

NEIGHBORHOOD ISSUE IN SINGLE-FRAME IMAGE SUPER-RESOLUTION

Kevin (Xu) Su¹, Qi Tian¹, Qing Xue¹, Nicu Sebe², Jingsheng Ma³

¹Department of Computer Science, University of Texas at San Antonio, TX 78249
{xsu, qitain, qxue}@cs.utsa.edu

² Faculty of Science, University of Amsterdam, The Netherlands, nicu@science.uva.nl

³ Institute of Petroleum Engineering, Heriot-Watt University, Edinburgh, UK,
Jingsheng.ma@pet.hw.ac.uk

ABSTRACT

Super-Resolution is the problem of generating one or a set of high-resolution images from one or a sequence of low-resolution frames. Most methods have been proposed for super-resolution based on multiple low resolution images of the same scene, which is called multiple-frame super-resolution. Only a few approaches produce a high-resolution image from a single low-resolution image, with the help of one or a set of training images from scenes of the same or different types. It is referred to as single-frame super-resolution. This article reviews a variety of single-frame Super-Resolution methods proposed in the recent years. In the paper, a new manifold learning method: locally linear embedding (LLE) and its relation with single-frame super-resolution is introduced. Detailed study of a critical issue: “*Neighborhood Issue*” is presented with related experimental results and analysis. And possible future research is given.

1. INTRODUCTION

With the development of image processing applications, there is a big demand for high-resolution (HR) images since HR images not only give the viewer a pleasing picture but also offer additional detail that is important for the analysis in many applications. The current technology to obtain HR images mainly depends on sensor manufacturing technology that attempts to increase the number of pixels per unit area by reducing the pixel size. However, the cost for high-precision optics and sensors may be inappropriate for general purpose commercial applications, and there is a limitation to pixel size reduction due to shot noise encountered in the sensor itself. Therefore, a resolution enhancement approach using signal processing techniques to obtain one or a set of HR images from observed single or a sequence of low resolution (LR) frames has been a great concern in many areas, and it is called super-resolution (SR).

Super-Resolution image reconstruction is proved to be useful in many practical cases including medical imaging, satellite imaging, video applications, image enlarging in web pages, and restoration of old historic photographs etc.

Most methods have been proposed for super-resolution based on multiple low resolution images of the same scene, which is called *multiple-frame super-resolution*. Only a few approaches have been proposed to generate a high-resolution image from a single low-resolution image, with or without the help of one or a set of training images from scenes of the same or different types. We refer to this as *single-frame super-resolution* problem.

Although several articles have surveyed the different classical multiple-frame super-resolution methods and compared their performances, till now there is no article that reviews the works on single-frame super-resolution presented in the recent years. The intention of this article is to pinpoint the various difficulties inherent to single-frame super-resolution problems, especially neighborhood issue. For the purpose of this article, super-resolution will refer to the single-frame image super-resolution problem in the rest of this paper.

The organization of this article is as follows. In Section 2, we will discuss related previous work. Section 3 describes the relation between locally linear embedding (LLE) and super-resolution, and the common *neighborhood issue* in learning-based approaches. Section 4 shows related experiments and analysis, and Section 5 concludes the paper.

2. RELATED WORK

Simple resolution enhancement methods based on smoothing and interpolation techniques for noise reduction have been commonly used in image processing. Smoothing is usually achieved by applying various spatial filters such as Gaussian, Wiener, and median filters.

Commonly used interpolation methods, such as bicubic and cubic spline interpolation [1, 2], approximate an unknown continuous function by a set of local functions, which can then be discretely sampled to the desired resolution. Interpolation methods usually give better performance than simple smoothing methods. While both methods are usually fast and easy to implement, the resulting images often have blurred edges.

Image sharpening techniques have been proposed to ameliorate the results from functional interpolation methods [3, 4]. These methods result in sharper images, but may contain haloing artifacts.

An alternate solution involves deconvolving the blurring filter [5, 6, 7, 8, 9]. While the results are quite good, deconvolution methods, as well as image sharpening methods, only enhance features that are present in the low resolution image.

More recently, some learning-based methods have been proposed by different researchers [11, 12, 13, 14, 15, 16]. Prior information is used to enhance the solution space. These methods add realistic image details, but can also add artifacts. Graphical models are also designed for image super-resolution [10]. The flexibility of the graphical model incorporates the statistics of natural images to define compatibility functions between the high-resolution pixels being estimated.

3. LLE AND NEIGHBORHOOD ISSUE

3.1. Problem formulation

The single-frame super-resolution problem can be formulated as follows. Given a low-resolution image X_L as input, the desired high-resolution image X_H is estimated with the help of a set of training low resolution images Y_L^i and the corresponding high-resolution images Y_H^i , where $i = 1..M$, M is the number of training image pairs.

Each low- or high-resolution image can be represented as a set of small image patches with or without overlap. Each low-resolution image Y_L^i and the corresponding high-resolution image Y_H^i have the same number of patches. We denote the low-resolution patches in X_L and desired high-resolution patches in X_H as $\{x_L^p\}_{p=1}^{N_x}$ and $\{x_H^p\}_{p=1}^{N_x}$ respectively, where N_x is number of patches in input image. Low- and high-resolution patches in the training image pairs can be represented as $\{y_L^q\}_{q=1}^{N_y}$ and $\{y_H^q\}_{q=1}^{N_y}$ respectively, where N_y is the total number of patches in the training images. Obviously, N_x and N_y depend on the patch size and the degree of overlap between adjacent patches.

In super-resolution, for each low-resolution patch x_L^p from input image X_L , we need to estimate x_H^p with the help of training image patches $\{y_L^q\}_{q=1}^{N_y}$ and $\{y_H^q\}_{q=1}^{N_y}$.

The intuitive way is to find the y_H^q which is most similar to x_H^p from $\{y_H^q\}_{q=1}^{N_y}$, and let $x_H^p = y_H^q$, or find several high-resolution patches in $\{y_H^q\}_{q=1}^{N_y}$ that are closest to x_H^p in high-resolution space, and estimate x_H^p from these patches, which are called nearest neighbors of x_H^p . However, the nearest neighbors of x_H^p in $\{y_H^q\}_{q=1}^{N_y}$ can not be easily determined since we don't know x_H^p and can not use some kind of distance metric to determine these nearest neighbors. This is called *neighborhood issue* in this paper.

Chang et al. [11] tried to address neighborhood issue based on the new manifold learning method: locally linear embedding (LLE).

3.2. LLE

LLE is a promising manifold learning method that has aroused a lot of interest in machine learning. It computes low-dimensional, neighborhood-preserving embeddings of high-dimensional inputs and recovers the global nonlinear structure from locally linear fits.

High-resolution image patches can represent points in a high dimensional data space, and the corresponding low-resolution image patches are points in a low dimensional data space. In this case, LLE is to estimate low-resolution patches, given the corresponding high-resolution patches, and LLE can be thought as a reverse procedure of solving super-resolution problems. In LLE, nearest neighbors of data point in high dimensional space will still be nearest neighbors in low dimensional space. The neighborhood will be preserved in both high and low dimensional spaces. Chang et al.'s work [11] is based on the assumption that the high dimensional data points whose corresponding low dimensional data points are neighbors will still be neighbors in high dimensional space. However in super-resolution, this is not true. If the low-resolution image patches are in neighborhood, their corresponding high-resolution patches are not in neighborhood anymore, and vice versus. Figure 1 illustrates that the corresponding high-resolution patches of 4-nearest neighbors of input low-resolution patch x are not the same as 4-nearest neighbors of desired high-resolution patch y .

4. EXPERIMENTS AND ANALYSIS

4.1. Feature representation

In our experiments, we use the YIQ color model where Y channel represents luminance and I and Q channels represent chromaticity, since humans are more sensitive to changes in luminance than to changes in color. And only the luminance values from the Y channel are used to define features.

For the high-resolution image patches, the feature of each patch is represented by its luminance values of the pixels in the patch.

For the low-resolution image patches, we use two schemes to define the feature vector. The first method is to define the feature vectors for each patch based on the luminance values of the pixels in the patch. The second method is to consider the relative luminance changes within a patch, same as in [11]. More specifically, we use the first-order and second-order gradients of the luminance as features.

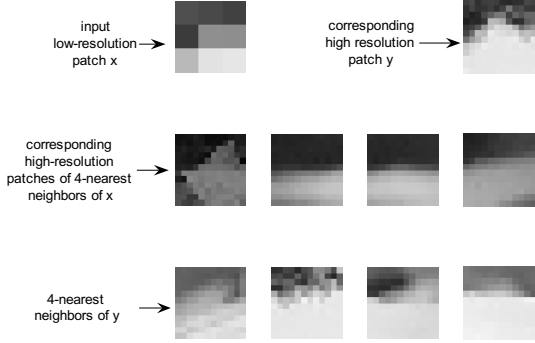


Figure 1. Neighborhood preservation example

4.2. Testing image set generation

To generate our testing image set, we start from a collection of high-resolution botanic images and degraded each of them to obtain corresponding low-resolution images by blurring and downsampling. Some of testing images are shown in Figure 2. Then each image pairs are divided into small patches.

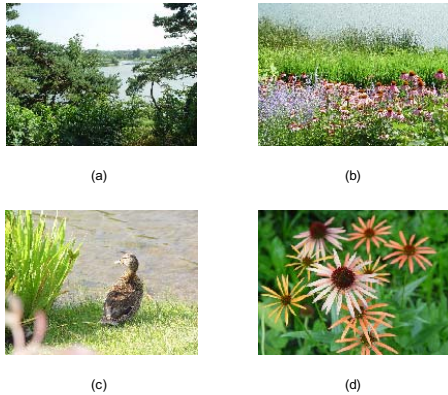


Figure 2. Sample testing images

In order to consider the effect of downsampling factors and downsampling methods on neighborhood preservation, we use 3 downsampling factors: 2, 4 and 8

in our experiments. And we also test different downsampling methods, such as arithmetic mean (A), median (M), harmonic mean (H), geometric mean (G), random selection (R), and specific position (S).

4.3. Experimental results

First randomly select 1000 low-resolution image patches from testing set. For each low-resolution patch y_L^q , we find its K-nearest neighbors among all the low-resolution patches in the testing set, and find the K-nearest neighbors of the corresponding high-resolution patch y_H^q from all the high-resolution patches. After comparison, find out how many corresponding high-resolution patches of K-nearest neighbors belong to K-nearest neighbors of y_H^q , average this number for 1000 tests, then we get K-nearest neighborhood preservation rate. In our experiments, the default patch size is 3x3, patch overlap size is 1, downsampling factor is 2, and downsampling method is arithmetic mean (A).

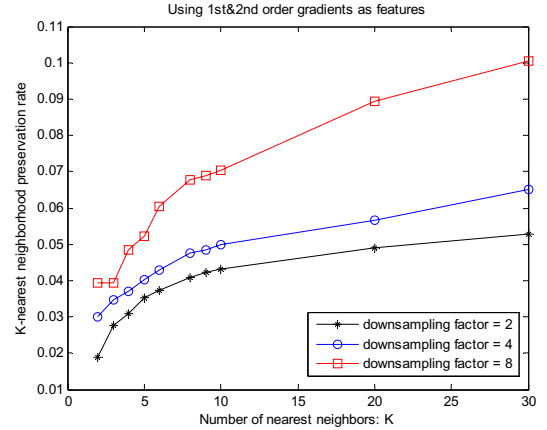


Figure 3. Neighborhood preservation with different downsampling factors

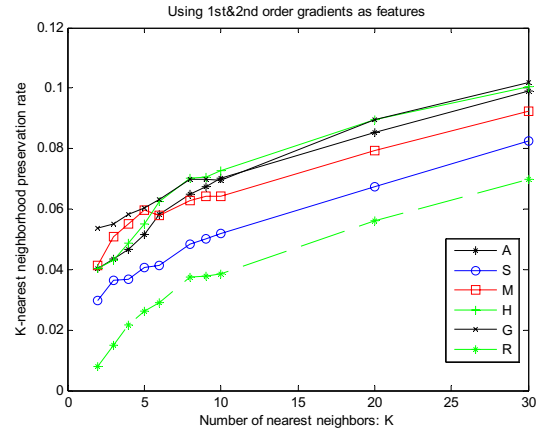


Figure 4. Neighborhood preservation with different downsampling algorithms

Figure 3 and Figure 4 study the effect on neighborhood preservation using different downsampling factors and different downsampling methods respectively. In both cases, first-order and second-order gradients of the luminance values of the pixels are used as features for low-resolution patches. In Figure 5 and Figure 6, the feature of each patch is represented by its luminance values of the pixels in the patch. From all these figures, we can see that all neighborhood preservation rates are small, although using luminance values of the pixels as features gives better neighborhood preservation.

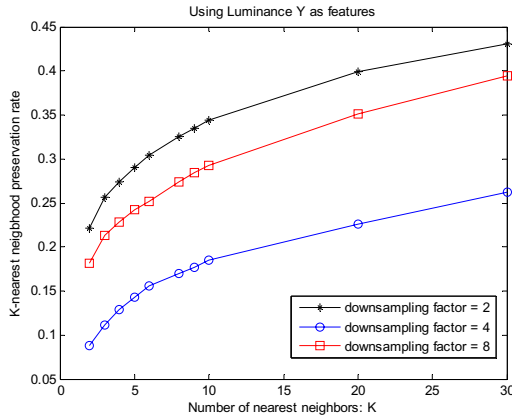


Figure 5. Neighborhood preservation with different downsampling factors

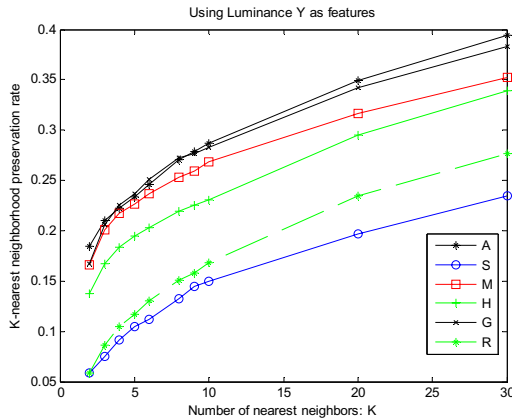


Figure 6. Neighborhood preservation with different downsampling algorithms

5. CONCLUSION AND FUTURE WORK

In this paper, we review a variety of single-frame Super-Resolution methods proposed in the recent years. The new manifold learning method locally linear embedding (LLE) and its relation with super-resolution is introduced. And a critical issue: “Neighborhood Issue” is presented and discussed. Related experimental results and analysis are given to show that neighborhood preservation for low- and high-resolution patches rarely holds.

In order to increase neighborhood preservation, there are two possible ways. One is to select a good feature to represent image patch, which can preserve neighborhood better. The other is to select a good reconstruction function given some high-resolution neighbor patches.

REFERENCES

- [1] H.H. Hou and H.C. Andrews, “Cubic splines for image interpolation and digital filtering,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26(6):508-517, 1978.
- [2] R.G. Keys, “Cubic convolution interpolation for digital image processing,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 29(6):11153-1160, 1981.
- [3] H. Greenspan, C. Anderson, and S. Akber, “Image enhancement by nonlinear extrapolation in frequency space,” *IEEE Transactions on Image Processing*, 9(6), 2000.
- [4] B. Morse and D. Schwartzwald, “Image magnification using level set reconstruction,” *Proc. of Int’l Conf. on Computer Vision*, pp. 333-341, 2001.
- [5] M. Belge, M. Kilmer, and E. Miller, “Wavelet domain image restoration with adaptive edge-preserving regularity,” *IEEE Transactions on Image Processing*, 9(4):597-608, 2000.
- [6] Y. Zhu, S.C. Schwartz, M.T. Orchard, “Wavelet domain image interpolation via statistical estimation,” *Proc. of IEEE Int’l Conf. on Image Processing*, 2001.
- [7] Mário A. T. Figueiredo, and Robert D. Nowak, “An EM algorithm for wavelet-based image restoration,” *IEEE Transactions on Image Processing*, 2003.
- [8] Yi Wan, R.D. Nowak, “A wavelet-based statistical model for image restoration,” *Proc. of IEEE Int’l Conf. on Image Processing*, 2001.
- [9] Shubin Zhao, Hua Han and Silong Peng, “Wavelet-domain HMT-based image superresolution,” *Proc. of IEEE Int’l Conf. on Image Processing*, 2003.
- [10] Marshall F. Tappen, Bryan C. Russell, W.T. Freeman, “Efficient Graphical Models for Processing Images,” *Proc. of IEEE Conf. CVPR*, 2004.
- [11] Hong Chang, Dit-Yan Yeung, Yimin Xiong, “Super-Resolution Through Neighbor Embedding,” *Proc. of IEEE Conf. CVPR*, 2004.
- [12] W.T. Freeman, T.R. Jones, and E.C. Pasztor, “Example-based superresolution,” *IEEE Computer Graphics and Applications*, 2002.
- [13] W.T. Freeman, E.C. Pasztor, and O.T. Carmichael, “Learning low-level vision,” *International Journal of Computer Vision*, 2000.
- [14] Simon Baker and Takeo Kanade, “Hallucinating Faces,” *Fourth International Conference on Automatic Face and Gesture Recognition*, 2000.
- [15] C. Liu, H.Y. Shum, and C.S. Zhang, “A two-step approach to hallucinating faces: global parametric model and local nonparametric model,” *Proc. of IEEE Conf. CVPR*, 2001.
- [16] J. Sun, N.N. Zheng, H. Tao, and H.Y. Shum, “Image hallucination with primal sketch priors,” *Proc. of IEEE Conf. CVPR*, 2003.