SIMILARITY MODULATED BLOCK ESTIMATION FOR IMAGE INTERPOLATION

Jie Ren, Jiaying Liu†*, Wei Bai and Zongming Guo*

Institute of Computer Science and Technology, Peking University, Beijing 100871, China

ABSTRACT

Modeling the nonstationarity of image signals is one of the challenging issues for image interpolation. In this paper, we propose a similarity probability modeling to faithfully characterize the nonstationarity of image signals, and present a novel image interpolation algorithm based on the proposed model. The missing pixels are estimated in groups by weighted block estimation. The weight of each pixel inside the block is defined as the similarity probability between itself and the centered to-be-interpolated pixel. It is demonstrated by the experimental results that the proposed method preserves the edge structures of the interpolated images better than the state-of-the-art interpolation methods. Annoying artifacts nearby the sharp edges are also greatly reduced.

Index Terms— Image interpolation, similarity probability modeling, block estimation

1. INTRODUCTION

Image interpolation is the process of producing high-resolution (HR) images from its low-resolution (LR) counterparts. Understanding and modeling the inherent image structures partly reflected by LR images is the key task and challenge for image interpolation. Conventional interpolation methods, such as bilinear and bicubic interpolation regard the ground-truth HR images to be continuous and smooth. These methods work well in smooth regions but fail to capture the fast varying property around edge structures, suffering from the problem of aliasing, blurring or ringing artifacts.

It is well recognized that edge structure is one of the most important image features in natural images. To address the problems of conventional interpolation methods, many spatial adaptive interpolation methods have been proposed to preferably preserve the edge structures. A major challenge for developing such adaptive interpolation algorithm is modeling the nonstationarity of image signals, in particular the edge structures. Fortunately, geometric regularity property of natural images (*i.e.*, image intensity field evolves more slowly along the edge orientation than across the edge orientation) gives the prior knowledge to model the image signals. Edge-adaptive methods [1, 2, 3] explicitly extract the edge structural information, such as edge directions and widths. Then missing pixels are fitted by directional interpolating along the detected edge directions to enhance the geometric regularity property of the interpolated image. However, it is difficult to carry out this process when irregular textures or noises are present in the LR images. Instead of using explicitly extracted geometry information, other methods [4, 5, 6, 7] implicitly exploit the inherent image structures in the LR images through statistical models. Li and Orchard [4] proposed a covariance-based method based on the autoregressive (AR) model to produce better interpolation results along arbitrarily oriented edge. The HR covariances are approximated by the LR covariances which are estimated under the assumption that the statistics are kept stationary in a local window. Zhang and Wu [7] further introduced more correlations on different directions into the AR model and jointly estimated the missing HR pixels with the imposed statistical consistency constraint in the local window. These implicit methods are grounded when the piecewise statistical stationary assumption in a local window is tenable. However, the richness and varying scales of textures and edges in natural images violate the validity of the stationary assumption.

In this paper, we propose a new implicitly statistical method for image interpolation. First, a similarity probability modeling is proposed to faithfully model the image signal which is adaptive to the local image structure at different scales. Then, according to this modeling, an improved interpolation algorithm with the similarity modulated block estimation is presented. Missing HR pixels in a block (local window) are jointly estimated by minimizing an energy function of model fitting errors. Experimental results show that the proposed method is more adaptive to the local image structures, and thus reduces most of the artifacts caused by the large disparity in local image structures.

The rest of the paper is organized as follows. Sec. 2 gives brief reviews and remarks to the AR model. Sec. 3 describes the new proposed image interpolation algorithm based on a novel similarity probability modeling. Experimental results and a comparison study with the state-of-the-art interpolation techniques are presented in Sec. 4. Finally, concluding remarks are given in Sec. 5.

†Corresponding author

This work was supported by the National Basic Research 973 Program of China under Contract 2009CB320907, National Natural Science Foundation of China under contract No. 60902004 and Beijing Municipal Natural Science Foundation under contract No. 4102025.

978-1-4577-1302-6/11/$26.00 ©2011 IEEE

2. AR INTERPOLATION MODEL

AR model is a type of random process in statistics and signal processing, which is often used to model and predict various types of natural phenomena. A 2-D image signal I can be modeled as AR process as follows,

$$I_{(i,j)} = \sum_{(x,y)\in\Omega} \psi_{(x,y)} I_{(i+x,j+y)} + \varepsilon_{(i,j)}, \tag{1}$$

where $\psi_{(x,y)}$ are the model parameters, $\varepsilon_{(i,j)}$ is a white noise, and Ω is a local neighborhood around the pixel $I_{(i,j)}$. The number of neighbors in Ω decides the order of the model. To model the nonstationarity of image signals, $\psi_{(x,y)}$ need to be adaptive to the local pixel structure around pixel $I_{(i,j)}$, *e.g.*, giving large coefficients to the neighbors along the local edge direction. To compute the optimal $\psi_{(x,y)}$, a local window is assigned and the statistics are considered to be stationary in the local window [4, 7]. However, the fixed size of local window can not adapt to the local edge features at different scales. Especially when the scale of local edge feature is smaller than the selected local window size, the stationarity assumption is violated.

An example of the irregular statistical stationarity of image signals within a local window is shown in Fig.1(a). The statistics vary significantly in different regions, while they stay stationary or nearly stationary within each region. When interpolating the centered to-be-interpolated pixel y_c, the reasonable local structure information is provided by the samples within *red* region (area a_1). However, samples in *yellow* and *green* regions (area a_2 and a_3) will violate the accuracy statistics estimation at y_c. A natural idea is to explicitly exclude those pixels with large disparity in local structures (*e.g.*, pixels inside a_2 and a_3) and only remain those pixels with similar local structures (*e.g.*, pixels inside a_1). However, the hard-decision of whether a pixel belongs to the statistics stationary area is a difficult task when only the LR image pixels are available. Besides, spatial template of interpolation becomes too complicated to be implemented. Therefore, a probability-based method is proposed in this paper to implicitly define the irregular area which contains similar local structures to that of pixel y_c in the following section.

3. THE PROPOSED INTERPOLATION ALGORITHM

In this section, we present a new implicitly statistical method for image interpolation. According to the analysis of Section 2, a similarity probability modeling is first proposed to model the nonstationarity of image signals. Then the new interpolation algorithm is described based on the proposed similarity probability modeling.

Specifically, all the missing HR pixels are interpolated by two passes: the first pass is to interpolate the diagonal missing pixels (gray squares in Fig.1(b)) by using available LR pixels (black squares in Fig.1(b)); the second pass is for the remaining missing pixels (white squares in Fig.1(b)) by using LR pixels and first-pass generated HR pixels. These two passes differ only in orientation and scale. For convenience, we shall take the first pass to elaborate our proposed method.

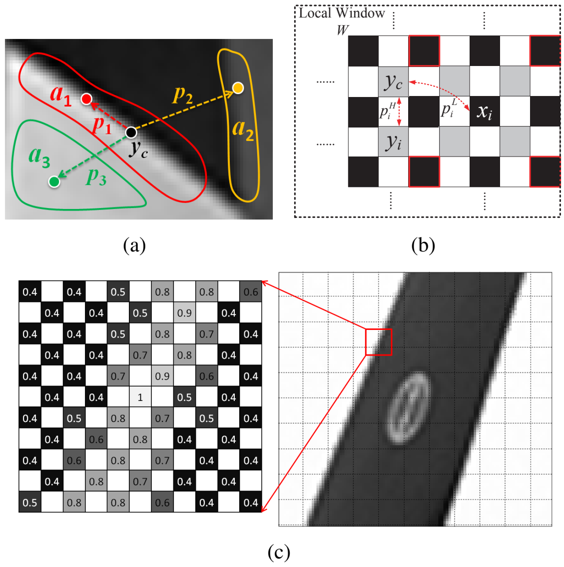

Fig. 1. Nonstationarity of image signals and similarity probability modeling. (a) The irregularity of statistics stationarity. (b) Similarity probability modeling at high and low resolution in the first pass. The black and gray squares represent the LR pixels and missing HR pixels, respectively. (c) An example of similarity probability distribution in a local window.

3.1. Similarity probability modeling

First, we introduce the similarity probability to represent the degree of local structural similarity between two pixels in a local window. Similarity in structures between two pixels is defined in the sense that they have similar relationship with their neighbors. In this paper, we use the the four diagonal neighbors from the 8-connected neighborhood of pixel z_i, forming a vector $\vec{\mathcal{N}}_{z_i}$, to represent its local image structure. Then the similarity probability between two arbitrary pixels z_i and z_j is modeled as the Gaussian function of the Euclidean distance between their local structure vectors, *i.e.*,

$$p(z_i, z_j) = \exp\left(-\left\|\vec{\mathcal{N}}_{z_i} - \vec{\mathcal{N}}_{z_j}\right\|_2^2 / h^2\right), \tag{2}$$

where the only parameter h controls the shape of the exponential function. In particular, to make the structural similarity between two pixels independent to the image intensity field, the local structure vector $\vec{\mathcal{N}}_{z_i}$ is normalized, $\vec{\mathcal{N}}_{z_i} = \frac{\vec{\mathcal{N}}_{z_i}+\epsilon}{max(\vec{\mathcal{N}}_{z_i})+\epsilon}$, where operator $max(\cdot)$ extracts the maximum value among the vector elements, and ϵ is some positive constant to avoid the divide-by-zero error.

The similarity probability between each pixel and the center pixel y_c is computed throughout the local window. Let x_i and y_i be the LR pixels and HR pixels in local window W, respectively. The similarity probability between x_i and y_c are written as p_i^L. Similarly, p_i^H represents the similarity probability between y_i and y_c. Fig.1(b) gives the spatial configuration for computing similarity probability for HR and LR pixels. For computing p_i^H, $\vec{\mathcal{N}}_{y_i}$ is directly consist of the four nearest diagonal LR neighbors on the HR image grid. However, when computing p_i^L, representing $\vec{\mathcal{N}}_{x_i}$ of LR pixel x_i is a little problematic since the nearest diagonal neighbors of x_i on the HR image grid are the missing HR pixels which are still unknown. Instead, we construct the vector $\vec{\mathcal{N}}_{x_i}$ by taking the diagonal neighbor pixels on the LR image grid (black squares with red border in Fig.1(b)) , based on the assumption that local structures keep stationary at different scales.

Finally, the distribution of the similarity probability within the local window implicitly characterize the profile of the statistical stationary area. Fig.1(c) shows an example of similarity probability distribution on portion of the 'Airplane' image. The pixels with large disparity in local structures have significantly lower probabilities, which is shown with darker intensity in the left image. The profile of stationary area characterized by the probability distribution is well consistent with the image structure in the local window (red rectangle on the right image).

3.2. Similarity modulated block estimation

With above similarity modeling, we propose a new interpolation method, in which the missing HR pixels are jointly estimated by weighted block estimation. The whole HR image can be modeled as an AR process based on the Eq.(1) described in Section 2. To void the risk of data overfitting, we use two AR models with model parameters $\mathbf{a} = \{a_t\}$ and $\mathbf{b} = \{b_t\}(t = 1, 2, 3, 4.)$ instead of an AR model of order 8 to characterize the diagonal and cross-direction correlations, respectively,

$$z_i = \sum_{t=1}^{4} a_t z_{i\otimes t} + \varepsilon_i^{\otimes}, z_i = \sum_{t=1}^{4} b_t z_{i\oplus t} + \varepsilon_i^{\oplus}, \tag{3}$$

where z_i refers to either LR pixels x_i or HR pixels y_i in the local window W. $z_{i\otimes t}$ and $z_{i\oplus t}$ are the diagonal and cross-direction neighbors around z_i pixel, respectively. $\varepsilon_i^{\otimes}$ and $\varepsilon_i^{\oplus}$ are the fitting error of diagonal and cross-directional AR model. The spatial configurations of these two AR models are figured as Fig.2.

A block of missing HR pixels in W are jointly estimated by minimizing the following energy function Eq.(4),

$$E(\mathbf{y}, \mathbf{a}, \mathbf{b}|\mathbf{x}) = E_d(\mathbf{y}, \mathbf{a}|\mathbf{x}) + \lambda E_c(\mathbf{y}, \mathbf{b}|\mathbf{x}), \tag{4}$$

where λ is a parameter to balance the energy functions of diagonal fitting errors $E_d(\mathbf{y}, \mathbf{a}|\mathbf{x})$ and cross-directional fitting errors $E_c(\mathbf{y}, \mathbf{b}|\mathbf{x})$. According to the modeling described in Section 3.1, the similarity probability p_i actually indicates the consistency of AR model parameters for each pixel z_i in W. Therefore, the model fitting error at each pixel z_i is weighted by p_i, giving the following energy function formalization,

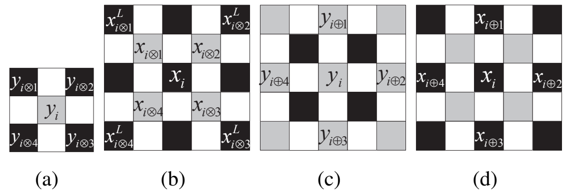

Fig. 2. The spatial configuration of AR model. (a) and (b) illustrate the diagonal relationship between HR and LR pixels. For LR pixel x_i in (b), we also indicate its diagonal neighbors $x_{i\otimes t}^L$ on the LR image grid. (c) and (d) illustrate the cross-direction relationship for HR pixel and LR pixel.

$$\begin{aligned} E_d(\mathbf{y}, \mathbf{a}|\mathbf{x}) &= \sum_{z_i \in W} p_i \left(z_i - \sum_{t=1}^{4} a_t z_{i\otimes t} \right)^2, \\ E_c(\mathbf{y}, \mathbf{b}|\mathbf{x}) &= \sum_{z_i \in W} p_i \left(z_i - \sum_{t=1}^{4} b_t z_{i\oplus t} \right)^2. \end{aligned} \tag{5}$$

The model parameters $\mathbf{a}$, $\mathbf{b}$ and the missing HR pixels $\mathbf{y}$ are both treated as variables in the non-linear optimal estimation in Eq.(4). Iterative methods, such as gradient descent, are needed to solve the optimization of Eq.(4). To reduce the computational cost, the parameters $\mathbf{a}$, $\mathbf{b}$ can be approximately estimated by the available LR pixels x_i in local window W ahead of time,

$$\begin{aligned} \hat{\mathbf{a}} &= \arg\min_{\mathbf{a}} \sum_{x_i \in W} p_i^L \left(x_i - \sum_{t=1}^{4} a_t x_{i\otimes t}^L \right)^2, \\ \hat{\mathbf{b}} &= \arg\min_{\mathbf{b}} \sum_{x_i \in W} p_i^L \left(x_i - \sum_{t=1}^{4} b_t x_{i\oplus t} \right)^2, \end{aligned} \tag{6}$$

where $x_{i\otimes t}^L$ is the diagonal neighbor pixels on the LR image grid (Please refer to the graphical illustration in Fig.2(b)). This approximation reduces the problem to a linear estimation,

$$\hat{\mathbf{y}} = \arg\min_{\mathbf{y}} \left\{ E_d'(\mathbf{y}|\mathbf{x}, \hat{\mathbf{a}}) + \lambda E_c'(\mathbf{y}|\mathbf{x}, \hat{\mathbf{b}}) \right\}, \tag{7}$$

where $E_d'(\mathbf{y}|\mathbf{x}, \hat{\mathbf{a}}) = \sum_{z_i \in W} p_i \left(z_i - \sum_{t=1}^{4} \hat{a}_t z_{i\otimes t} \right)^2$, $E_c'(\mathbf{y}|\mathbf{x}, \hat{\mathbf{b}}) = \sum_{z_i \in W} p_i \left(z_i - \sum_{t=1}^{4} \hat{b}_t z_{i\oplus t} \right)^2$.

Only the center pixel y_c is output in one block estimation process. For practical purpose, we perform the proposed algorithm in areas of high activities (local variances above 100). For the pixels in low activities area, we still use the bicubic interpolation due to its simplicity.

4. EXPERIMENTAL RESULTS

The proposed interpolation algorithm is implemented on MATLAB 7.6 platform and compared with conventional

Table 1. PSNR(dB) results of four interpolation methods.

Images	Bicubic	NEDI[4]	SAI[7]	Proposed
Lena	33.86	33.80	34.63	**34.67**
Pepper	32.77	33.30	33.51	**33.56**
Tulip	33.69	34.16	35.66	**35.82**
Cameraman	25.18	25.34	25.70	**25.78**
Monarch	31.72	31.68	32.90	**33.16**
Airplane	30.70	31.21	31.65	**31.72**
Caps	33.67	34.05	**34.48**	**34.48**
Bike	25.93	25.97	26.97	**27.02**

Bicubic interpolation method and two state-of-the-art interpolation methods: new edge-directed interpolation (NEDI) in [4]and soft-decision adaptive interpolation (SAI) in [7][1]. We have tested the proposed algorithm on a large image set, including the Kodak database and many standard test images also used in previous papers on image interpolation.

To compare the objective quality of different interpolation methods, the original HR images are first directly downsampled by a factor of two to generate the input LR images. Then different interpolation methods are applied to interpolate the input LR images to their original resolutions. Table 1 tabulates the PSNR results of the four interpolation methods on several images in our experiments. From Table 1, we can see that the proposed method produces comparable or often better PSNR results than other methods. It is worth noticing that for 'Monarch' and 'Tulip' image, the proposed method gains 0.26 dB and 0.16 dB respectively over the second-best SAI algorithm.

We also compare the visual quality of different interpolation methods. Bicubic interpolation blurs the edges in Fig.3. While the edge-directed method such as NEDI[4] preserves the long edge structure well, it produces annoying artifacts nearby the fast-evolving edges. SAI[7] and the proposed method achieve nearly the same good visual quality for most of the test images. However, comparing them in detail, we can see that the proposed method produces smaller interpolation errors than other methods, especially on the edge structures pointed by the red arrows. Our proposed method achieves the best visual quality and is more adaptive and robust to the fast evolving edge structures. Another interesting property of our method is that it tends to produce some continues edges which is visually plausible, even though these edges may not be faithful to the original images. This phenomena is demonstrated within the yellow boxes in Fig. 3.

5. CONCLUSION

In this paper, we propose a novel interpolation method based on an adaptive image signal model. The irregularity of statistics stationarity of image signals is implicitly characterized by the proposed similarity probability modeling. The proposed

[1]We thank the authors of [4, 7] for providing their source codes or executable program.

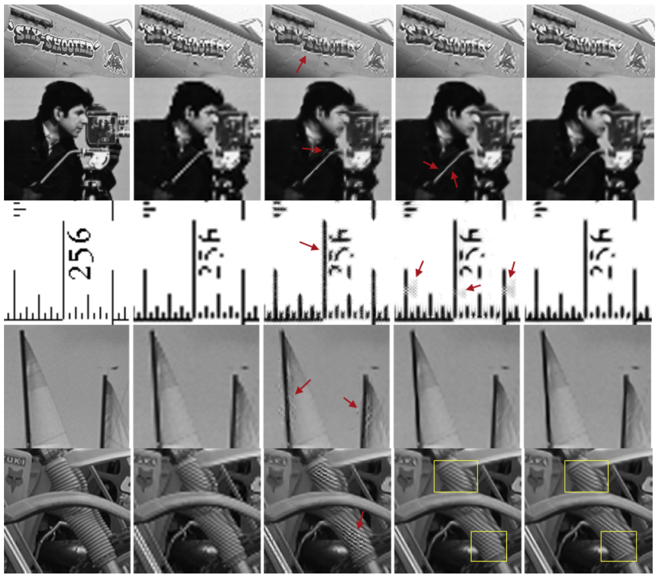

Fig. 3. Visual comparisons: Portions from various interpolated images using different methods. From top to bottom: airplane, cameraman, ruler, sailboat, bike. From left to right: ground truth, Bicubic, NEDI[4], SAI[7], proposed method.

interpolation method is more adaptive to the local fast evolving structures. Experimental results show that our proposed method preserves the edge structures in images better than other popular interpolation methods and greatly reduces the artifacts nearby sharp edges.

6. REFERENCES

[1] K. Jensen and D. Anastassiou, "Subpixel edge localization and the interpolation of still images," *IEEE Trans. Image Processing*, vol. 4, no. 3, pp. 285–295, March 1995.

[2] J. Allebach and P. W. Wong, "Edge-directed interpolation," in *Proc. Int. Conf. Image Processing*, September 1996, vol. 3, pp. 707–710.

[3] B.S. Morse and D. Schwartzwald, "Isophote-based interpolation," in *Proc. Int. Conf. Image Processing*, Oct. 1998, vol. 3, pp. 227–231.

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

[5] M. Li and T.Q. Nguyen, "Markov random field model-based edge-directed image interpolation," *IEEE Trans. Image Processing*, vol. 17, no. 7, pp. 1121–1128, July 2008.

[6] R. Xiong and W. Gao, "Tanner graph based image interpolation," in *Data Compression Conference (DCC)*, 2010, pp. 376–385.

[7] X. Zhang and X. Wu, "Image interpolation by adaptive 2-d autoregressive modeling and soft-decision estimation," *IEEE Trans. Image Processing*, vol. 17, no. 6, pp. 887–896, June 2008.