}_{5\times 16}]$$

$$D_1 = \{d_1(i,j)\}, \quad i = 1, 2, \ldots, 12, \ \ j = 1, 2, \ldots, 21$$

where $d_1(i,j) = \hat{a}_t$ if x_j is the neighbor $x_{i\diamond t}^{(8)}$ of $y_i, t = 1, 2, 3, 4$; otherwise, $d_1(i,j) = 0$.

Therefore, the estimated block of pixels are, as the solution of (10)

$$\boldsymbol{y} = \mathbf{F}\boldsymbol{x}, \text{ where } \mathbf{F} = (\mathbf{C}^T\mathbf{C})^{-1}\mathbf{C}^T\mathbf{D}. \tag{13}$$

As can be seen from (13), each missing pixel y_i is estimated as a linear combination of all the available LR pixels $\boldsymbol{x}$ in the

TABLE I
PSNR (DECIBELS) RESULTS OF THE RECONSTRUCTED HR IMAGES BY DIFFERENT METHODS. NUMBER IN PARENTHESES IS THE RANK OF THE METHOD. FOR THE SAI ALGORITHM, ITS GAIN IN DECIBELS OVER THE SECOND-BEST METHOD IS ALSO GIVEN

Image	Bicubic	Method in [3]	Method in [5]	Method in [7]	SAI
Lena	33.92 (3)	33.09 (5)	33.76 (4)	33.92 (2)	**34.74** (1; 0.81dB)
Flower	32.30 (2)	30.18 (5)	30.47 (4)	31.89 (3)	**32.51** (1; 0.21dB)
Leaves	30.52 (4)	30.25 (5)	30.68 (3)	30.73 (2)	**32.16** (1; 1.43dB)
Baboon	22.92 (3)	22.84 (4)	23.10 (2)	22.79 (5)	**23.28** (1; 0.18dB)
Bush	26.34 (2)	25.79 (3)	25.53 (5)	25.79 (4)	**26.56** (1; 0.22dB)
Necklace	31.63 (2)	30.75 (5)	30.84 (4)	31.35 (3)	**31.89** (1; 0.36dB)
Bike	25.53 (3)	25.39 (4)	25.35 (5)	25.64 (2)	**26.53** (1; 0.89dB)
Parrot	35.80 (2)	35.26 (5)	35.61 (3)	35.46 (4)	**36.32** (1; 0.52dB)
Average	29.87 (2)	29.19 (5)	29.42 (4)	29.70 (3)	**30.51** (1; 0.64dB)

block, where the weights are specified by the ith row of the matrix $\mathbf{F}$ which is constructed by matrices C and D. Although the autoregression model parameters $\boldsymbol{a}$ and $\boldsymbol{b}$ appear to relate y_i to its immediate 8-connected or 4-connected neighbors only, the net effect of soft-decision block estimation is to interpolate y_i using all known pixels $\boldsymbol{x}$ in a local window. This is equivalent to an adaptive nonseparable 2-D interpolation filter whose order is the same as the block size of soft-decision estimation, which distinguishes the SAI technique from the existing adaptive image interpolation methods.

Clearly, the computation bottleneck of the SAI algorithm is in solving (13). Inverting the 12×12 matrix (13) is expensive. Instead, we use the steepest descent method, which ensures global minimum for the objective function (10) is convex. In particular, by exploiting the overlaps of moving windows, we can launch the steepest descent algorithm from a good initial point to achieve fast convergence. Referring to Fig. 5, we see that eight of the twelve unknown pixels in the current octagonal window have at least one estimate obtained when the adjacent windows to the north, northwest, northeast and west were processed. Due to spatial coherence of the HR image, these estimates are statistically good initial values of the corresponding variables in the objective function. For the other four unknown pixels which have no estimates yet, we use results of a traditional interpolation method (e.g., bicubic interpolation) as the initial estimates. With this initialization the steepest descent algorithm can converge in three iterations on average in our experiments. Also, note matrix $\mathbf{C}$ is quite sparse with only 49 out of 252 elements being nonzero. This sparsity of $\mathbf{C}$ can be exploited to save computations.

Another way of reducing computation complexity is to perform soft-decision estimation only in areas of high activities, because simple methods, such as bicubic interpolation, suffice to interpolate smooth 2-D waveforms. We classify the high activity areas based on the local variances estimated from LR pixels. If the local variance is above a threshold, the pixel is declared in the area of high activities. We empirically find that the variance threshold value of 100 realizes a good tradeoff between complexity and performance. Fortunately, since most natural images have an exponentially decaying power spectrum, only a small fraction of pixels (10% to 25% under the above threshold) need to be interpolated by the soft-decision method to ensure good visual quality.

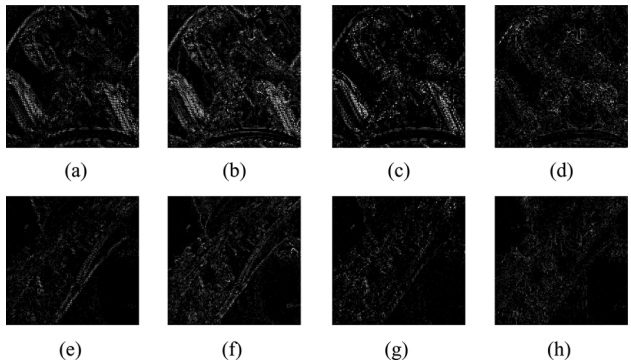


Fig. 7. Pixel locations where the SAI algorithm has smaller interpolation errors than the competing methods (only those of difference 3 or greater are plotted) on Bike and Lena images. The intensity level represents the magnitude of reduction in interpolation error by the proposed method from the competing method. (a) Bicubic. (b) Method [3]. (c) Method [7]. (d) Method [5]. (e) Bicubic. (f) Method [3]. (g) Method [7]. (h) Method [5].

VI. EXPERIMENTAL RESULTS AND REMARKS

Extensive experiments were conducted to evaluate the proposed new image interpolation technique in comparison with its predecessors. For thoroughness and fairness of our comparison study, we selected a large set of test images, including some of more difficult cases for image interpolation. Fig. 6 lists eight example images in our test set, some of which were also used as test images in previous papers on image interpolation.

The comparison group includes four other image interpolation methods: bicubic interpolation [1], subpixel edge localization [3], edge-directed interpolation [5], and fused bidirectional interpolation [7]. Table I tabulates the PSNR results of the five different methods when applied to the eight test images of Fig. 6. On all instances, the proposed SAI algorithm consistently ranks the first among all methods in terms of PSNR performance. On images of rich high frequency components, such as Leaves and Bikes, the SAI algorithm exceeds the PSNR values of the second best method by 1 dB or more. Since PSNR is an average quality measure, we plot in Fig. 7 the spatial locations where the SAI algorithm produces significantly smaller interpolation errors than the competing methods for more localized image quality assessment. The intensity level of the plots represents the magnitude of reduction in interpolation error by the new method from the competing method. Fig. 7 clearly demonstrates the advantage

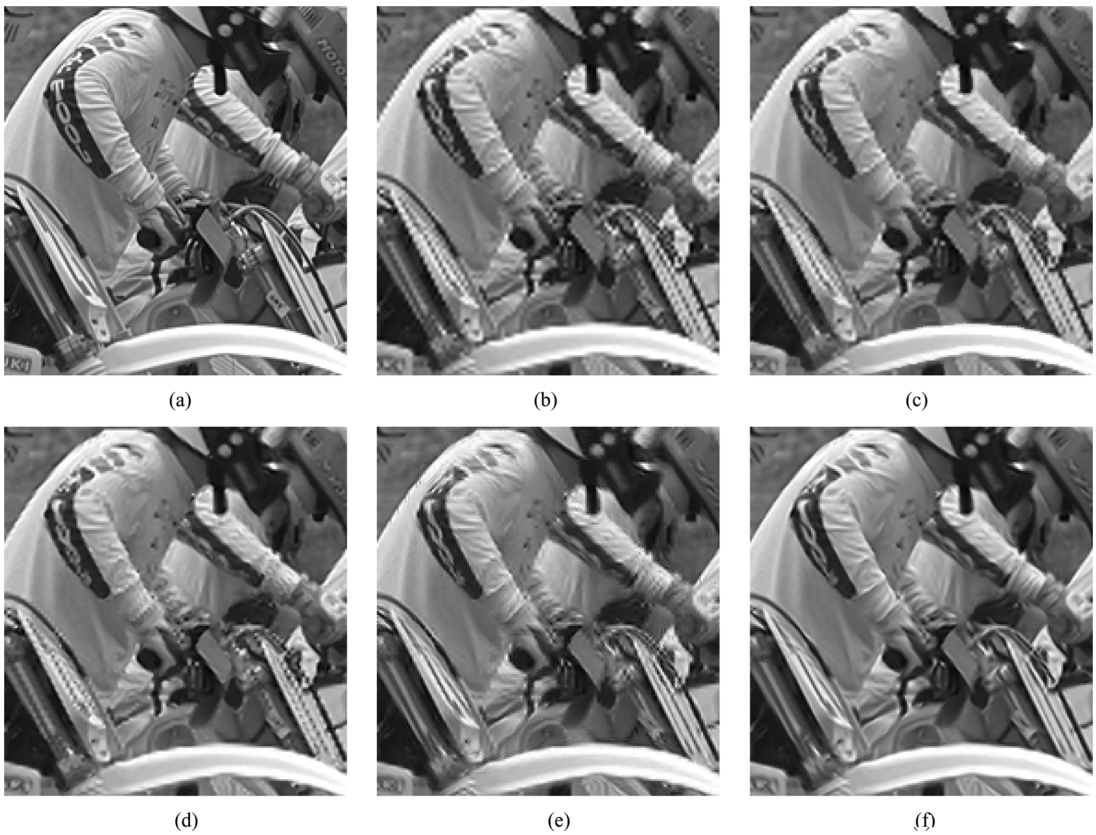


Fig. 8. Comparison of different methods on Bike image. (a) Original HR image. (b) Bicubic interpolation. (c) Method in [3]. (d) Method in [7]. (e) Method in [5]. (f) SAI.

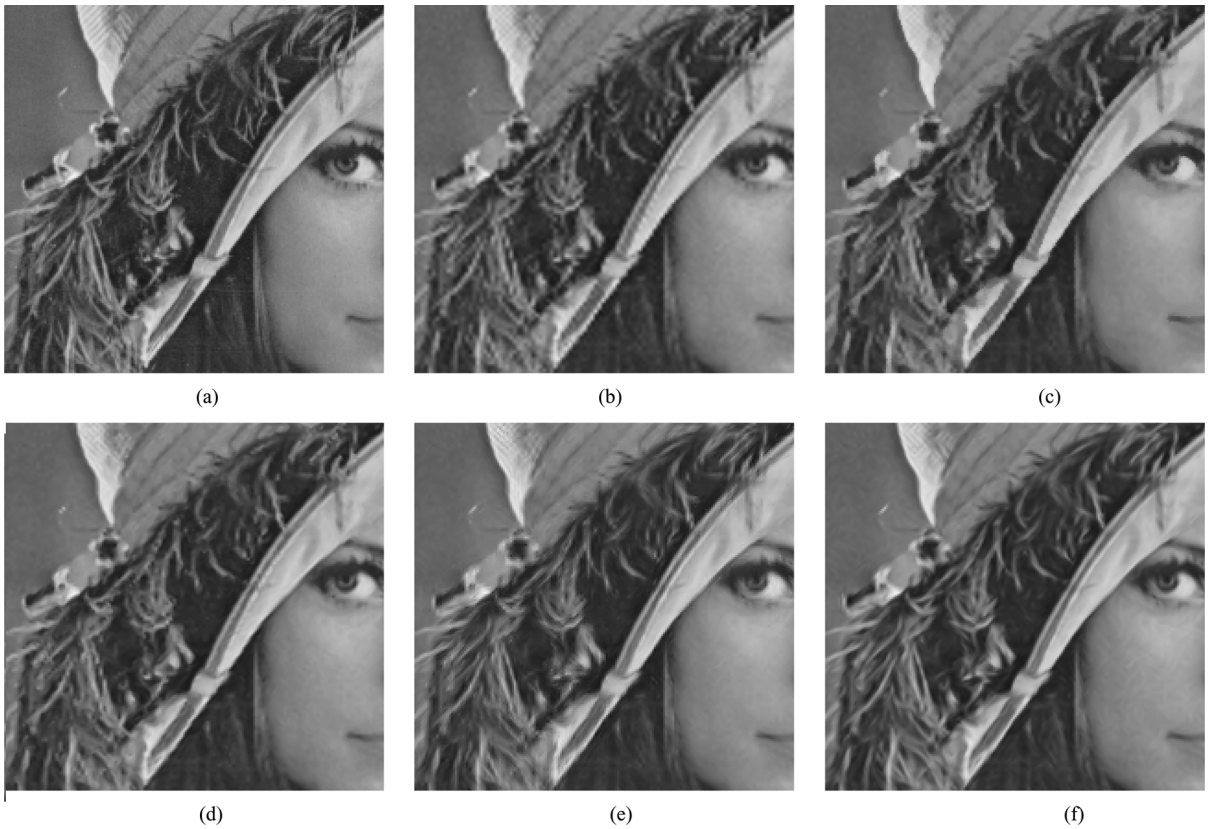


Fig. 9. Comparison of different methods on Lena image. (a) Original HR image. (b) Bicubic interpolation. (c) Method in [3]. (d) Method in [7]. (e) Method in [5]. (f) SAI.

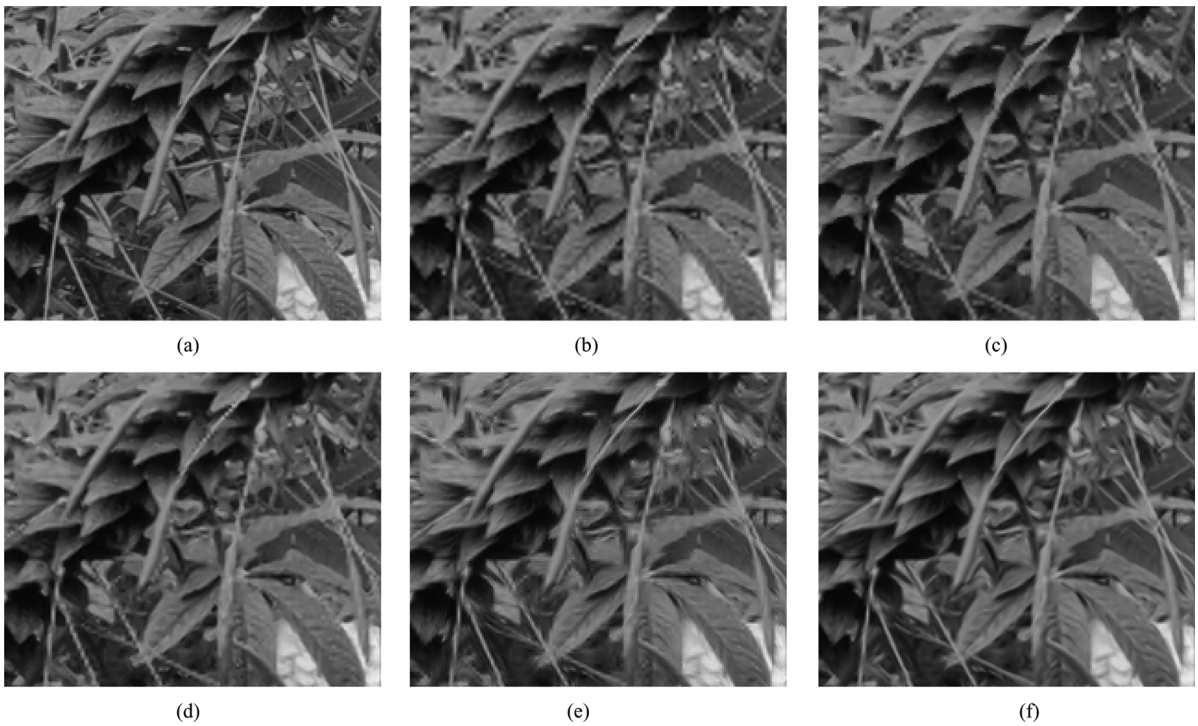


Fig. 10. Comparison of different methods on Bush image. (a) Original HR image. (b) Bicubic interpolation. (c) Method in [3]. (d) Method in [7]. (e) Method in [5]. (f) SAI.

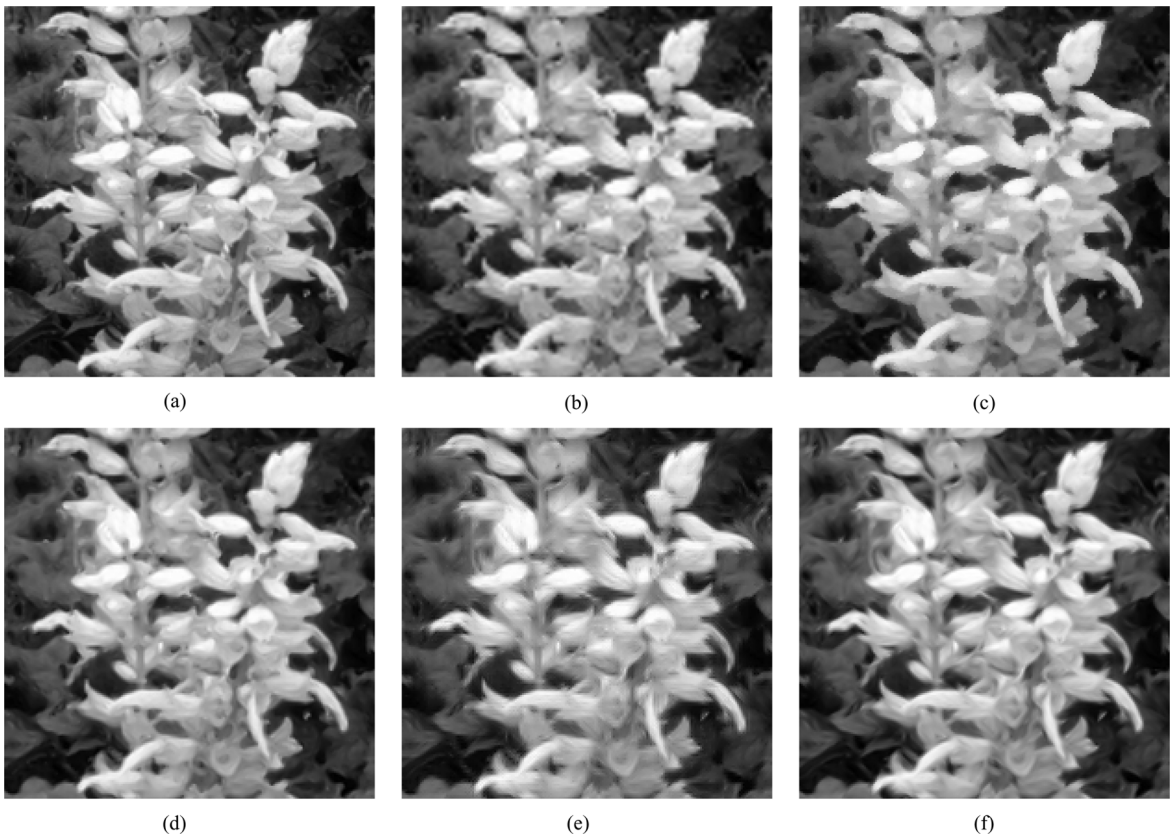


Fig. 11. Comparison of different methods on Flower image. (a) Original HR image. (b) Bicubic interpolation. (c) Method in [3]. (d) Method in [7]. (e) Method in [5]. (f) SAI.

of the SAI algorithm in reproducing the high frequency image constructs (edges and textures) over the other methods.

Given the fact that the human visual system is sensitive to errors near edges which signify object shapes and interobject relationship, one can expect from the spatial patterns of Fig. 7

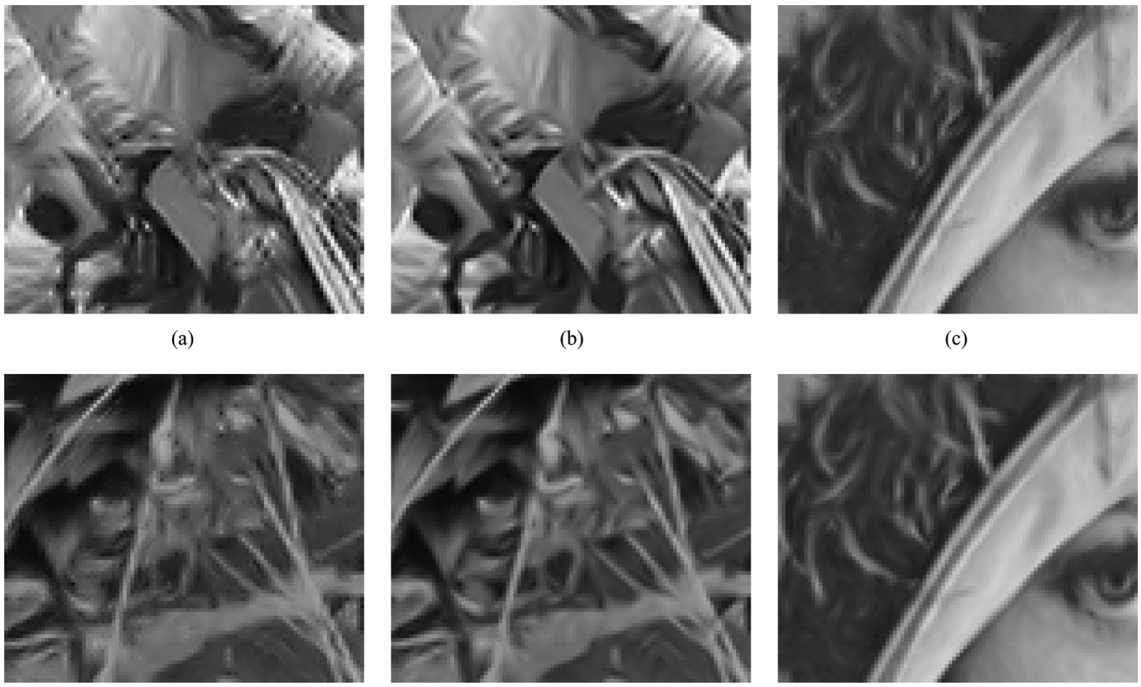


Fig. 12. Reconstructed images enlarged to compare the edge-directed interpolation method and the SAI algorithm. (a) Method in [5]. (b) SAI. (c) Method in [5]. (d) Method in [5]. (e) SAI. (f) SAI.

that the SAI algorithm should achieve superior visual quality. Figs. 8–11 compare the results of the five different image interpolation methods on test images Bike, Lena, Bush, and Flower, respectively. Different visual characteristics of the evaluated methods are exhibited near edges and fine textures in the test images.

The bicubic interpolation method tends to blur the image details more than other methods, and it also generates prominent jaggies along sharp edges. This method is in general inferior to the others in visual quality despite its PSNR measure is the second highest on average, next only to the proposed new method.

The method of subpixel edge localization [3] reproduces sharp edges, but the reconstructed edges are somewhat contrived and at times unnatural. This problem is exemplified by the reconstructed flower petals in Fig. 11(c), and by the rim of hat in test image Lena Fig. 9(c). This method consistently ranks lower than others in terms of PSNR.

The edge-directed interpolation method [5] is very competitive in terms of visual quality. This is primarily because it preserves long edges well. However, in high activity areas where features have small curvatures or multiple edges intersect, this method sometimes generates speckle interpolation noises and ringing artifacts, as shown in Figs. 8(e) and 11(e), and more clearly in close-up parts of the reconstructed images in Fig. 12. Such flaws are the reason that this method only ranks the forth in PSNR in the comparison group of five methods, although it achieves pleasant visual effects on large edge structures.

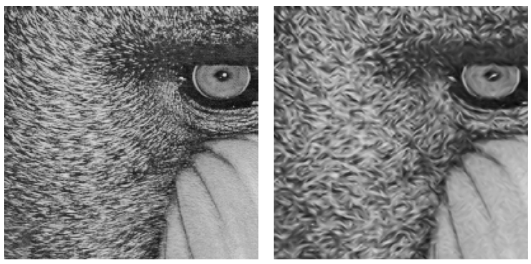
The fused bidirectional interpolation method [7] takes a middle ground between the direction-less bicubic interpolation and edge-directed interpolation. It reproduces sharper large scale edges than the bicubic method, but the reconstruction is not as good as the method of [5] when the LR image contains enough information to correctly learn the edge direction. However, on small–scale features where the edge direction learnt from the LR image is not reliable, this method can reduce speckle interpolation noises via the MMSE fusion process. It is interesting to note that the fusion method ranks in between the bicubic method and the method of [5] in PSNR.

As is evident in Figs. 8–11, the SAI algorithm eliminates most of the visual defects associated with the other methods. It reproduces visually more pleasant HR images than the edge-directed interpolation method [5] that is considered as one of the best so far. Please refer to Fig. 12 for a side-by-side comparison between the two methods. Of particular significance is the fact that the SAI algorithm obtains superior visual quality on image features of large and small scales alike. This robustness is attributed to the soft-decision estimation technique that enforces the spatial coherence of estimated pixels according to the PAR model.

Although the proposed SAI algorithm is presented to double the horizontal and vertical resolutions, it can be readily generalized to scaling factor $S = 2^z$ with z being a positive integer. One can simply apply the proposed SAI algorithm z times to scale the input image by S times. For an arbitrary scaling factor S (not an integer power of two), the interpolation can be done by first using the SAI algorithm to expand the input image by 2^z times such that $2^z < S < 2^{z+1}$, and then applying a conventional image interpolation algorithm, such as bicubic or bilinear interpolation, to scale up the output image of the new method by s times such that $2^z s = S$. Fig. 13 is an example for scaling factor of 3, and it shows that the image interpolated by the SAI algorithm followed by bicubic interpolation has better visual quality than the image by bicubic interpolation only.

(a) (b)

Fig. 13. Images scaled up three times using different methods: (a) the bicubic method; (b) the SAI algorithm followed by the bicubic interpolation (i.e., first scaling by a factor of 2 and then by a factor of 1.5).



(a) (b)

Fig. 14. Visual effects of over-fitting of LR data. (a) Original image. (b) Output image of the SAI algorithm without safe guard against over-fitting.

As described in Section IV, the estimation of model parameters assumes that the spatial correlation between the HR pixels is approximately the same as between LR pixels. When this assumption is violated, the SAI algorithm may introduce some false edges or textures due to over-fitting of LR data. This over-fitting problem can be seen in Fig. 14. For this particular test image, the spatial correlation changes after down sampling, which causes the SAI algorithm to produce erroneous textures. There is a way to detect where the over-fitting problem will likely occur. If the value of the cost function (5) at the convergence of the SAI algorithm is above a threshold, then chances are that the PAR model is not valid in the locality. As a recourse, one can switch to a more conservative method such as bicubic interpolation to prevent the generation of false pixel structures. However, this also tends to blur some sharp edges. Further research is needed to find a better solution to the problem.

VII. Conclusion

A novel soft-decision approach is proposed for adaptive image interpolation. When coupled with a piecewise autoregressive image model, the soft-decision approach estimates a block of missing pixels jointly by imposing an adaptively learnt spatial sample relation not only between known pixels and missing pixels but also between missing pixels themselves. This new image interpolation technique outperforms the existing methods in both PSNR measure and subjective visual quality over a wide range of scenes, by preserving the spacial coherence of the reconstructed HR image on features of large and small scales alike.

References

[1] R. G. Keys, "Cubic convolution interpolation for digital image processing," *IEEE TranS. Acoust., Speech, Signal Process.*, vol. ASSP-29, no. 6, pp. 1153–1160, Dec. 1981.

[2] H. S. Hou and H. C. Andrews, "Cubic splines for image interpolation and digital filtering," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-26, no. 6, pp. 508–517, Dec. 1978.

[3] 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.

[4] S. Carrato and L. Tenze, "A high quality 2× image interpolator," *IEEE Signal Process. Lett.*, vol. 7, no. 6, pp. 132–135, Jun. 2000.

[5] X. Li and M. T. Orchard, "New edge-directed interpolation," *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1521–1527, Oct. 2001.

[6] D. D. Muresan and T. W. Parks, "Adaptively quadratic (aqua) image interpolation," *IEEE Trans. Image Process.*, vol. 13, no. 5, pp. 690–698, May 2004.

[7] L. Zhang and X. Wu, "Image interpolation via directional filtering and data fusion," *IEEE Trans. Image Process.*, vol. 15, no. 8, pp. 2226–2238, Aug. 2006.

[8] 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.

[9] Y. Zhu, S. C. Schwartz, and M. T. Orchard, "Wavelet domain image interpolation via statistical estimation," in *Proc. Int. Conf. Image Processing*, Sep. 2001, vol. 3, pp. 840–843.

[10] D. D. Muresan and T. W. Parks, "Prediction of image detail," in *Proc. Int. Conf. Image Processing*, Sep. 2000, vol. 2, pp. 323–326.

[11] X. Wu and X. Zhang, "Image interpolation using texture orientation map and kernel fisher discriminant," in *Proc. IEEE Int. Conf. Image Processing*, Sep. 2005, vol. 1, pp. I–49–52.

[12] 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.

[13] X. Wu, K. U. Barthel, and W. Zhang, "Piecewise 2-D autoregression for predictive image coding," in *Proc. Int. Conf. Image Processing*, Chicago, IL, Oct. 1998, vol. 3, pp. 901–904.

[14] X. Wu and N. Memon, "Context-based adaptive lossless image coding," *IEEE Trans. Commun.*, vol. 45, no. 4, pp. 437–444, Apr. 1997.

[15] B. Meyer and P. Tischer, "TMW—A new method for lossless image compression," presented at the Picture Coding Symp., Berlin, Germany, Oct. 1997.

[16] A. Said and W. Pearlman, "An image multiresolution representation for lossless and lossy compression," *IEEE Trans. Image Process.*, vol. 5, no. 9, pp. 1303–1330, Sep. 1996.

[17] I. Matsuda, H. Mori, J. Maeda, and S. Itoh, "Design and evaluation of minimum-rate predictors for lossless image coding," (in Japanese) *Trans. IEICE (DII)*, vol. J85-D-II, no. 3, pp. 448–456, Mar. 2002.

Xiangjun Zhang received the B.Sc. degree from Xi'an Jiaotong University, Xi'an, China, in 1997, and the M.S. degree in electrical engineering from Tsinghua University, Beijing, China, in 2002. She is currently pursuing the Ph.D. degree at the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada.

Her research interests include image processing, multimedia communication, and signal compression.

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

He started his academic career in 1988, and has since been on the faculty at the University of Western Ontario, New York Polytechnic University, Brooklyn, and currently McMaster University, Hamilton, ON, where he is a Professor in the Department of Electrical and Computer Engineering, and where he holds the NSERC-DALSA Industrial Research Chair in Digital Cinema. His research interests include image processing, multimedia compression, joint source-channel coding, multiple description coding, and network-aware visual communication. He has published over 180 research papers and holds two patents in these fields.

Dr. Wu is an Associate Editor of IEEE Transactions on Multimedia.