

An Improved Super-Resolution with Manifold Learning and Histogram Matching

Tak Ming Chan¹ and Junping Zhang^{1,2}

¹ Shanghai Key Laboratory of Intelligent Information Processing,
Department of Computer Science and Engineering,
Fudan University, 200433, China
{0272366, jpzhang}@fudan.edu.cn

² The Key Laboratory of Complex Systems and Intelligence Science,
Institute of Automation, Chinese Academy of Sciences,
Beijing, 100080, China

Abstract. Biometric Person Authentication such as face, fingerprint, palmprint and signature depends on the quality of image processing. When it needs to be done under a low-resolution image, the accuracy will be impaired. So how to recover the lost information from down-sampled images is important for both authentication and preprocessing. Based on Super-Resolution through Neighbor Embedding algorithm and histogram matching, we propose an improved super-resolution approach to choose more reasonable training images. First, the training image are selected by histogram matching. Second, neighbor embedding algorithm is employed to recover the high-resolution image. Experiments in several images show that our improved super-resolution approach is promising for potential applications such as low-resolution mobile phone or CCTV (Closed Circuit Television) image person authentication.

1 Introduction

The super-resolution problem arises in a number of biometric applications, for example, person authentication from a low-resolution input such an image sent by mobile phones or taken from CCTV. However, a low-resolution image loses detailed information of important features in biometric person authentication such as suspect identification. Therefore, how to recover lost information from a low-resolution image to a high-resolution one is important for building effective image based biometric applications.

Classical recovery methods include interpolation and smoothing approaches [1]. However, images may suffer from block effect and aliasing and lose details such as facial texture and edges. Better methods of super-resolution [2, 3] are developed. Recently, a novel and outstanding method with manifold learning is proposed [4]. In the paper, neighbor embedding with training images is adopted to recover the super-resolution image. One disadvantage of the approach is that the recovery of super-resolution image is easily affected by the training image which needs to be

selected within related contents manually. Meanwhile, the original paper didn't consider how to apply the approach into the preprocessing of biometric person authentication.

Considering those mentioned, we propose an improved approach where the training image is automatically selected based on histogram matching from a set of unlabeled images. Neighbor embedding is then employed. Experiments in several facial images show that the proposed approach has a potential ability to choose the reasonable training image to reconstruct super-resolution images better.

The rest of the paper is outlined as follows. In Section 2 we propose the improved super-resolution with manifold learning and histogram matching. Experimental results are reported in Section 3. In the final section we make a conclusion on this paper.

2 An Improved Super-Resolution with Manifold Learning and Histogram Matching

Our proposed approach is based on super-resolution through manifold learning. For better understanding our work, the original approach will be briefly introduced in the following subsection.

2.1 Super-Resolution Through Manifold Learning Approach

From the manifold learning point of view, data in low-dimensional subspace should have as similar neighborhood relationship as corresponding one in high-dimensional observation space [5,6]. Therefore, one patch in a low-resolution image can be represented by locally linear weighted sum of neighbor patches. And weights can be calculated based on least square criterion of locally linear embedding algorithm [7]. Similarly, the weights combining neighbor patches of some high-resolution image are adopted to reconstruct the unknown low-resolution image. It is the main idea and more details can be seen in [4].

2.2 The Proposed Approach

The disadvantage of super-resolution with manifold learning is that training images need to be manually selected with similar contents. When there are large-scale images, one automated way to select proper training image is desirable. Hence we propose the simple but powerful histogram matching approach to select the training image from a collection of images.

Histogram applies the probability of pixels to represent some statistical properties hidden in an image. The basic formulation in gray level is as follows:

$$h(r_k) = n_k \quad (1)$$

$$p(r_k) = n_k/n \quad n_k < n, r_k = 0, 1, \dots, L-1 \quad (2)$$

where r_k is the k th gray level, n_k is the number of pixels in the image having gray level r_k , and L is the number of gray level, $p(r_k)$ is an estimate of the

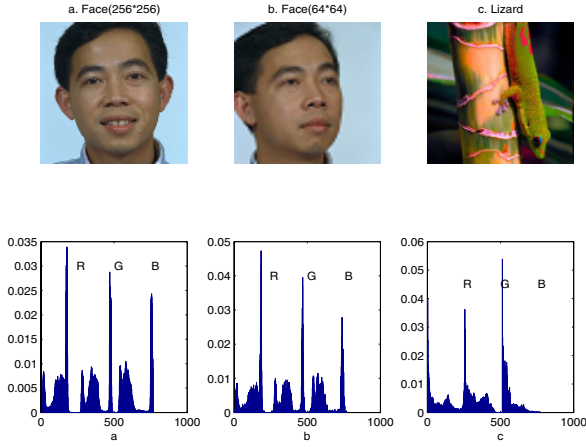


Fig. 1. Color histograms based on Y component of YIQ color space

probability of gray level r_k in an image. While conceptually simple, histogram can partially represent the contents of an image.

From a and b from Figure 1 it is obvious that there are similar normalized histograms between high-resolution and low-resolution, frontal and lateral view-point face images. Furthermore, when objects belong to different classes, for example, b and c in Figure 1, the normalized histograms will have remarkable differences. Considering the aforementioned properties, we employ histogram matching for the automated selection of relative training image from a collection of unlabeled images.

In this paper, color histograms are adopted to perform histogram matching. The color space of a image is discretized into n distinct (discretized) colors. A color histogram H is a set of vector h_1, h_2, \dots, h_{n_i} , in which each bucket h_j contains the number of pixels of color j in the image. For a given image I , the color histogram H_I is a compact summary of the image. A database of images can be queried and the most similar image to I , the image I_0 will be returned with the most similar color histogram H_{I_0} . We use the measurement of the sum of squared differences (L_2 -norm), which is formulated as follows:

$$H(I, I')_{L_2} = \|H_I - H_{I'}\|_{L_2} = \sum_{j=1}^n (H_I(j) - H_{I'}(j))^2 \tag{3}$$

Then the most similar image to image I would be the one I_0 minimizing distances among I and images from the collection set. The objective criterion is:

$$C(I) = \min_j H(I, I_j) \quad j = 1, 2, \dots, M \tag{4}$$

Where M denotes the number of training images, $C(I)$ denotes the final training image selected based on Eq. 4. We need a set of images other than only one image. The pseudo-code of the proposed approach is tabulated as in Table 1.

Table 1. The Pseudo-code of The Proposed Algorithm

Input: low resolution image X_t , training set T_r , Neighbor Numbers k , Patch Size s and Magnification Factor n .

Procedure 1: Histogram matching

1. Compute the normalized histogram H_t of low input X_t .
2. For each image Y_i from T_r , do
 - {
 - Compute the normalized histogram H_i of Y_i
 - Compute the $H(t, i)_{L_2}$ between H_t and H_i
 - }
3. Select the image Y_I which has the minimum $H(t, I)_{L_2}$ to be Y_s and blur and downsample it by $1/n$ as X_s .

Procedure 2: Super-Resolution through Neighbor Embedding

1. Cut X_t and also X_s into patches of size s by s with overlapping by one or two pixels.
2. Cut Y_s into patches of size $n \times s$ by $n \times s$ with overlapping by n or $n \times 2$ pixels accordingly
3. For each patch x_t^q from X_t , do
 - {
 - Find k nearest neighbors among all patches from X_s
 - Compute the reconstruction weights to minimize the error of reconstructing x_t^q
 - Compute the high-resolution embedding y_t^q using the reconstruction weights combining the patches in Y_s corresponding to the k nearest neighbors in X_s .
 - }
4. Enforce local compatibility and smoothness constraints between adjacent patches among all y_t^q and get Y_t .

3 Experiments

To evaluate the preprocessing performance of the proposed approach for biometric person authentication, experiments are performed on the pool of images from figure 2. There are 10 eye region images [8] and 6 images of irrelevant topics for testing. Each time we take one out of the set, downsample it to be the low resolution input and leave the rest (15 images) to make up the training set.

The low input size is 70×20 pixels and our goal is to compute the $4X$ magnification. We set the parameters as the paper [4] does, using 5 nearest neighbors, patches size of 3×3 , overlapping 2 pixels, according to its satisfactory performance. With histogram matching we can compute a series of ranking subplots of choices of training images. The ranking k can be apprehended that without those images having ranking higher than k in the training set, the one with ranking k is the choice as training image. Examples of reconstruction of high resolution images ranked by histogram matching are shown in Figure 3 and Figure 4.

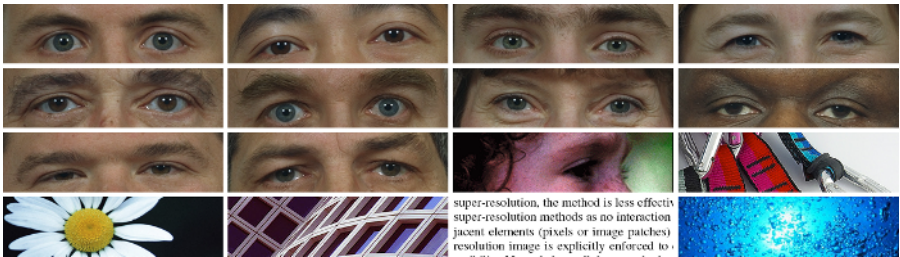


Fig. 2. Training images pool, from left to right, top to bottom: labeled No.1 to No. 16



Fig. 3. Results of (YIQ) histogram matching and neighbor embedding. Rankings of the results descending from left to right, top to bottom. The corresponding numbers in training set are: 5, 10, 2, 4, 3, 9, 6, 7, 16, 11, 8, 14, 12, 13 and 15.



Fig. 4. a: High resolution target (Label No.1 in our pool). b: Low resolution input. c: Parts of results, Training images used: Left: No.5, Middle: No.16, Right: No. 15.

We can see that histogram matching chooses topic related training images prior to those irrelevant images (No.12 to No.16 in training set). Furthermore, it is easy to find that mosaic effect is reinforced as ranking increases in Figure 4.

To quantitative analyze the performance of the reconstruction of super-resolution image, RMS (Root Mean Square) errors are introduced which have the formulation of

$$RMSe = \left(\sum_{i=1}^n \frac{(\hat{y}_i - y_i)^2}{n} \right)^{\frac{1}{2}} \tag{5}$$

Where \hat{y}_i stands for the values of pixel in the ideal target Y and y_i stands for the values of corresponding pixels in output Y_t . And n stands for the number of total pixels in Y .

According to the ranking of histogram matching, average ranked RMS errors and standard deviations can be computed from 15 images of which each is not the low-resolution test image. The ranked RMS errors are illustrated as in Figure 5.

Although the method with histogram matching may not always choose the optimal training image left in the training set, it chooses image good enough and only increases the RMS error by a trivial little comparing to the optimal one.

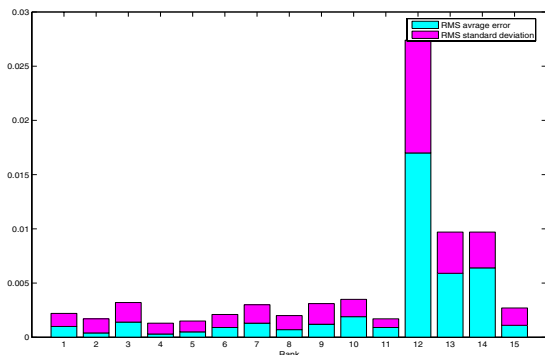


Fig. 5. Average RMS errors and standard deviations of normalized histogram matching-based ranking with 16 test images



Fig. 6. a: High-resolution target; b: Low-resolution input(Enlarged); c: Histogram matching result (No.4 chosen, $RMSe=0.0535$); d: Optimal choice result (No. 10 chosen, $RMSe=0.0519$)

The method of histogram matching is efficient enough to automatically choose training image instead of choosing manually.

At last we show an example of a whole face using YIQ histogram matching and compare the performance with optimal RMS choice. The results are illustrated in c, d in Figure 6. Our training set is identical to our experiment pool while the size of each image is a little smaller to save running time. Notice that our choice of training image based on histogram matching are second best in all RMS errors, i.e., it is the optimal choice if no.10 does not exist in training set.

It is worth noting that the result is not as good in overall details as before. One reason is that we just use parts of eye region, which are not very so similar with the whole face. During our further research of using geometric division to separate facial features and choose training images for each feature under the same principle, in other regions good recovery is obtained as that of eye regions.

4 Conclusion

In this paper, to carry out super-resolution of face images, we improve the novel method of Super-Resolution through Neighbor Embedding. We indicate the problem of the choice of the training image affecting the quality of results. In-

stead of selecting the training image manually, we propose the automatic method of histogram matching to choose the proper image from the training set and obtain fairly good results. And it is effective and costless to carry out and as a result it explores the capacity of the training set with limited images. The proposed approach has potential application in the preprocessing of biometric person authentication.

Several problems deserve to make further research. First, the performance of neighbor embedding can be further improved and specified for biometric authentication. Second, histogram matching only provides a principle and coarse approximation to the selection of training image, more elaborate methods are under our research. Finally, the practical combination of our proposed super-resolution approach and biometric person authentication systems is desirable.

Acknowledgement

The authors are very grateful to PhD Hong Chang and Professor Dit-Yan Yeung for generous providing source code and invaluable comments. And Portions of the research in this paper use the Gray Level and Color database of the FERET program.

References

1. R. C. Gonzalez and R. E. Woods, *Digital Image Processing* (Second Edition), Prentice Hall, 2002.
2. William T. Freeman, Thouis R. Jones, and Egon C. Pasztor, "Example-Based Super-Resolution," in *Proceedings of Computer Graphics and Applications, IEEE*, March/April 2002, pp. 56–65.
3. Simon Baker, Takeo Kanade, "Limits on Super-Resolution and How to Break Them," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.24, NO.9, September 2002.
4. H. Chang, D. Y. Yeung, Y. Xiong, "Super-resolution through neighbor embedding," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, vol.1, pp.275-282, Washington, DC, USA, 27 June - 2 July 2004.
5. J. Zhang, S. Z. Li, and Jue Wang, "Manifold Learning and Applications in Recognition," in *Intelligent Multimedia Processing with Soft Computing*. Yap Peng Tan, Kim Hui Yap, Lipo Wang (Ed.), Springer-Verlag, Heidelberg, 2004.
6. J. Zhang, "Several Problems in Manifold Learning," *Machine Learning and Applications*, Zhi-Hua Zhou et. al. (Eds.), Tsinghua University Press, 2005.
7. S. T. Roweis and K. S. Lawrence, "Nonlinear Dimensionality Reduction by Locally Linear Embedding," *Science*, 290, pp. 2323-2326, 2000.
8. P. J. Phillips and H. Moon and S. A. Rizvi and P. J. Rauss, "The FERET Evaluation Methodology for Face Recognition Algorithms," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Volume 22, October 2000, pp. 1090-1104.