IMAGE INTERPOLATION USING ACROSS-SCALE PIXEL CORRELATION

Tao Chen, Hong Ren Wu, and Bin Qiu

School of Computer Science and Software Engineering
Monash University, Clayton, VIC 3168, Australia
{tchen,hrw,bq}@csse.monash.edu.au

ABSTRACT

In this paper, a novel method is proposed for image interpolation. It is assumed that the pixel correlation between local regions across scales would remain similar. In addition, this *a priori* similarity could be extracted from a set of available image data that have the same content but different resolutions. A simple architecture is devised to estimate the correlation efficiently, which is then used to predict the unknown pixel values in a high-resolution image. Evaluation shows a promising performance of the proposed algorithm.

1. INTRODUCTION

Image interpolation pervades many areas, such as surveillance, medical imaging, automatic target recognition, and consumer electronics [1, 2]. In multimedia applications, for instance, images of high resolution require a large media space for their storage, or a large bandwith/long duration for their transmission. In addition, low-resolution images can often be retrieved by a user with relative ease for browsing scenes of interest. In order to make image browsing more efficient, it is also important for a user to be able to perceive the available data with their richest and visually most pleasing way. In this case and others, effective interpolation techniques are highly desirable.

With ideal linear and stationary systems, the *sinc* function is an optimal means for interpolation. It permits exact reconstruction of a bandlimited signal; however, it is impossible to realize this function physically. Furthermore, images in general contain sharp edges and other detailed features, and thus they are not strictly bandlimited [2]. To combat these, different approximations have been devised for interpolation. Conventionally, the amplitude of an unknown or missing pixel in a high-resolution image is estimated based on its known neighbors available from the low-resolution version. The well-known bilinear and bicubic [1] interpolation methods are the simple forms in this class. The common case of image magnification by an integer scaling factor can also be achieved through convolution of the zero-interleaved image pixels with a single kernel, such as the pyramid and cubic B-spline kernels [1].

Due to the ill-posed nature of interpolation or super-resolution problem [2], additional information apart from the collected data is required in order to obtain a desirable solution. Usually, reasonable assumptions about the characteristics of a true image are made, and one of the most common assumptions is smoothness. Recently, interpolation approaches based on *a priori* knowledge of natural images have been the subject of active research. Image interpolation is dealt with through the constrained minimization of a function in variational schemes [3, 4]. Directional approaches [5, 6] evaluate local edge information in an image and then conduct interpolation along the edge rather than across the edge. POCS (projection onto convex sets) methods solve the problem using a regularized iterative projection [7]. The multiresolution model in [8] provides a solution to image magnification by enhancing the result of the bilinear interpolation with an adaptive filter.

This paper proposes a novel technique by assuming that the pixel correlation of local structures (or small regions) would remain similar across different scales. In addition, this *a priori* information could be extracted from the available or given image data that have the same scene (content) but different resolutions. As such, the problem now is to devise an interpolation scheme which is able to capture the pixel correlation across scales. A simple architecture is formulated to fulfill this purpose efficiently. Interpolation is then achieved by predicting the values of unknown image pixels with the use of the estimated correlation.

The remainder of the paper is organized as follows. Section 2 discusses the new image interpolation scheme and its implementation. Simulation results are presented in Section 3. Finally, Section 4 concludes the paper.

2. THE PROPOSED ALGORITHM

Let $\mathcal{I}(i_1, i_2)$ denote the pixel values of image $\mathcal{I}$, where i_1 and i_2 are the coordinates in vertical and horizontal directions, respectively. Also, let H and W be the image height and width, respectively, i.e., $1 \le i_1 \le H$ and $1 \le i_2 \le W$. The problem of image interpolation can be viewed as to obtain a high-resolution image $\mathcal{I}_h$ of size $S_1 H \times S_2 W$, where

$S_1 > 1$ and $S_2 > 1$ denote the scaling factors in vertical and horizontal directions, respectively. Integer scaling factors are considered in this work, i.e., $S_1, S_2 = 2, 3, \cdots$.

In order to predict the unknown sample values in image $\mathcal{I}_h$ from the pixels in $\mathcal{I}$ using the local pixel correlation across scales, the pixel correlation of small regions between a low-resolution image $\mathcal{I}_l$ and image $\mathcal{I}$ is estimated first. In particular, the low-resolution image $\mathcal{I}_l$ is of size $H/S_1 \times W/S_2$. Ideally, the low-resolution image would encompass the same scene as in $\mathcal{I}$ and has been physically generated by hardware with a known resolution setting. Since such a data collection is not common, the low-resolution version of image $\mathcal{I}$ is simulated by decimation instead. The image acquisition model discussed in [4] is taken here. A neighborhood is defined in image $\mathcal{I}$ as follows, especially corresponding to the pixel (i_1, i_2) in image $\mathcal{I}_l$:

$$\mathcal{N}_{i_1,i_2} = \left\{ (u_1, u_2) \middle| \begin{array}{l} S_1 i_1 \le u_1 \le S_1(i_1+1) - 1, \\ S_2 i_2 \le u_2 \le S_2(i_2+1) - 1 \end{array} \right\}. \tag{1}$$

Then, the low-resolution version is obtained by

$$\mathcal{I}_l(i_1, i_2) = \frac{1}{S_1 S_2} \sum_{(u_1,u_2) \in \mathcal{N}_{i_1,i_2}} \mathcal{I}(u_1, u_2). \tag{2}$$

In this case, each low-resolution pixel value $\mathcal{I}_l(i_1, i_2)$ corresponds to the average of the samples located within neighborhood $\mathcal{N}_{i_1,i_2}$ in its high-resolution version, i.e., image $\mathcal{I}$.

With respect to the similarity of pixel correlation across scales, overlapping $R_1 \times R_2$ regions in the low-resolution image are first considered. Here, $R_1 = 2r_1 + 1$ and $R_2 = 2r_2 + 1$ (r_1 and r_2 are non-zero integers) both are odd for the concern of symmetry in implementation. Specifically, a local region is defined at each position (i_1, i_2) in the low-resolution image as

$$\mathcal{R}_{i_1,i_2} = \left\{ (u_1, u_2) \middle| \begin{array}{l} i_1 - r_1 \le u_1 \le i_1 + r_1, \\ i_2 - r_2 \le u_2 \le i_2 + r_2 \end{array} \right\}. \tag{3}$$

The sample vector observed via this region is given by

$$\begin{aligned} \mathbf{X}_{i_1,i_2} = [\, & \mathcal{I}_l(i_1 - r_1, i_2 - r_2), \\ & \mathcal{I}_l(i_1 - r_1, i_2 - r_2 + 1), \cdots, \\ & \mathcal{I}_l(i_1 - r_1, i_2 + r_2), \cdots, \\ & \mathcal{I}_l(i_1 + r_1, i_2 + r_2) \,]^T, \end{aligned} \tag{4}$$

where the pixels are rearranged in a lexicographic order, i.e., row by row. In the above equation, $\mathbf{X}_{i_1,i_2}$ is a $p \times 1$ vector, where $p = R_1 \times R_2$.

Further, each $R_1 \times R_2$ region in the low-resolution image $\mathcal{I}_l$ is paired up with its homologous $(2S_1 - 1) \times (2S_2 - 1)$ region in its high-resolution counterpart, i.e., image $\mathcal{I}$. The high-resolution region in $\mathcal{I}$ is described specifically by

$$\tilde{\mathcal{R}}_{i_1,i_2} = \left\{ (u_1, u_2) \middle| \begin{array}{l} S_1 i_1 - S_1 + 1 \le u_1 \le S_1 i_1 + S_1 - 1, \\ S_2 i_2 - S_2 + 1 \le u_2 \le S_2 i_2 + S_2 - 1 \end{array} \right\}. \tag{5}$$

Correspondingly, the sample vector of this region is obtained in a similar way as in Eq.(4) and is denoted as $\mathbf{Y}_{i_1,i_2}$, which is a $q \times 1$ vector and $q = (2S_1 - 1) \times (2S_2 - 1)$.

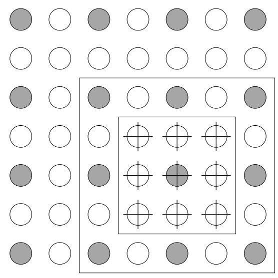

Figure 1: Local regions of different resolutions where $R_1 = R_2 = 3$ and $S_1 = S_2 = 2$. The shaded circles in the large square denote the pixels of a low-resolution region, while the crossed circles in the small square are those of a high-resolution region.

The aforementioned pair of low and high resolution regions are shown in Figure 1, where the shaded circles together denote a low-resolution image and the white circles are subject to interpolation to form a high-resolution version. In the case as shown in the figure, the values of those crossed circles around the center of the high-resolution region (i.e., the small square) are estimated based on the values of the shaded circles in the large square available from the low-resolution image. That is, each group of $R_1 \times R_2$ low-resolution circles (shaded) are used to reconstruct the $(2S_1 - 1) \times (2S_2 - 1)$ high-resolution circles (crossed) around the same center. In this case, the interpolated pixels that overlap are subject to be averaged to form a high-resolution image. This averaging operation improves the reliability of the final high-resolution samples. For the center pixel (shaded and crossed), one can choose either to replace it with a new value or to leave it unaltered. Extension to other scaling factors and region size is straightforward.

The correlation between the two sample vectors of the low and high resolution regions is characterized as follows.

$$\mathbf{Y}_{i_1,i_2} = \mathbf{W}\mathbf{X}_{i_1,i_2}, \tag{6}$$

where $\mathbf{W}$ is a weighting matrix of size $q \times p$. The weighting matrix can be thought of as a convolution kernel as in the conventional methods [1]. In other words, with an established estimation function, each region as defined previously in a given image is multiplied by $\mathbf{W}$ to generate a high-resolution region. Those estimated high-resolution regions are then overlapped and averaged to reconstruct a high-resolution image. The reconstruction kernel $\mathbf{W}$ is obtained as follows.

Consider a practical scenario where image $\mathcal{I}$ is given and is to be interpolated to form a high-resolution image. Based on the decimation model in Eq.(2), its low-resolution version is obtained as $\mathcal{I}_l$. With this pair of images available, the convolution kernel is optimized by exploiting the pixel correlation between all the paired up regions observed from the two images. In this case, the given image $\mathcal{I}$ can be viewed as the high-resolution counterpart of image $\mathcal{I}_l$.

Now, given the sample vector, $\mathbf{X}_{i_1,i_2}$, of a low-resolution region in $\mathcal{I}_l$, its corresponding high-resolution region is reconstructed or predicted by

$$\hat{\mathbf{Y}}_{i_1,i_2} = \mathbf{W}\mathbf{X}_{i_1,i_2}. \tag{7}$$

The optimal weighting matrix is sought to minimize the mean-squared error (MSE), $\mathcal{E}$, which is given as follows.

$$\mathcal{E} = E\left\{\left[\mathbf{Y}_{i_1,i_2} - \hat{\mathbf{Y}}_{i_1,i_2}\right]^2\right\}, \tag{8}$$

where $\mathbf{Y}_{i_1,i_2}$ is the sample vector of the region available from the given image $\mathcal{I}$. One can take the partial derivative of $\mathcal{E}$ with respect to each weighting coefficient in $\mathbf{W}$, and consider the instantaneous gradient of this partial derivative at the current location. Using the steepest gradient method, the weighting matrix can be derived recursively via the well-known least mean square (LMS) algorithm [9]. The updating formula of LMS recursion is then given by

$$\mathbf{W}' = \mathbf{W} + \mu\left(\mathbf{Y}_{i_1,i_2} - \hat{\mathbf{Y}}_{i_1,i_2}\right)\mathbf{X}^T_{i_1,i_2}, \tag{9}$$

which is conducted at pixel location (i_1, i_2). Here, μ denotes the learning rate. Clearly, the update at the next pixel location is conducted based on $\mathbf{W}'$. The simple raster scanning across the image is assumed in this work.

3. SIMULATION RESULTS

The performance of the proposed technique has been evaluated for image magnification. For comparison, the bilinear, bicubic and cubic B-spline interpolators, are also simulated. Interpolation performance is quantitatively measured by the *peak signal-to-noise ratio* (PSNR), which is defined as

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{\sum_{(i_1,i_2)} 255^2}{\sum_{(i_1,i_2)}\left[\mathcal{I}(i_1,i_2) - \hat{\mathcal{I}}(i_1,i_2)\right]^2}\right)\mathrm{dB},$$

where $\mathcal{I}$ and $\hat{\mathcal{I}}$ denote the original and the interpolated high-resolution images, respectively.

In our experiments, the scaling factors $S_1 = S_2 = 2$. In other words, each of the 512×512 test images is first subsampled by decimation to a 256×256 low-resolution version. Then the decimated image is interpolated back to the size of 512×512.

Concerning the derivation of weighting matrix in the proposed algorithm, three cases are considered as follows: (i) the weights employed in interpolating a 256×256 image are derived with the use of the its original 512×512 version; (ii) the weights are derived with the use of the given 256×256 image that is to be interpolated; (iii) the weights derived using the 512×512 *Lena* image are employed to interpolate the other images, and the *Lena* image is interpolated using the weights derived from the 512×512 *Man* image. Note that, the first case represents an impractical scenario, while the latter two are practically feasible.

Table 1: Comparative results in PNSR (dB).

Filters	Images		
	Lena	*Man*	*Peppers*
Bilinear	29.80	28.27	28.90
Bicubic	30.05	28.50	29.07
Cubic B-spline	29.83	28.08	29.35
Proposed (i)	32.17	30.39	30.40
Proposed (ii)	32.03	30.31	30.31
Proposed (iii)	32.16	30.34	30.33

The PSNR results of performance comparison are presented in Table 1. It is apparent that the proposed technique yields noticeable gains over the other methods. The performance of the proposed approach in cases (ii) and (iii) has not degraded too much as compared with that in case (i). Its satisfactory performance in case (ii) indicates that the proposed algorithm is well able to capture the pixel correlation between images with the same content across scales. The good results in case (iii) demonstrate the robustness of the new method in interpolating images other than that used in deriving the weights. In addition, the proposed technique produces visually more pleasing result. Some of the resultant *Man* images are shown in Figure 2, where the image interpolated using the new method looks sharper than the results provided by the other methods.

4. CONCLUSIONS

A new image interpolator is formulated by assuming that similar pixel correlation of regions across scales could be extracted and predicted. In addition, a simple architecture is proposed to capture this similarity which is used to estimate the values of missing pixels in a high-resolution image.

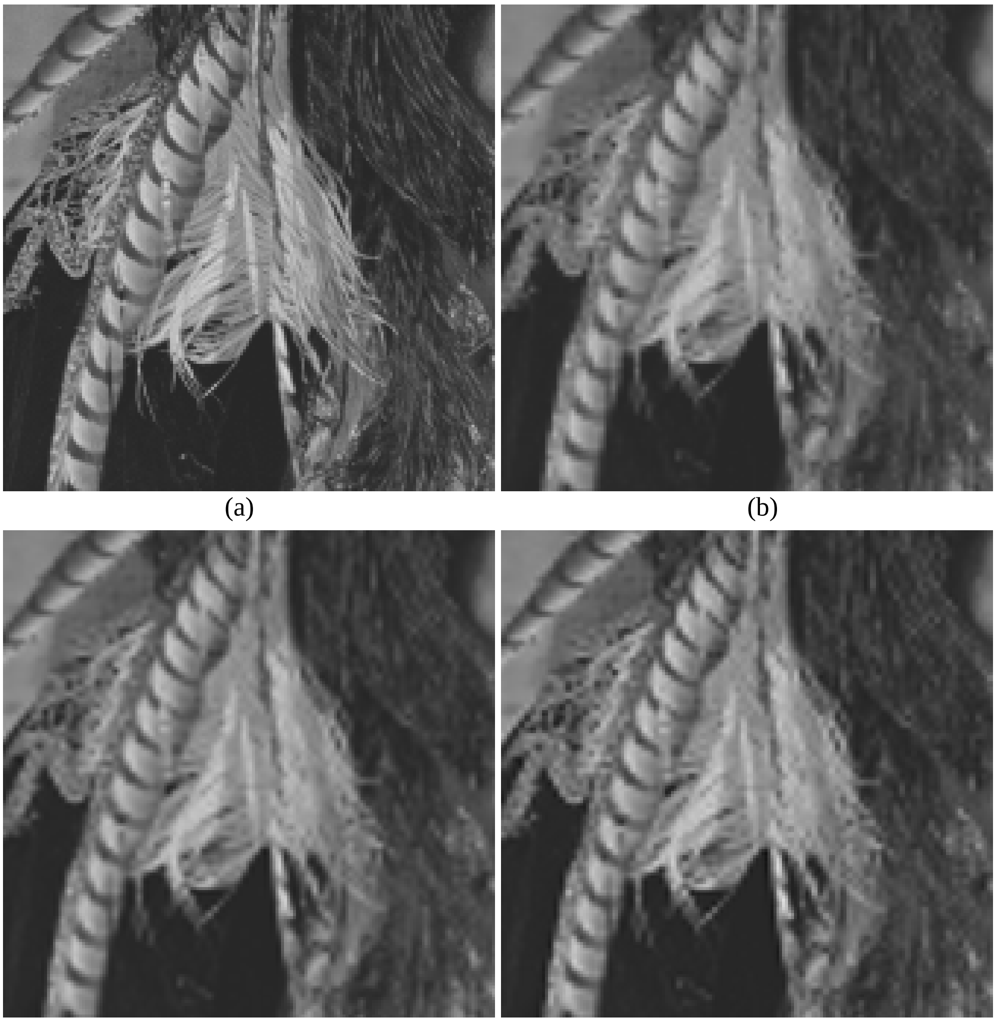

Figure 2: The enlarged portions of (a) original *Man* image, and the resultant images by (b) bilinear (28.27dB), (c) bicubic (28.50dB), and (d) proposed (30.31dB) interpolation algorithms, respectively.

Promising results have been produced, though blurring was inevitably brought in by the new method due to the linear operation. Local signal attributes are to be considered to form an adaptive, location-variant, scheme. Further analysis of the new technique is also under investigation.

5. REFERENCES

[1] W. K. Pratt, *Digital Image Processing.* 2nd ed., Wiley, New York, 1991.

[2] A. M. Tekalp, *Digital Video Processing.* Prentice Hall, Englewood Cliffs, NJ, 1995.

[3] N. B. Karayianni and A. N. Venetsanopoulos, “Image interpolation based on variational principles,” *Signal Processing*, vol. 25, pp. 259–288, 1991.

[4] R. R. Schultz and R. L. Stevenson, “A Bayesian approach to image expansion for improved definition,” *IEEE Trans. Image Processing*, vol. 3, no. 3, pp. 233–242, 1994.

[5] S. D. Bayrakeri and R. M. Mersereau, “A new method for directional image interpolation,” in *Proc. ICASSP*, vol. IV, pp. 2383–2386, 1995.

[6] K. Jesen and D. Anastassiou, “Sub-pixel edge localization and the interpolation of still images,” *IEEE Trans. Image Processing*, vol. 4, no. 3, pp. 285–295, 1995.

[7] K. Ratakonda and N. Ahuja, “POCS based adaptive image magnification,” in *Proc. ICIP*, vol. III, pp. 203–207, 1998.

[8] F. A. C. M. Cardoso, D. S. Arantes, and J. M. Romano, “A multiresolution model for image interpolation with adaptive filtering,” in *Proc. ICASSP*, vol. IV, pp. 2203–2206, 2000.

[9] S. Haykin, *Neural Networks: A Comprehensive Foundation.* 2nd ed., Prentice-Hall, NJ, 1999.