Research Journal of Applied Sciences 9 (7): 443-451, 2014
ISSN: 1815-932X

Adaptive Locally Linear Embedding for Multiview Face Hallucination

[1]Metee Thongdee, [1]Siriporn Supratid and [2]Parinya Sanguansat
[1]Faculty of Information Technology, Rangsit University, 12000 Pathum-Thani, Thailand
[2]Faculty of Engineering and Technology, Panyapiwat Institute of Management, Nonthaburi, Thailand

Abstract: The traditional Locally Linear Embedding (LLE) technique was applied for face hallucination. This technique determines the optimal weights by the fixed number of neighbors for every point. The previous research, named an Adaptive Locally Linear Embedding (ALLE), referred to a modified version of LLE was proposed to apply with frontal view face hallucination; it uses a threshold of similarity for selecting the neighbors of each point. However, frontal face is barely captured in the real world. Therefore, this study proposes a novel ALLE for multiview face hallucination. The main objective is to generate high quality of frontal and non-frontal face images. The processing steps, according to the proposed method are operated as follows; first, a Low Resolution (LR) face in one of front, up, down, left or right views is fed as an input then the other views of such an LR image are generated by ALLE which applies a threshold of similarity for selecting the neighbors of each point and High Resolution (HR) face images in all views of the same input object are achieved afterwards. The experimental results show that the proposed method yields the better image quality of the reconstructed frontal and non-frontal face images over the baseline methods.

Key words: Face hallucination, super-resolution, multiview, locally linear embedding, adaptive locally linear embedding

INTRODUCTION

In these days, video surveillance cameras have been widely used in many places such as banks, stores and parking lots where intensive security is critical. Details of facial features obtained from the surveillance video are important for identifying personal identity. However, in many cases, the images obtained from the surveillance cameras cannot be well identified due to low resolution of facial images that cause some losses of the facial features. Thus, in order to obtain detailed facial features for a purpose of personal recognition, it is necessary to infer a Low-Resolution (LR) facial image to a High-Resolution (HR) one by a technique called face hallucination or face super-resolution (Wang and Tang, 2005). Such techniques are applied in a variety of important sectors, e.g., medical imaging, satellite imaging, surveillance system, image enlarging in web pages and restoration of old historic photographs. A super resolution is a technique to produce a single or multiple High-Resolution (HR) images from the Low-Resolution (LR) image sequences or a single LR image. Under some circumstances, it is impossible to obtain image sequences. Due to limited information of image identification, reconstruction and expression analysis is a challenge to both human and computer. Several Super-Resolution Reconstruction (SRR) researches have been proposed, relying on two approaches: reconstruction-based and learning-based approach. Reconstruction-based approach employs multiple LR images of the same object as an input for reconstructing an HR image whereas learning-based approach uses numbers of training samples from same domain with different objects to reconstruct the HR image. An advantage of the learning-based approach is its ability to reconstruct the HR image from the single LR image. In this research, learning-based super-resolution which is also known as "face hallucination" (Baker and Kanade, 2000, 2002; Elad and Feuer, 1999; Freeman and Pasztor, 1999; Hardie *et al*., 1997; Lin and Shum, 2004; Park *et al*., 2003; Hu *et al*., 2011a, b; Gao and Yang, 2013) is focused for applying to human facial images. A number of related facial hallucination methods have been proposed in recent years. Indeed, learning based methods have acquired greater attention as they can achieve high magnification factor and produce positive super-resolved results compared to other methods.

Ma *et al*. (2009, 2010a, b) proposed a Facial Hallucination Method based on position-patch to improve the image quality. Such Ma's Method performs one-step facial hallucination based on position-patch instead of neighbor-patch. A patch position is one of the learning-based approaches that utilizes facial image as well as image features to synthesize high-resolution facial images from the low-resolution images. In comparison,

Corresponding Author: Metee Thongdee, Faculty of Information Technology, Rangsit University, 12000 Pathum-Thani, Thailand

neighbor patches is widely used in face hallucination (Jia and Gong, 2008; Zhuang *et al*., 2007; Liu *et al*., 2005; Park and Savvides, 2008; Fan and Yeung, 2007; Zhou *et al*., 2012, 2014; Jung *et al*., 2011). In their research, the reconstruction of a high-resolution facial image could be created based on a set of high and low resolution training image pairs. The high resolution image obtained from their proposed method was generated based on the same position image patches of each training images. The result of this research shows better image quality than many other methods including Cubic B-spline, Wangs Eigen-Transformation Method (Wang and Tang, 2005), ChangsNeighbor Embedding (Chang *et al*., 2004) and Zhuang *et al*. (2007)'s Locality Preserving Method. By Zhou *et al*. (2012) have extended the research by Ma *et al*. (2010a, b) by using bilateral patches. By Hu *et al*. (2011b), local pixel structure is learnt from nearest neighbors (KNN) faces. However, there are some uncontrollable problems regarding this method, i.e., the facial images captured by the camera are LR and non-frontal facial images where such a Facial Hallucination Method is limited to the frontal faces. Therefore, it is practically significant to study how to create the HR multi-viewed faces from the LR non-frontal images.

As a sequence, Ma *et al*. (2010a, b) proposed a Multi-Viewed Facial Hallucination Method based on the position-patch (Ma *et al*., 2009, 2010a, b). It is a simple face transformation method that converts an LR face image to a globally, predicting LR multiple views of that given LR one. Based on the synthesized LR faces, facial details are incorporated by using the local position-patch (Ma *et al*., 2009). Nevertheless, the traditional Locally Linear Embedding (LLE) technique, applied in such a Hallucination Method still confronts a problem, relevant to the determination of the optimal weights. Such weights are defined using fixed number of neighbors for every points. This is not practical for real world data because the number of neighbors in each point is not equal to others.

In this study, a Multiview Face Hallucination (MFH) using Adaptive Locally Linear Embedding (ALLE) technique is proposed for efficiently reconstructing high-resolution face images from a single low-resolution one. The optimal weights determination problem, addressed in LLE is resolved by using the previously proposed technique, ALLE (Thongdee *et al*., 2012). The ALLE is applied to manipulate the non-frontal facial details. By feeding a single LR face in one of up, down, left or right views to the proposed method, the HR images will then be generated in all views.

LOCALLY LINEAR EMBEDDING FOR FACE HALLUCINATION

The LLE (Roweis and Saul, 2000) algorithm is employed to discover non-linear structure in the data. The algorithm is developed under the assumption that any object is "nearly" flat on small scales. The basic of LLE is described by Sun *et al*. (2010). The original LLE try to map a data from one space globally to another space by embedding the locations of the neighbors for each point. For face hallucination, it is about LR and HR face image spaces. The main idea of LLE is to minimize the reconstruction error of the set of all local neighborhoods in the data set. The LLE process can be categorized into two steps as follows:

Step 1: Locally fitting hyper-planes around each object x based on its K-nearest neighbors.

Step 2: Calculating the reconstruction weights by minimizing reconstruction error by the cost function:

$$\varepsilon_w = \sum_{i=1}^{n} \left| \mathbf{x}_i - \sum_{j=1}^{K} w_{i,j} \mathbf{x}_{N(j)} \right| \tag{1}$$

In this step, x_i is linearly reconstructed in terms of its neighbors $x_{N(1)}, x_{N(2)}, ..., x_{N(K)}$. If x_i and x_j are not in the same neighborhood then $w_{i,j} = 0$. Weights $w_{i,j}$ are computed according to the least square principle and they sum up to 1. Weights $w_{i,j}$ are stored in an n×n sparse matrix W.

For face hallucination such reconstruction weights are computed in LR face image space and then applied to HR space to reconstruct the HR face image that corresponding to the LR input.

ADAPTIVE LOCALLY LINEAR EMBEDDING

In this study, the Adaptive Locally Linear Embedding (ALLE) is proposed. Normally, the weight matrix W is computed from the neighborhood in the same space of the training samples. Similar to LLE, ALLE computes the weights from the space of training samples which is adaptively built from only the neighborhood of each input, not all training samples. Since, some information that contains in all training samples can make the optimizer misleading to another optimal values. The threshold of similarity θ is defined for building each subspace by:

$$\frac{\sum_{i=1}^{K} d_i}{\sum_{i=1}^{n} d_i} \le \theta \tag{2}$$

where, d_1 is the distance between input sample and the 1st training sample in which $d_1 \leq d_2 \leq d_3 \leq \ldots \leq d_n$. That means the number of the nearest neighbors (K) is not the same value for all input samples.

THE PROPOSED NOVEL MULTIVIEW FACE HALLUCINATION

The position-patch based multiview face hallucination has been proposed in MFH (Ma *et al*., 2010a, b). The framework is summarized in Fig. 1 and the detail of processing steps is described as follows:

Step 1: Input LR face image at view p, LR training set and HR training set in all views.

Step 2: Estimate the linear combination to obtain the optimal weights W_p of LR input from LR training set at view p by LLE.

Step 3:

- For each view: synthesize other views of LR face images by the LR training set with the optimal weights W_p by LLE
- Divide multiview LR face images, multiview LR training face images and multiview HR training face images, respectively into position patches
- Concatenate and integrate the hallucinated high resolution patches to form a facial image which is the target high resolution face image

Face image is represented as a column vector of all pixel values. Based on structural similarity, face image can be synthesized using the linear combination of training objects (Wang and Tang, 2005). In other words, a face image at unfixed view can also be reconstructed using linear combination of other objects in the same view aligned as shown in Fig. 2:

$$I_p = W_p L_p \tag{3}$$

Where:

I_p = The given LR face at view p

W_p = The construction coefficients at view p that can be determined by Eq. 1

L_p = The training faces at view p

Since, the LR face images for other views L_o are required. From the assumption by Ma *et al*. (2010a, b), the LR face images for other views I_o can be approximated by:

$$I_o \approx W_p L_o \tag{4}$$

Since, there is no information for determining the construction coefficients at other views. After obtaining LR in all views, these LR images are hallucinated to HR face images of all views by position-patch approach.

The proposed framework for multiview face hallucination is to replace the linear combinations in all processes by using ALLE instead of LLE in order to improve the performance of hallucinating face image in multiple views. That means the number of the neighbors can be adapted for each input patches. This would relieve the deviation from optimal values.

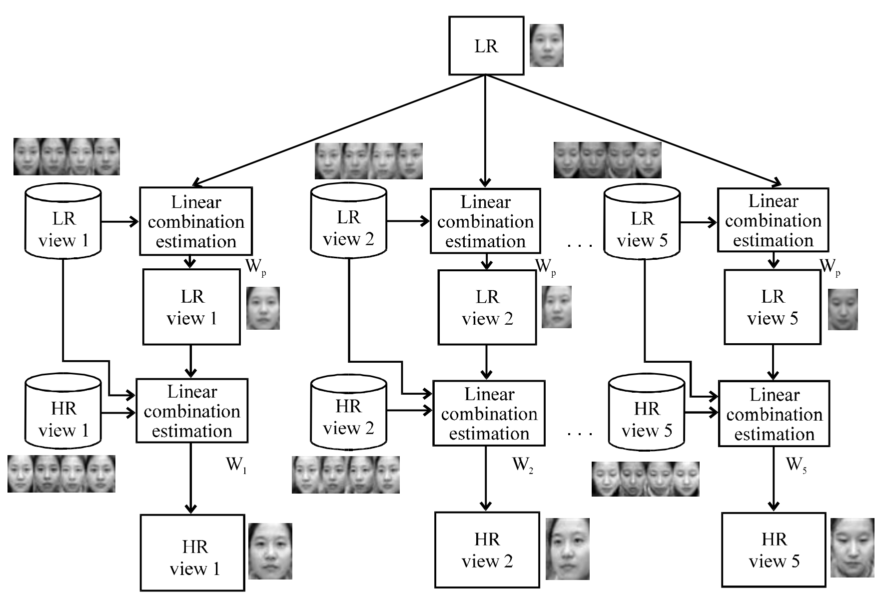

Fig. 1: Multiview face hallucination framework

EXPERIMENTAL RESULTS

The Face Hallucination algorithm was performed on the CAS-PEAL-R1 face database (Gao *et al*., 2008). The 5,055 face images of 1011 different individuals are randomly selected. Each individual has five different views (left, right, up, down and frontal views) under the same light condition.

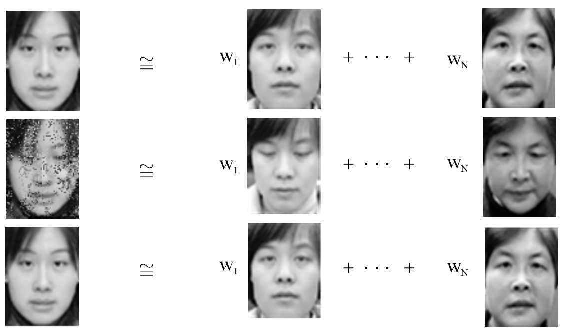

Fig. 2: Linear combination for multiview face

These face images were aligned manually using the locations of three points: centers of left and right eyeballs and center of the mouth (Ma *et al*., 2009). Figure 3 shows some aligned face images which are cropped to 32×24 pixel for low resolution face images and to 128×96 pixel for high resolution face images. Some cropped face images are shown in Fig. 4 and 5. Based on the same training sets, the Proposed Method is compared to Baseline Method, Bicubic and Ma's *et al*. (2010a) Method. Input face images are shown in Fig. 6 including frontal view, up view, down view, left view and right view. Each individual input face image was generated into five different outputs of LR, synthesized LR face images and HR face images according to the framework. Face image outputs, generated by the Proposed Method were compared with Ma *et al*. (2010a, b)'s Method as shown in Fig. 7-11. The experimental results show the difference between the results of high-resolution facial image by varying the number of neighbors, K as 100, 200 and 300 regarding LLE along with 0.1, 0.2 and 0.3 similarity threshold (θ) according to ALLE, respectively. With regard to ALLE,

Fig. 3: Aligned face images

Fig. 4: Low resolution cropped face images

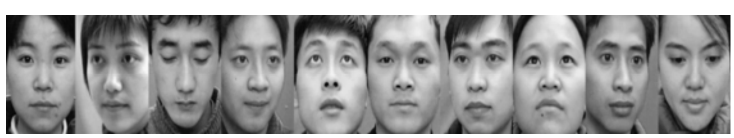

Fig. 5: High resolution cropped face images

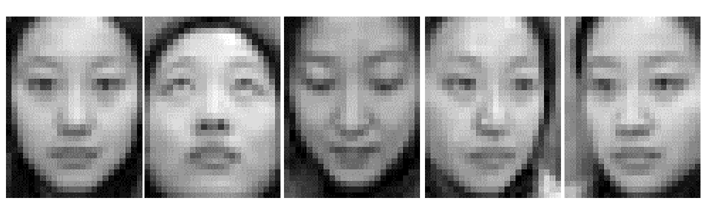

Fig. 6: Example input a face image on five different single views

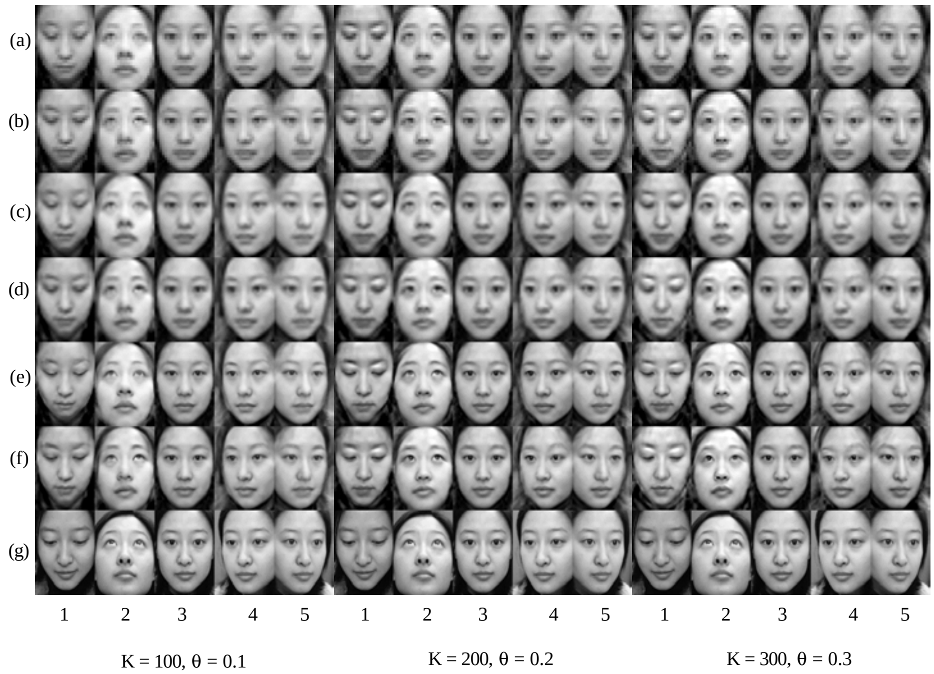

Fig. 7: Input a face image of frontal view Fig. 6a; a) synthesized low-resolution images by Ma's Method; b) synthesized low-resolution images by the proposed method; c) bicubic interpolation of (a, d) bicubic interpolation of (b, e) final HR results by Ma's Method; f) final HR results by the Proposed Method and g) original images

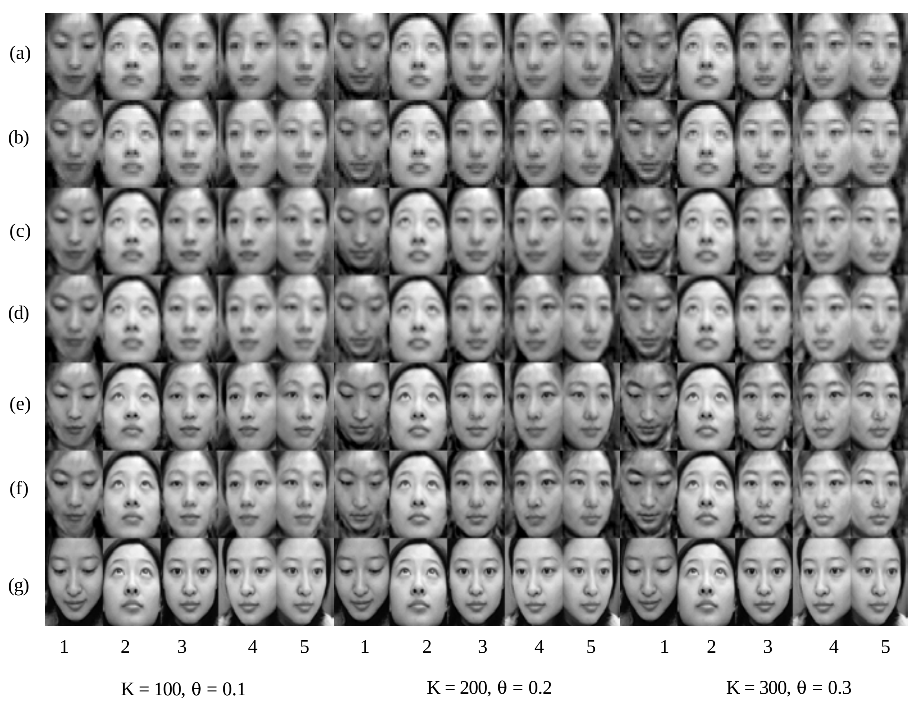

Fig. 8: Input a face image of up view Fig. 6b; a) synthesized low-resolution images by Ma's Method; b) synthesized low-resolution images by the proposed method; c) bicubic interpolation of (a, d) bicubic interpolation of (b, e) final HR results by Ma's Method; f) final HR results by the Proposed Method and g) original images

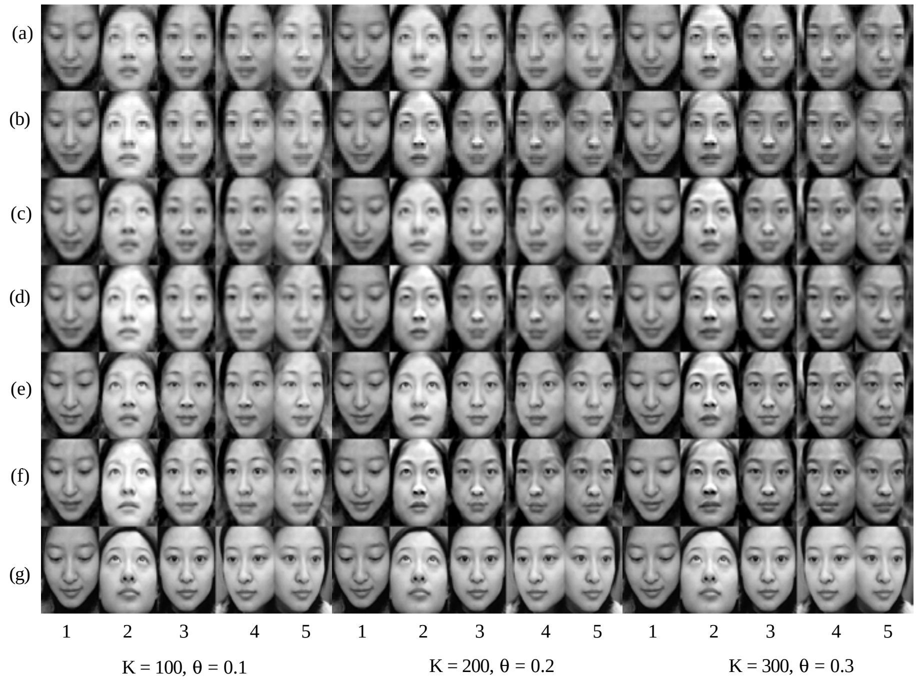

Fig. 9: Input a face image of down view Fig. 6c; a) synthesized low-resolution images by Ma's Method; b) synthesized low-resolution images by the proposed method; c) bicubic interpolation of (a, d) bicubic interpolation of (b, e) Final HR results by Ma's Method; f) final HR results by the Proposed Method and g) original images

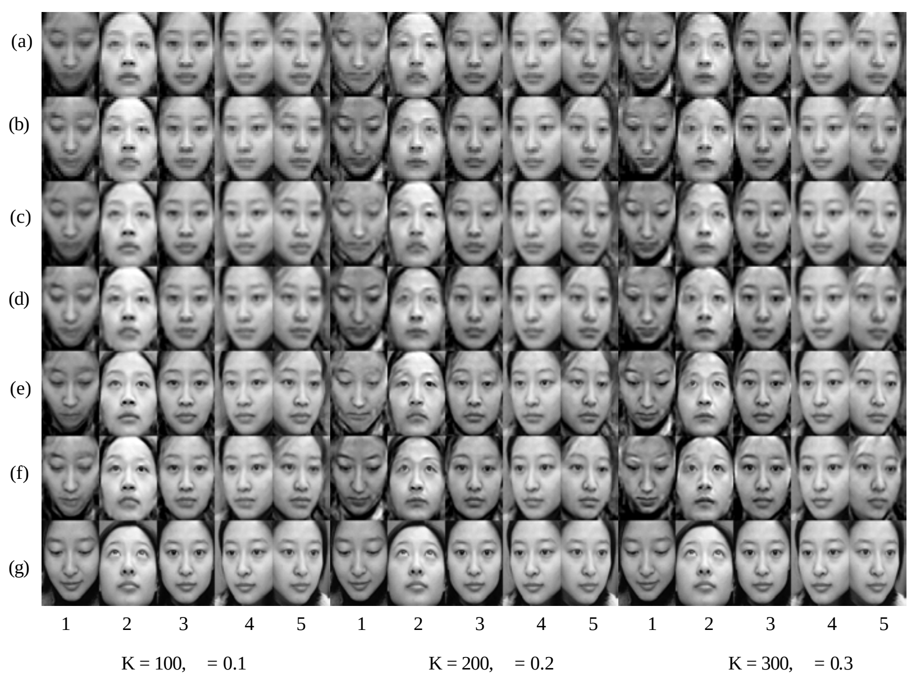

Fig. 10: Input a face image of left view Fig. 6d; a) synthesized low-resolution images by Ma's Method; b) synthesized low-resolution images by the proposed method; c) bicubic interpolation of (a, d) bicubic interpolation of (b, e) final HR results by Ma's Method; f) final HR results by the Proposed Method and g) original images

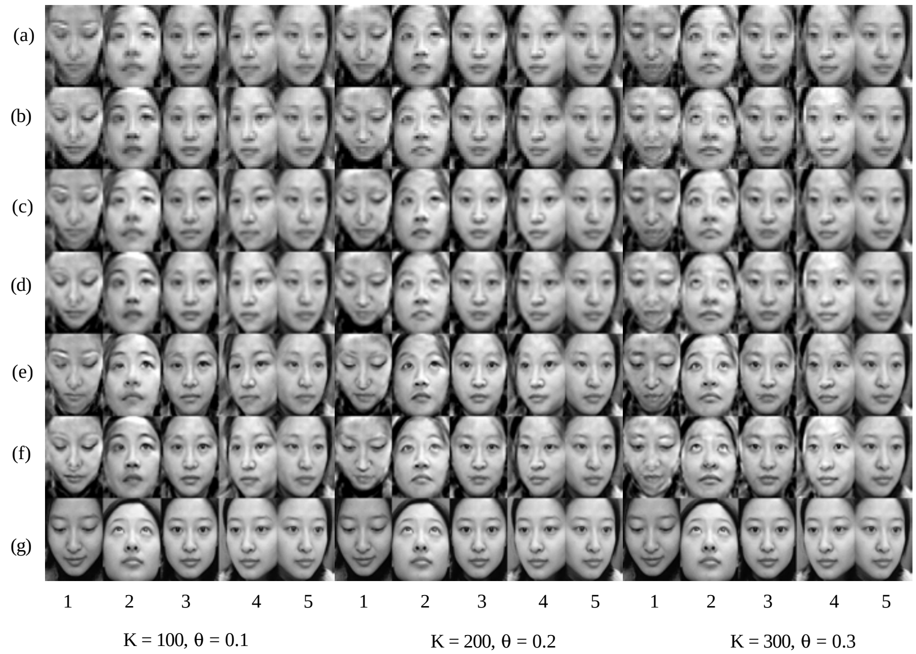

Fig. 11: Input a face image of right view Fig. 6e; a) synthesized low-resolution images by Ma's Method; b) synthesized low-resolution images by the proposed method; c) bicubic interpolation of (a, d) bicubic interpolation of (b, e) Final HR results by Ma's Method; f) final HR results by the Proposed Method and g) original images

Table 1: List of PSNR of Hallucinated image (input a face image on frontal view) K = 100, θ = 0.1

Views	Ma's Method	Our Method
Frontal	32.49	32.89
Up	26.98	29.94
Down	26.59	26.80
Left	31.13	32.47
Right	28.40	30.19

Table 2: List of PSNR of Hallucinated image (input a face image on up view) K = 200, θ = 0.2

Views	Ma's Method	Our Method
Frontal	30.04	31.30
Up	33.64	34.18
Down	28.74	28.88
Left	32.92	34.68
Right	31.36	31.75

Table 3: List of PSNR of Hallucinated image (input a face image on down view) K = 200, θ = 0.2

Views	Ma's Method	Our Method
Frontal	29.61	32.93
Up	28.34	30.75
Down	33.20	34.17
Left	34.61	34.68
Right	32.70	34.47

Table 4: List of PSNR of Hallucinated image (input a face image on left view) K = 200, θ = 0.2

Views	Ma's Method	Our Method
Frontal	31.67	32.39
Up	30.55	31.84
Down	28.19	30.82
Left	33.44	34.29
Right	32.28	32.30

Table 5: List of PSNR of Hallucinated image (input a face image on right view) K = 200, θ = 0.2

Views	Ma's Method	Our Method
Frontal	26.98	28.76
Up	30.54	30.56
Down	26.23	26.24
Left	28.56	31.46
Right	33.12	33.75

Table 1 indicates that the best quality of frontal view image can be found at 0.1 value of similarity threshold θ while Table 2-5 points that the best quality of other views is found at 0.2. In all cases of inputs, final HR images generated by the proposed method show superior image quality over Ma's Method in every view.

CONCLUSION

In this study, a novel multiview Face Hallucination Method using Adaptive Locally Linear Embedding (ALLE) technique. Experimental results show the higher quality of reconstructed image of the proposed framework over those enhanced with the baseline in both interpolation and learning methods based on the same training set. High-resolution face images of five different views are generated from a single low-resolution face image. According to the experimental results, the reconstructed image will be more accurate if the view of the input is same as that of the output. Especially, the frontal, up and down views achieve better estimation than others. The result of the proposed method show superior reconstruction quality of the HR face image over other related methods in both visualization and PSNR. However, the alignment on low-resolution face images may not be accurate. An automatic Face Alignment algorithm will be development in the future research.

ACKNOWLEDGEMENTS

This research has been supported by Research Grant for New Scholar (MRG5480162) from TRF (Thai Research Fund), CHE (Commission on Higher Education) and Panyapiwat Institute of Management, Thailand.

REFERENCES

Baker, S. and T. Kanade, 2000. Hallucinating faces. Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition, March 28-30, 2000, Grenoble, France, pp: 83-88.

Baker, S. and T. Kanade, 2002. Limits on super-resolution and how to break them. IEEE Trans. Pattern Anal. Mach. Intell., 24: 1167-1183.

Chang, H., D. Yeung and Y. Xiong, 2004. Super-resolution through neighbor embedding. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Volume 1, June 27-July 2, 2004, Washington, DC., USA., pp: 275-282.

Elad, M. and A. Feuer, 1999. Super-resolution reconstruction of image sequences. IEEE Trans. Pattern Anal. Mach. Intell., 21: 817-834.

Fan, W. and D.Y. Yeung, 2007. Image hallucination using neighbor embedding over visual primitive manifolds. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 17-22, 2007, USA., pp: 1-7.

Freeman, W.T. and E.C. Pasztor, 1999. Learning low level vision. Proceedings of the 7th IEEE International Conference on Computer Vision, Volume 2, September 20-27, 1999, Kerkyra, pp: 1182-1189.

Gao, G.W. and J. Yang, 2013. A novel sparse representation based framework for face image super-resolution. Neurocomputing. 10.1016/j.neucom. 2012. 12.059.

Gao, W., B. Cao, S.G. Shan, X.L. Chen, D.L. Zhou, X.H. Zhang and D.B. Zhao, 2008. The CAS-PEAL large-scale Chinese face database and baseline evaluations. IEEE Trans. Syst. Man Cybernet. A Syst. Humans, 38: 149-161.

Hardie, R.C., K.J. Barnard and E.E. Armstrong, 1997. Joint map registration and high-resolution image estimation using a sequence of undersampled images. IEEE Trans. Image Process., 6: 1621-1633.

Hu, Y., K.M. Lam, T. Shen and W. Wang, 2011a. A novel kernel-based framework for facial-image hallucination. Image Vision Comput., 29: 219-229.

Hu, Y., K.M. Lam, G. Qiu and T. Shen, 2011b. From local pixel structure to global image super-resolution: a new face hallucination framework. IEEE Trans. Image Process, 20: 433-445.

Jia, K. and S. Gong, 2008. Generalized face super-resolution. IEEE Trans. Image Process., 17: 873-886.

Jung, C., L. Jiao, B. Liu and M. Gong, 2011. Position-patch based face hallucination using convex optimization. IEEE Signal Process. Lett., 18: 367-370.

Lin, Z. and H.Y. Shum, 2004. Fundamental limits of reconstruction-based super resolution algorithms under local translation. IEEE Trans. Pattern Anal. Mach. Intell., 26: 83-97.

Liu, W., D. Lin and X.O. Tang, 2005. Neighbor combination and transformation for hallucinating faces. Proceedings of the IEEE International Conference on Multimedia and Expo, July 6-8, 2005, Netherlands, pp: 478-484.

Ma, X., J.P. Zhang and C. Qi, 2009. Position-based face hallucination method. Proceedings of the IEEE International Conference on Multimedia and Expo, June 28-July 3, 2009, New York, USA., pp: 290-293.

Ma, X., H. Huang, S. Wang and C. Qi, 2010a. A simple approach to multiview face hallucination. IEEE Signal Process. Lett., 17: 579-582.

Ma, X., J.P. Zhang and C. Qi, 2010b. Hallucinating face by position-patch. Pattern Recognit., 43: 2224-2236.

Park, S.C., M.K. Park and M.G. Kang, 2003. Super-resolution image reconstruction a technical overview. IEEE Signal Process. Mag., 20: 21-36.

Park, S.W. and M. Savvides, 2008. Robust super-resolution of face images by iterative compensating neighborhood relationships. Proceedings of the Biometrics Symposium, September 11-13 2007, USA., pp: 1-5.

Roweis, S.T. and L.K. Saul, 2000. Nonlinear dimensionality reduction by locally linear embedding. Science, 290: 2323-2326.

Sun, B.Y., X.M. Zhang, J.Y. Li and X.M. Mao, 2010. Feature fusion using locally linear embedding for classification. IEEE Trans. Neural Networks, 21: 163-168.

Thongdee, M., C. Kimpan and P. Sanguansat, 2012. Adaptive locally linear embedding for face hallucination. Proceedings of the Asian Conference on Information Systems, December 6-8, 2012, Siem Reap, Cambodia.

Wang, X.G. and X.O. Tang, 2005. Hallucinating face by eigentransformation. IEEE Trans. Syst. Man Cybern C Appl. Rev., 35: 425-434.

Zhou, F., B. Wang and Q. Liao, 2012. Super-resolution for face image by bilateral patches. Electron. Lett., 48: 1125-1126.

Zhou, F., B. Wang and Q. Liao, 2014. Super-resolution for facial image using multilateral affinity function. Neurocomputing, 133: 194-208.

Zhuang, Y.T., J. Zhang and F. Wu, 2007. Hallucinating faces: LPH super-resolution and neighbor reconstruction for residue compensation. Pattern Recognit., 40: 3178-3194.