Novel Human Face Hallucination for Real World Setting

HAJU MUHAMED MUHAMED NALEER*, YAO LU**, AND FEROZ SHAH***

RECEIVED ON 21.06.2010 ACCEPTED ON 21.06.2012

ABSTRACT

The face hallucination research, in past ten years, has proposed a number of human face hallucination techniques. The face hallucination method does not follow the automatic face alignment in their experimental stages. In other words, these methods are not adequate for real world setting. Mainly some factors affect the final results of such methods, such as face pose, alignment, and lighting correction. Furthermore, the experimental images are aligned manually, and it may neither be accurate, nor added to the frame work. The main consideration in this paper is, how to produce the hallucinated human face image from single input low resolution image, using two step example base learning method with consideration of face pose and alignment in automatic frame work. We propose that the novel human face hallucination automatic frame work is excellent agreement with real world input low resolution images.

Key Words: Face Hallucination, Alignment, Pose, Example Base, Learning Methods.

1. INTRODUCTION

In modern years, the ordinary learning-based face super-resolution algorithms typically engage two steps. Domain knowledge about facial images was used to generate high resolution facial images for face hallucination. Global face image is generated in first step keeping the main characteristics of the real image by learning method like, LLE (Locally Linear Embedding) and probabilistic method in a MAP (Maximum a Posterior) frame or manifold. The next step produces residual image to compensate the results of the first step.

In addition, Baker, et. al. [1] were the earliest to build up a hallucination method underneath a Bayesian formulation and anticipated the term "Face Hallucination". Liu, et. al. [2] recommended a two-step move to integrate a global parametric model with Gaussian assumption and a local nonparametric model based on MRF (Markov Random Field). Both of these methods compose the degradation function into the formulation in order to solve the concluding hallucinated result. Also, probabilistic model based face hallucination algorithms [2-4] could do with large number of training images. Su, et. al. [5] extract multi-orientation and multi-scale information of low-level facial characteristics from the input of low-resolution and high-resolution faces using a steerable pyramid.

They adopted the pyramid parent configuration and local most excellent counterpart to optimize the prior information in order to solve a Bayesian MAP problem. Enlightened by Liu's work [2], most of the existing research considered face hallucination as a two-step problem [6-8]. Firstly, a global face image was generated which held the main

* Department of Mathematical Science, Faculty of Applied Sciences, South Eastern University, Sri Lanka.
** Beijing Laboratory of Intelligent Information Technology, School of Computer Science, Beijing Institute of Technology, Beijing, China.
*** Assistant Professor, Department of Electrical Engineering, Mehran University of Engineering & Technology, Jamshoro.

characteristics of the original high-resolution face but lacked some comprehensive features. Secondly, a residue face image having the high-frequency image information has been synthesized and added to the global face image to achieve the decisive results.

For instance, in a two-phase face hallucination model residential by Zhuang, et. al. [6], the locality preserving hallucination algorithm united LPP and RBF (Radial Basis Function) regression collectively to hallucinate a global high-resolution face. Facts of the synthesized high-resolution face were highly enhanced by residue compensation based on Neighbor Embedding [9].

Wang, et. al. [10] recommended the exploit of PCA to correspond to the structural likeness of face images; On the other hand this can hardly keep up the global efficiency and visual consistency, more than ever at locations around the face contour and margin of the mouth. Inspired by LLE, a famous manifold learning method, Chang, et. al. [9] used the Neighbour Embedding algorithm based on the postulation that the training low- and high-resolution images form manifolds with comparable local geometry in two diverse feature spaces.

Motivated by Chang's work [9], a number of face hallucinations methods [6-12,14] were urbanized, based on Neighbor Embedding or by means of neighbor patch. Conversely, Su's [15] experiments showed that neighbourhood preservation for low-resolution and high-resolution patches rarely holds. Jia, et. al. [16] applied hierarchical tensor (multilinear) algebra to face hallucination to work out dissimilar views of faces. Multilinear analysis [17-18] can be engaged to model the association between two spaces with different views of the faces. Multilinear investigation is a wide-ranging expansion of habitual linear methods and two-dimensional tensor is analogous to PCA.

Even though Jia, et. al. [16] provided a novel way to represent dissimilar views of faces, his super-resolution advance still follows a two-step framework, integrating parametric global models with Gaussian assumption. Yang, et. al. [19] applied the perspective of Compressed Sensing [20] to super-resolution. The high-resolution image created by the sparse illustration come close to might not satisfy the acquisition process assumed. The ending result was generated by extra residue compensation footstep by means of back-projection method.

Park, et. al. [21] anticipated a novel example-based face hallucination method with a comprehensive morphable face model. Since the global way in the first step of Park's method [21] frequently grades in low re-enactment accuracy, the algorithm anticipated supplementary methods to compensate for residual errors and protect characteristics. This method has the privilege that it enhances the resolution of face image and getting better the performance of face recognition. Instead of using a probabilistic model or manifold learning model, Xiang, et. al. [22] with the help of coefficient learning anticipated an example-based two-step face image super-resolution method.

In this experimental, Xiang, et. al. [22] performed face hallucination method on CAS-PEAL Face Database. These face images and input image were aligned manually and under same light condition. In residual compensation stage, each low-resolution position patch covers 3x3 pixels by Xiang, et. al. [22] experimental result is an excellent agreement with original images. However, mainly some factors are affected in the final results for hallucination methods, such as input low-resolution face pose and alignment. Even though, this research work including Xiang, et. al. [22] does not work for real world setting. The methods discussed above do not pertain to automatic frame work for real world low-resolution input image to get the hallucinated face image as a result. This is the main argument in this paper. We propose the hallucinated human face image from single input low resolution based on learning method with consideration of face pose and alignment in automatic frame work.

We have developed a novel automatic frame work for face hallucination. Next section will describe in detail the proposed system frame work. In Section 2 the experimental setup is explained.

2. PROPOSED SYSTEM FRAME WORK

To overcome the face hallucination for multiview low-resolution input face image, the proposed eight-orientation-difference method and SVM (Support Vector Machin) classifier method can be considered respectively. In this study, the automatic face alignment and patch method are having the major role of face example base hallucination. Therefore, initially the proposed eight-orientation-difference method will be considered. Later, automated face cropping and alignment will be considered.

2.1 Two Step Face Hallucination Method

Our proposed two step hallucination method is mostly related to the Xiang, et. al. [22]method. But we made the major modification based on the local residual compensation. It will affect the final result of this method.

In the stage of residual compensation, the image patches play major role in this system. Here, we explore facial prior information embedded among un-overlapped regions of patches to adequate for the face hallucination learning method. The differences between un-overlapped regions of eight-connected neighbor patches maintain recognizable facial structure information.

In this method, the input low-resolution face image is broken into overlapped patches, and a set of appropriate high-resolution candidates are collected from training database for each patche. In order to handle various facial patterns aesthetically while maintaining the smoothness of high-resolution image, the patch size is generally very small. The neighbor patches' affiliation via overlapped regions is barely used to assure smoothness of hallucinate image, so the prior is not strong enough to regularize super-resolution when experiential low-resolution image loose facial structure information.

We consider that a high-resolution image is constructed by lots of overlapped small patches using correspondence low-resolution patch, and the un-overlapped regions in these patches are shown in Fig. 1. If we calculate the difference between un-overlapped regions of each patch with its eight-connected neighbors, and construct these differences into image, as shown Fig. 2, and we can use that eight-orientation-difference images to face hallucination method.

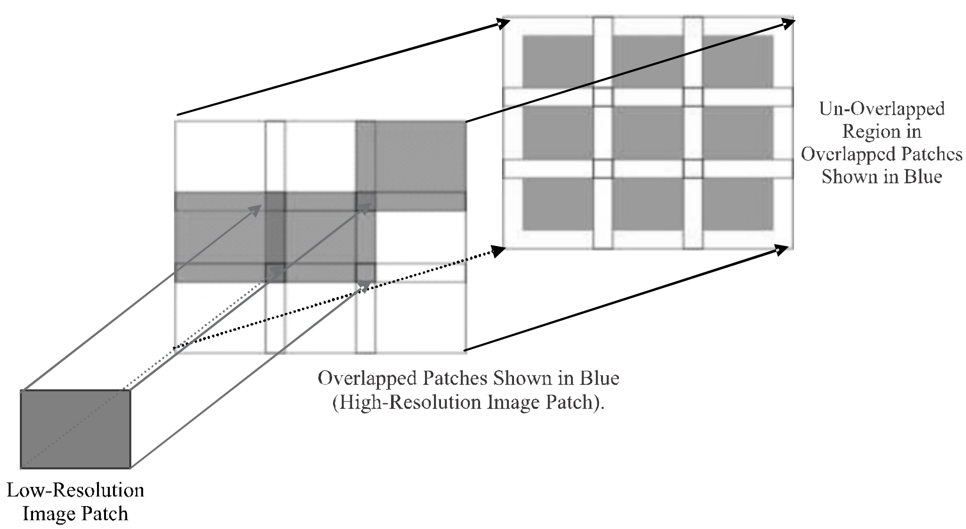

FIG. 1. OUR PATCHES TECHNIQUE INCORPORATES TO TWO STEP METHOD

The proposed learning method for face hallucination is used in error compensation of the residual, every low-resolution training image could be obtained as shown in Fig. 3.

2.2 Automatic Face Alignment

This technique is based on analyzing the high dimensional feature space for categorization. The SVR classifiers are also called SVM. SVM based system automatically aligns faces in differing poses. To automate the process, We first use the Viola-Jones face detector [23] to automatically crop faces from images. We now present work on SVM based alignment procedure that aligns given images reasonably well for a large number of poses. The procedure learns canonical alignments for a set of discretely separated poses based on provided manual alignments as training data. The alignment algorithm uses a set of "alignment SVM's for each pose.The algorithm works as follows:

- Assemble a training set of manually aligned faces in pose n.
- For each image, divide the entire image into K consistently sized regions over the whole image.
- For each region r^k_n, train an SVM A^k_n (using a radial basis function kernel) to organize all examples of region r^k_n as positive and any other region r^l_n where $1 \neq k$ as negative.
- Using the alignment scoring procedure, a search for a good alignment can be done simply by a brute force search over the alignment parameters of translation, scale, and in-plane rotation.

2.3 Proposed Automatic Face Hallucination Frame Work

Our proposed automatic face hallucination method for different pose low resolution input image frame work as

(a) (b) (c) (d) (e) (f) (g) (h) (i)

FIG. 2(a). ORIGINAL FACE IMAGE (b-i) EIGHT-ORIENTATION-DIFFERENCE IMAGES, CONSTRUCTED BY DIFFERENCE BETWEEN NEIGHBOR PATCHES' UN-OVERLAPPED REGIONS

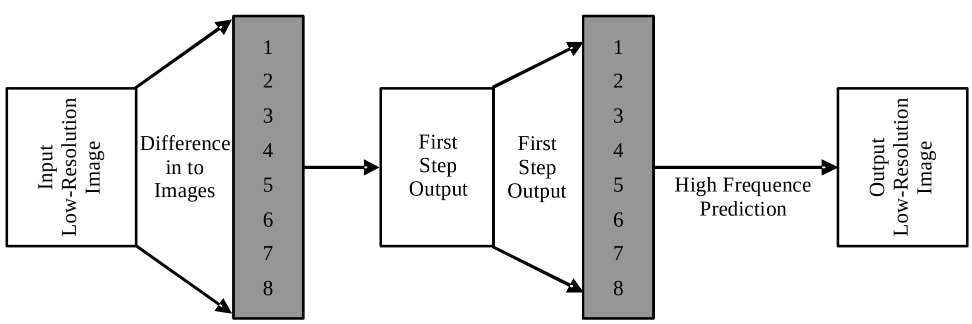

FIG. 3

shown in Fig. 4. The following steps are followed in this system:

- *Step-1:* Create the training data set for automatic face alignment manually as training data set for alignment process.
- *Step-2:* Produce the input low- resolution image using SVM with Step-1.
- *Step-3:* Create the aligned low-resolution data set for two step face hallucination method using Step-2.
- *Step-4:* Create the alignment high-resolution image data set using SVM alignment method for each image.
- *Step-5:* For residual compensation for two step method, create the patches as described in Section 2.1 for each image in data sets (low and high -resolution).

3. EXPERIMENTAL RESULT

We conducted our proposed frame work on the FERET and CAS-PEAL face databases. About 250 images of the people in the same light conditions were selected. The low-resolution images (36x24) were obtained from same high-resolution images by down sampling. In this experiment, all high-resolution images were (128x96). Further, 100 images were selected and manually aligned for automatic face alignment method in different poses. As many as 30 images have been selected for the purpose of testing our proposed frame work.

Experimental data sets were aligned by automatic face alignment method as descried above. In the residue compensation stage the low-resolution images were treated with the difference between un-overlapped regions of each patch with its eight-connected neighbours. We considered 3x3 image patches on low-resolution images and 12x12 image patches on high-resolution images. We had used different pose input low-resolution images in to Xiang, et. al. [22] system and our proposed automatic two-step face hallucination frame work system. The results obtained on both systems are shown in Figs.5-6 for the comparison purpose.

We can predict that our proposed frame work is in excellent agreement with original face image from results given in Fig. 5. Even though Xiang, et. al. [22] method is used for two-step hallucinating method, But, it is not well adequate for different pose of low-resolution input images. Whereas; our method works for any pose low-resolution input images. Fig. 6 shows that, our proposed frame work is working well in automatic two-step face hallucination for any real world images in different pose low-resolution input images.

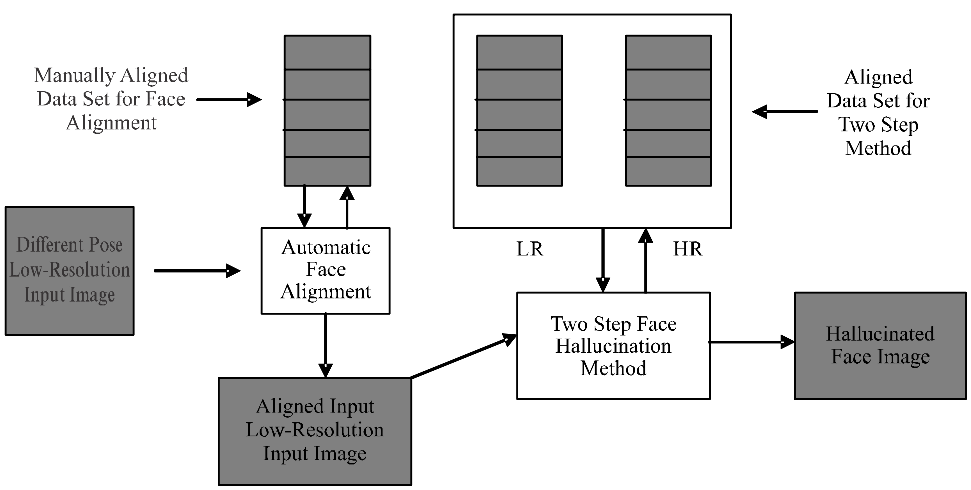

FIG. 4. THE PROPOSED AUTOMATIC FRAME WORK FOR FACE ALIGNMENT AND HALLUCINATING SYSTEM.

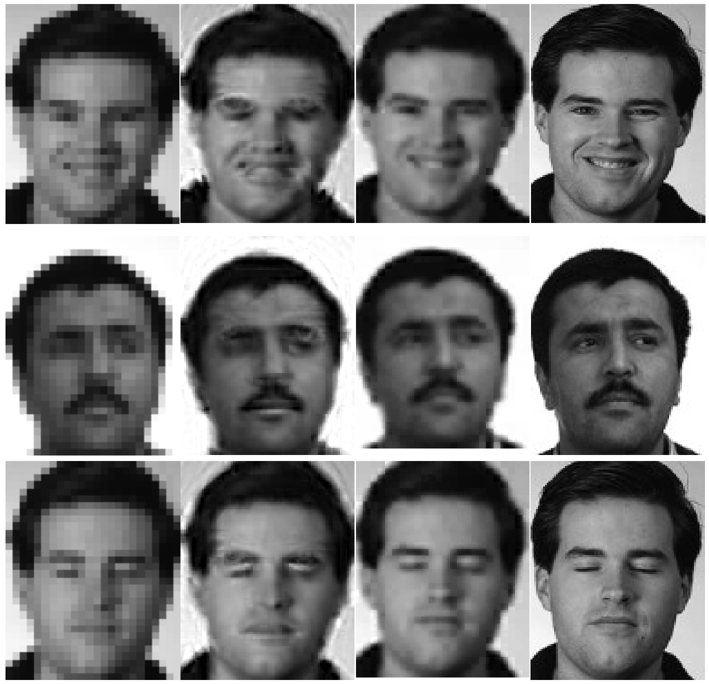

FIG. 5(a) THE LOW-RESOLUTION INPUT (b) XIANG MA'S RESULTS (c) OUR PROPOSE METHOD (d) ORIGINAL IMAGE

(a) LOW-RESOLUTION INPUT IMAGE (b) AUTOMATIC OUTPUT FROM OUR PROPOSED FRAME WORK (c) ORIGINAL IMAGE

FIG. 6. IN DIFFERENT POSE

4. CONCLUSION

We presented a method for using region-based pose relation support vector machines to learn aligning and the differences between un-overlapped regions of eight-connected neighbor patches. It also maintains recognizable facial structure information. Then we incorporated these two techniques to automatic two-step face hallucination frame work. Our proposed automatic two-step face hallucination is not only applicable for frontal face images but also it can be used to different pose images. The proposed method is slightly not good towards the computational cost, because of the patches making the sparse matrix during the experimental work for compensation stage. We can say with confidence that one direction of future research might focus on computational cost. That is, the sparse residual compensation can be adapted to two-step face hallucination. In the first step, a global face image is constructed by optimal coefficients of the interpolated training images. Then, it can be used over complete patches dictionary and adopt the sparse representation to find the high-resolution residual image (local face image) in the second step.

ACKNOWLEDGEMENT

The authors would like to thank the Chinese Scholarship Council (CSC-2009) for financial support to do this research. The authors are also thankful to the reviewers for their valuables comments.

REFERENCES

[1] Baker, S., and Kanade, T., "Hallucinating Faces", International Journal of Computer Vision, Volume 20, No. 1, pp. 25-47, 2000.

[2] Liu, C., Shum, H., and Zhang, C., "A Two-Step Approach to Hallucinating Faces: Global Parametric Model and Local Nonparametric Model", Proceedings of CVPR, Volume 20, No. 3, pp. 192-198, Kauai Marriott, Hawaii, 2001.

[3] Li, Y., and Lin, X.Y., "An Improved Two-Step Approach to Hallucinating Faces", Proceedings of CVPR Conference, Volume 1, No. 7, pp. 275-252, Washingtan, 2004.

[4] Liu, C., Shum, H.Y., and Freeman, W.T., "Face Hallucination: Theory and Practice", International Journal of Computer Vision, Volume 52, No. 4, pp. 1289-1306, 2006.

[5] Su, C.Y., Zhuang, Y.T., Huang, L., and Wu, F., "Steerable Pyramid-Based Face Hallucination", Pattern Recognition, Volume 38, No. 6, pp. 813-824, 2005.

[6] Zhuang, Y.T., Zhang, J., and Wu, F., "Hallucinating Faces: LPH Super-Resolution and Neighbor Reconstruction for Residue Compensation", Pattern Recognition, Volume 40, No. 6, pp. 3178-3194, 2007.

[7] Liu, W., Lin, D., and Tang, X.O., "Neighbor Combination and Transformation for Hallucinating Faces", Proceedings of IEEE International Conference on Multimedia and Expo, Volume 3766, No. 5, pp. 221-230, Netherlands, 2005.

[8] Park, S.W., and Savvides, M., "Robust Super-Resolution of Face Images by Iterative Compensating Neighborhood Relationships", Proceedings of Biometrics Symposium, Volume 10, No. 4, pp. 321-333, USA, 2008.

[9] Chang, H., Yeung, D.Y., and Xiong, Y.M., "Super-Resolution through Neighbor Embedding", Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Volume 1, No. 3, pp. 275-282, Washington, 2004.

[10] Wang, X.G., and Tang, X.O., "Hallucinating Face by Eigentransformation", IEEE Transactions on Systems Man and Cybernetics, Volume 35, No. 3, pp. 425-434, 2005.

[11] Fan, W., and Yeung, D.Y., "Image Hallucination Using Neighbor Embedding Over Visual Primitive Manifolds", Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Volume 1, No. 9, pp. 275-282, USA, 2007.

[12] Liu, W., Lin, D.H., and Tang, X.O., "Hallucinating Faces: Tensor Patch Super-Resolution and Coupled Residue Compensation", Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Volume 2, No. 7, pp. 478-484, USA, 2005.

[13] Park, S.W., and Savvides, M., "Breaking the Limitation of Manifold Analysis for Super-Resolution of Facial Images", Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, Volume 1, No. 3, pp. 573-576, USA, 2007.

[14] Fan, W., and Yeung, D.Y., "Image Hallucination Using Neighbor Embedding Over Visual Primitive Manifolds", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Volume 19, No. 8, pp. 1-7, Maxico, 2007.

[15] Su, K., Tian, Q., Xue, Q., Sebe, N., and Ma, J.S., "Neighborhood Issue in Single-Frame Image Super-Resolution", Proceedings of IEEE International Conference on Multimedia and Expo, Volume 3, No. 11, pp. 1122-1125, Netherlands, USA, 2005.

[16] Jia, K., and Gong, S.G., "Generalized Face Super-Resolution", IEEE Transactions on Image Processing, Volume 17, No. 4, pp. 873-886, 2008.

[17] Vasilesescu, M.A.O., and Terzopoulos, D., "Multilinear Image Analysis for Facial Recognition", ICPR, Volume 2, No. 5, pp. 511-514, 2002.

[18] Vasilescu, M.A.O., and Terzopoulos, D., "Multilinear Analysis of Image Ensembles: Tensor Faces", Pattern Recognition, Volume 9, No. 8, pp. 511-514, 2002.

[19] Yang, J.C., Wright, J., Huang, T., and Ma, Y., "Image Super-Resolution as Sparse Representation of Raw Image Patches", Proceedings of CVPR, Volume 23, No. 17, pp. 1806-1816, 2008.

[20] Donoho, L.D., "Compressed Sensing", IEEE Transactions on Information Theory, Volume 52, No. 4, pp. 1289-1306, 2006.

[21] Park, J.S., and Lee, S.W., "An Example-Based Face Hallucination Method for Single-Frame, Low-Resolution Facial Images", IEEE Transactions Image Processing. Volume 17, No. 21, pp. 1806-1816, 2008.

[22] Ma, X., Zang, J.P., and Qi, C., "Hallucinating Faces: Global Linear Modal Based Super-Resolution and Position Based Residue Compensation", ICAP, Volume 11, No. 4, pp. 835-843, Spain, 2009.

[23] Viola, P., and Jones, M.J., "Robust Real-Time Face Detection", International Journal of Computer Vision, Volume 57, No. 3, pp. 137-154, 2004.