Markov Random Field Model-Based Edge-Directed Image Interpolation

Min Li and Truong Nguyen
ECE Dept., UC San Diego, La Jolla CA 92093-0407
http://videoprocessing.ucsd.edu

Abstract— **This paper presents an edge-directed image interpolation algorithm. In the proposed algorithm, the edge directions are implicitly estimated with a statistical-based approach. Consequently, the local edge directions are represented by length-16 vectors, which are denoted as weight vectors. The weight vectors are used to formulate geometric regularity constraint, which is imposed on the interpolated image through the Markov Random Field (MRF) model. Furthermore, the interpolation problem is formulated as a Maximum A Posterior (MAP)-MRF problem and, under the MAP-MRF framework, the desired interpolated image corresponds to the minimal energy state of a two-dimensional random field. Simulated Annealing method is used to search for the minimal energy state from a reasonable large state space. Simulation and comparison results show that the proposed MRF model-based edge-directed interpolation method produces edges with strong geometric regularity.**

Index Terms— **Image Interpolation, Markov Random Field, Edge-directed**

I. Introduction

Image interpolation is the process of producing high resolution image from its low resolution counterpart. The conventional image interpolation methods, such as bilinear, bicubic and splines, typically produce images with blur edges as a result of assuming continuity of the pixel intensity field, which is unrealistic. Spatial adaptive interpolation methods [1], [2], which adjust the interpolation coefficients according to the local pixel intensity properties, can be used to improve the interpolation performance. Among various spatial adaptive interpolation algorithms, interpolating according to local edge directions is an important idea. This is because edges are among most noticeable features in natural images. The visual quality near edge areas affects the visual quality of the overall interpolated image significantly. Interpolation along ideal step edges is not difficult since accurate edge direction information can be obtained explicitly from edge detectors [3], [4]. However, edges in natural images appear as spatially blurred edges due to sensor noises, focal blur, penumbral blur and shading, etc [5]. When edges are blur or noisy, it is difficult to explicitly specify the characteristics of edges, such as edge width, exact positions and digital directions [5]. This makes interpolation along natural edges difficult.

To avoid the difficulties with explicit edge-directed interpolation methods, implicit edge-directed interpolation methods are proposed. In this class of methods, the edge directions are not explicitly extracted. However, the interpolation is performed based on implicit edge direction information. Our proposed MRF model-based edge-directed interpolation method (MRF-EDI) is an implicit edge-directed interpolation method. In MRF-EDI, the edge directions of an edge pixel are indicated by the continuity strengths in all directions. Instead of labeling each direction as either edge or non-edge direction, we measure the continuity strength in each direction with a rational number between 0 and 1. Large values indicate strong continuities (along edge directions) while small values indicate weak continuities (across edge directions). These values are derived from the statistical properties of intensity variations within a local window. The relative continuity strengths of all the directions are used as edge direction information to formulate the geometric regularity (GR) spatial constraint, which can be summarized as smoothness along edge directions and sharpness across edge directions.

Another implicit edge-directed interpolation method that the proposed method is compared to is the New Edge-Directed Interpolation method (NEDI) [6]. In NEDI, the edge direction information is extracted as a low resolution covariance matrix and the interpolation of a high resolution covariance matrix from its low resolution counterpart is required. The edge direction information is reserved during the interpolation process. Compared to the conventional bilinear, bicubic and spline interpolation methods, the NEDI method shows impressive improvements of interpolated edges, which are less blur and smoother along edge directions. However, as a result of the fact that the interpolation of a pixel has been limited within its four nearest diagonal neighbors, the sharpness of the interpolated edges is not comparable to that of the original edges. Another limitation of this algorithm is that it has difficulties in dealing with textures. Typically, spurious minor edges are observed in the interpolated texture areas, which make the interpolated images look unnatural. As a result of losing fidelity of the original image, the interpolated image in the NEDI method suffers significant PSNR level decrease compared to the conventional methods. In this paper, we propose the MRF-EDI method to improve edges' geometric regularity further while maintaining the fidelity of the original image.

Review of MRF: Since MRF model is used in the MRF-EDI method, it is reviewed briefly here. MRF model provides a model to impose spatial constraints on the processed images and it proves to be a robust and accurate model for natural

This work is supported by a grant from Office of Navy Research. The emails of the authors are: {min_li, nguyent}@ece.ucsd.edu.

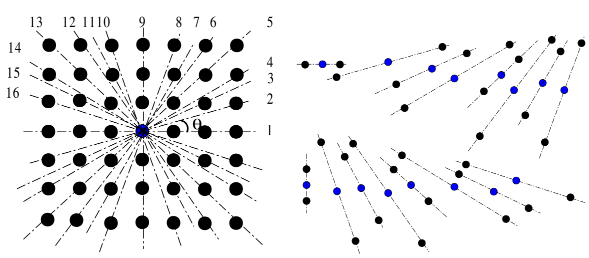
images [7]–[9]. The spatial constraint is imposed through the formulation of the energy function $U(\omega)$ in Gibbs distribution, $p(\omega) = \frac{1}{Z}\exp\{-\frac{U(\omega)}{T}\}$, where ω represents one configuration, T is a global control parameter of the convergent property and Z is a normalizing factor. Under the Maximum A Posterior (MAP)-MRF framework [10], the most desired configuration is the one that maximizes the probability $p(\omega)$ and, equivalently, minimizes the energy function $U(\omega)$.

The general modeling steps of spatial constraints in an MRF model are as follows. A neighborhood structure $\mathcal{N}_{i,j}$, which contains neighboring pixels of site (i,j) ((i,j) is not included), is first defined. Then a clique is defined on the neighborhood structure $\mathcal{N}_{i,j}$. A set of pixel sites c in $\mathcal{N}_{i,j}$ is a clique if all pairs of sites in c are neighbors. Lastly, a function V_c called potential function defines the interactions of pixel sites in clique c. Spatial constraints can be imposed on the processed image through the formulation of function V_c. The potential function is related to the energy function as $U(\omega) = \sum_{c\in\mathcal{C}} V_c(\omega)$.

Paper organization: The organization of the remaining of this paper is as follows. The implicit edge direction estimation algorithm is presented in Section II and the formulation of the GR spatial constraint is presented in Section III. Simulation results are presented and compared to other interpolation methods in Section IV. Section V concludes the paper.

II. Edge Direction Estimation

Suppose the high resolution image H comes directly from the low resolution image L as $H(2i-1, 2j-1) = L(i,j)$. For each pixel site in H, we consider sixteen discrete directions as indexed in Fig. 1(a). Each direction is represented by a vector



(a) (b)

Fig. 1: (a) The 7×7 neighborhood structure. Dots in black denote neighbor pixels. (b) The sixteen considered cliques.

$V = (v_r, v_c)$, which is the distance between the center pixel and its closest neighboring pixel in the corresponding direction. For example, direction 2 is represented by vector $(-1, 3)$. Each pixel in image H can be represented by one of the four coordinates, $H(2i-1, 2j-1), H(2i-1, 2j), H(2i, 2j-1)$ and $H(2i, 2j)$. We take pixel site $H(2i-1, 2j-1)$ as an example to present the calculation of the weights. Weights for other pixel sites can be obtained in a similar manner.

First a size 9×9 window W, which is centered at pixel site $H(2i-1, 2j-1)$, is formed from the high resolution image H and weights are learned from pixel intensity variations in the window W. Before calculating the pixel intensity variations, we temporarily interpolate the unavailable pixels in W by the bilinear method. This temporary interpolation is necessary. Otherwise, the intensity variations for some directions (e.g., directions 2, 6, 8, 10,12 and 16) are too few for the estimation of the statistical parameters. On the other hand, although bilinear interpolation blurs the original edges, edge directions are reserved during the interpolation.

When all pixels in W are available, the pixel intensity variations in all sixteen directions are calculated. Consequently, an intensity variation set for each direction is obtained as

$$\begin{aligned}\mathbf{\Delta I}(\mathbf{v_r}, \mathbf{v_c}) = \{&H(2(i+k)-1, 2(j+q)-1) - \\ &H(2(i+k)-1+v_r, 2(j+q)-1+v_c)\}, \\ &-1 \text{ or } -2 \le k, q \le 1 \text{ or } 2,\end{aligned} \tag{1}$$

where the exact range of k, q are determined by satisfying the condition that both $H(2(i+k)-1, 2(j+q)-1)$ and $H(2(i+k)-1+v_r, 2(j+q)-1+v_c)$ are in window W.

Let μ and σ^2 represent the mean and variance of $\mathbf{\Delta I}(\mathbf{v_r}, \mathbf{v_c})$, respectively, the weight for the discrete direction (v_r, v_c) is defined as $w = 1/\log_2(\sigma^2 + \mu^2)$, which can be interpreted as follows. The intensity variation samples that are obtained through (1) can be regarded as a short time realization of a wide sense stationary random process. The auto-correlation function $R(t)$ at $t = 0$ is used to represent the process and $\sigma^2 + \mu^2$ is the estimation for $R(0)$. The inverse operation is introduced so that weights are proportional to continuity strengths. The logarithm operation is used to keep all the sixteen weights of a pixel site in the same order of magnitudes. Otherwise, the weights decrease too quickly with respect to $R(0)$.

It can be concluded that the only case when the weight is significant is the one that both variance σ^2 and mean μ are small. This corresponds to cases where the directions are along edges. Consequently, the edge direction information of a pixel is indicated by a length-16 vector. Relatively large weights indicate relatively strong continuity.

III. MRF Model-Based Approach to Improve Edges' Geometric Regularity

This section discusses the formulation of the GR spatial constraint. The weights that are derived from Section II are incorporated in the constraint.

As reviewed in the introduction section, spatial constraints are imposed through the formulation of potential functions. The potential function in the proposed method is defined as

$$V_c(\Delta I_k) = w_k(-\gamma e^{-\frac{\Delta I_k^2}{\gamma}} + \gamma), \tag{2}$$

where w_k represents the weight in the kth discrete direction. Its calculation follows steps presented in Section II. ΔI_k represents the intensity variation in the kth discrete direction, which is

$$\begin{aligned}\Delta I_k \;=\; &\frac{1}{2}(|I(i+v_r, j+v_c) - I(i,j)| + \\ &|I(i,j) - I(i-v_r, j-v_c)|),\end{aligned} \tag{3}$$

where (v_r, v_c) denotes the kth discrete direction of pixel site (i, j). The neighborhood structure and cliques that are associated with the potential function (2) are shown in Fig. 1(a) and (b), respectively. The 48-pixel (7 × 7) neighborhood structure enables sixteen discrete edge directions. Sixteen three-pixel cliques are considered, one for each discrete direction.

The potential function (2) is obtained from a prototype function $g(\eta) = -\gamma e^{-\frac{\eta^2}{\gamma}}$, which is shown to be able to incorporate discontinuity-adaptive smoothness constraint in one-dimensional case [9]. Eq.(2) is a weighted extension of the prototype function. Potentials of all the sixteen directions are used to calculate the energy of the whole interpolated image. The GR spatial constraint will suppress high potentials along edge directions and reserve high potentials across edge directions. High potentials across edge directions might represent sharpness of edges.

The parameter γ in (2) is learned from the low resolution image. Its calculation can be summarized in the following two steps.

- Step 1: Intensity variations in the horizontal direction are calculated. $|\Delta I_h|$ denotes the data set, consisting of the magnitudes of the intensity variations and $|\Delta I_p|$ denotes samples, of which the magnitudes are among the top 10% of data set $|\Delta I_h|$.
- Step 2: Assume that $|\Delta I_p|$ represents significant discontinuity features in the image and let ΔI_t represent the minimal magnitude in set $|\Delta I_p|$, parameter γ is determined by relating a close-bound measure score to the intensity variation ΔI_t, for example,

$$-\gamma e^{-\frac{\Delta I_t^2}{\gamma}} + \gamma = .99\gamma. \tag{4}$$

After γ is obtained from (4), Simulated Annealing method [7] is used to search for the minimal energy state on the state space H. Note that the size of state space H can be large. Pixels in natural images can have any integer values between 0 and 255. Thus for an $M \times N$ high resolution image, the size of the realization space is $256^{\frac{3}{4} \times M \times N}$, which precludes any direct search methods. In order to lower the computational complexity, the following two strategies are proposed.

1) **Candidate sets***:* A candidate set for each pixel to be interpolated is proposed. With the proposed candidate set, each pixel to be interpolated can only have values from its candidate set. Thus the size of the state space that the minimal energy state is searched from reduces significantly. The candidate set is proposed based on the low resolution image and, basically, one candidate for each of the sixteen directions is proposed. If any two of the sixteen candidates are the same, only one of them is taken into the candidate set. Thus, plus one bicubic interpolation candidate, each pixel site will have at most seventeen candidates. The size of the state space H is reduced to a reasonable size.

Consider pixel site $(2i-1, 2j)$ as an example to show the formulation of its candidate set. As shown in Fig. 2(a), centered at the pixel site $(2i-1, 2j)$, a 7 × 7 window is formed in the high resolution image H. The candidates are proposed from pixels in this window. In discrete directions 1, 4, 7, 11 & 14, the two closest neighboring pixels of the center pixel are available and the average of the two neighboring pixels in the same direction is taken as one candidate pixel. As for the other directions, no closest neighboring pixels are available. In those cases, the pixels at the intersecting points (intersection of the direction and the central 4×2 square) are interpolated and their average is used as a candidate pixel. For example, direction 8 intersects the central square at points A and B. Pixel values at point A and B are first interpolated as $A = \frac{1}{8}I(2i-3, 2j-1) + \frac{7}{8}I(2i-3, 2j+1)$ and $B = \frac{7}{8}I(2i+1, 2j-1) + \frac{1}{8}I(2i+1, 2j+1)$, and then their average is used as a candidate pixel.

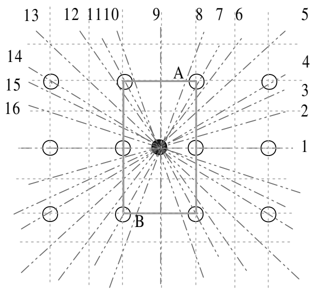
Proposing the candidates in a 7×7 local area is risky since 7×7 is relatively large. High frequency noises could be introduced easily if the intensity continuity of neighboring pixels is not maintained. In the proposed MRF-EDI method, the possible high frequency noises are suppressed efficiently by the GR spatial constraint.

2) **Discrimination between edge and non-edge pixels***:* To lower the complexity, only edge pixels are optimized iteratively while non-edge pixels are interpolated using deterministic interpolation methods, for example, bicubic. A pixel is identified as an edge pixel if only its two closest neighboring pixels in its strongest continuity direction (indicated by the largest weight) have their strongest continuities in similar directions. For example, if the strongest continuity of the pixel site being checked happens on direction d, its two closest neighboring pixels in direction d have to have their strongest continuities in directions d or $d \pm 1$ in order for this pixel to be declared as an edge pixel. Fig. 2(b)shows the identified edge pixels (in brightest intensity). It can be observed that major edges in the images are identified as edge pixels and optimized iteratively. Smooth or texture areas are not processed iteratively. The geometric regularity of major edges is improved iteratively during the iteration.

3) **Single-Pass Implementation***:* Adopting the one-pass algorithm in [11], we designed a single-pass implementation to replace the iterative optimization. The details are as follows. The low resolution image is initially interpolated using a convention method. Then for each edge pixel, the single-pixel related energy of each candidate is calculated and the one that has the minimal single-pixel related energy is the final output. Although no iteration is required in the single-pass implementation, its performance is almost as well as the iterative optimization. The main reason is that the initial state is very close to the global optimal state.

IV. Simulation Results

The parameter settings of the simulation are as follows. The low resolution images are obtained by directly downsampling the original images with a factor of two in both row and column dimensions. In the implementation of the proposed MRF-EDI method, the parameters are set as follows. The size of the local data window W is 9×9, in which the data is used to calculate weights of the sixteen discrete



(a) Candidate set for pixel site $(2i-1, 2j)$.

(b) Edge pixel map

Fig. 2: Circles represent pixels directly from the low resolution images.

(a) Original

(b) MRF-EDI (33.92 dB)

(c) NEDI (33.50 dB)

(d) Bilinear (32.36 dB)

Fig. 3: $4\times$ interpolation of "Foreman".

directions. The original value T is set as $T_0 = 50$. The parameters are set empirically and single pass implementation is used. Comparison results are shown in Fig. 3 and the local area zoom-in comparisons are shown in Fig. 4. One can observe that the major edges interpolated by the proposed MRF-EDI method are sharpest. More simulation results are available at `http://videoprocessing.ucsd.edu/~minli/interpolationdemo.htm`.

V. Conclusions

An implicit edge-directed interpolation algorithm for natural images is proposed in this paper. In this method, the interpolated image is modeled as an MRF and the most desired interpolated image is related to the minimal energy state of a two-dimensional random field. Edge direction information is incorporated in the formulation of the energy function in

(a) Original

(b) MRF-EDI

(c) NEDI

(d) Bilinear

Fig. 4: Zoom-in comparison.

the MRF model. Consequently, energy that is along the edge direction is strongly suppressed to achieve smoothness while energy that is across the edge direction is much less suppressed to maintain the sharpness of the edge. The interpolated edges have strong geometric regularity.

References

[1] S. Lee and J. Paik, "Image interpolation using adaptive fast B-spline filtering," *IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 5, pp. 177–180, Apr. 1993.

[2] A. Gotchev, K. Egiazarian, J. Vesma, and T. Saramaki, "Edge-preserving image resizing using modified B-splines," *IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 3, pp. 1865–1868, May 2001.

[3] K. Jensen and D. Anastassiou, "Subpixel edge localization and the interpolation of still images," *IEEE Trans. on Image Processing*, vol. 4, pp. 285–295, Mar. 1995.

[4] J. Allebach and P. Wong, "Edge-directed interpolation," *IEEE International Conference on Image Processing*, vol. 3, pp. 707–710, Sept. 1996.

[5] J. H. Elder and S. W. Zucker, "Local scale control for edge detection and blur estimation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 7, pp. 699–716, July 1998.

[6] X. Li and M. T. Orchard, "New edge-directed interpolation," *IEEE Trans. on Image Processing*, vol. 10, no. 10, pp. 1521–1526, 2001.

[7] S. Geman and D. Geman, "Stochastic relaxation, gibbs distribution, and the bayesian restoration of images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721–741, Nov. 1984.

[8] J. L. Marroquin, E. Arce, and S. Botello, "Markov random measure fields for image analysis," *IEEE International Conference on Image Processing*, vol. 1, pp. 765–768, 2002.

[9] S. Z. Li, "On discontinuity-adaptive smoothness priors in computer vision," *IEEE Trans. on Pattern and Machine Intelligence*, vol. 17, no. 6, pp. 576–586, June 1995.

[10] M. Li and T. Nguyen, "Discontinuity-adaptive de-interlacing scheme using markov random field model," *IEEE International Conference on Image Processing*, pp. 393–396, Oct. 2006.

[11] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based super-resolution," *IEEE Transaction on Computer Graphics and Application*, vol. 22, no. 2, pp. 56–65, Apr. 2002.