

Research Article

Alignment-Free and High-Frequency Compensation in Face Hallucination

Yen-Wei Chen,^{1,2} So Sasatani,² and Xian-Hua Han²

¹ College of Computer Science and Information Technology, Central South University of Forestry and Technology, Hunan 410004, China

² College of Information Science and Engineering, Ritsumeikan University, Shiga 525-8577, Japan

Correspondence should be addressed to Yen-Wei Chen; chen@is.ritsumei.ac.jp

Received 25 August 2013; Accepted 21 November 2013; Published 12 February 2014

Academic Editors: S. Bourennane and J. Marot

Copyright © 2014 Yen-Wei Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Face hallucination is one of learning-based super resolution techniques, which is focused on resolution enhancement of facial images. Though face hallucination is a powerful and useful technique, some detailed high-frequency components cannot be recovered. It also needs accurate alignment between training samples. In this paper, we propose a high-frequency compensation framework based on residual images for face hallucination method in order to improve the reconstruction performance. The basic idea of proposed framework is to reconstruct or estimate a residual image, which can be used to compensate the high-frequency components of the reconstructed high-resolution image. Three approaches based on our proposed framework are proposed. We also propose a patch-based alignment-free face hallucination. In the patch-based face hallucination, we first segment facial images into overlapping patches and construct training patch pairs. For an input low-resolution (LR) image, the overlapping patches are also used to obtain the corresponding high-resolution (HR) patches by face hallucination. The whole HR image can then be reconstructed by combining all of the HR patches. Experimental results show that the high-resolution images obtained using our proposed approaches can improve the quality of those obtained by conventional face hallucination method even if the training data set is unaligned.

1. Introduction

There is a high demand for high-resolution (HR) images such as video surveillance, remote sensing, and medical imaging because high-resolution images can reveal more information than low-resolution images. However, it is hard to improve the image resolution by replacing sensors because of the high cost, hardware physical limits. Super resolution image reconstruction (SR) is one promising technique to solve the problem [1, 2]. SR can be broadly classified into two families of methods: (1) the classical multiframe super resolution [2] and (2) the single-frame super resolution, which is also known as example-based or learning-based super resolution [3–5]. In the classical multiimage SR, the HR image is reconstructed by combining subpixel-aligned multiimages (LR images). In the learning-based SR, the HR image is reconstructed by learning correspondence between low and high-resolution image patches from a database.

Face hallucination is one of learning-based SR techniques proposed by Baker and Kanade [1, 6], which is focused on

resolution enhancement of facial images. To date, a lot of algorithms of face hallucination methods have been proposed [7–12]. Though face hallucination is a powerful and useful technique, some detailed high-frequency components cannot be recovered. In this paper, we propose a high-frequency compensation framework based on residual images for face hallucination method in order to improve the reconstruction performance. The basic idea of proposed framework is to reconstruct or estimate a residual image, which can be used to compensate the high-frequency components of the reconstructed high-resolution image. Three approaches based on our proposed framework are proposed. We also propose a patch-based alignment-free face hallucination method. In the patch-based face hallucination, we first segment facial images into overlapping patches and construct training patch pairs. For an input LR image, the overlapping patches are also used to obtain the corresponding HR patches by face hallucination. The whole HR image can then be reconstructed by combining all of the HR patches.

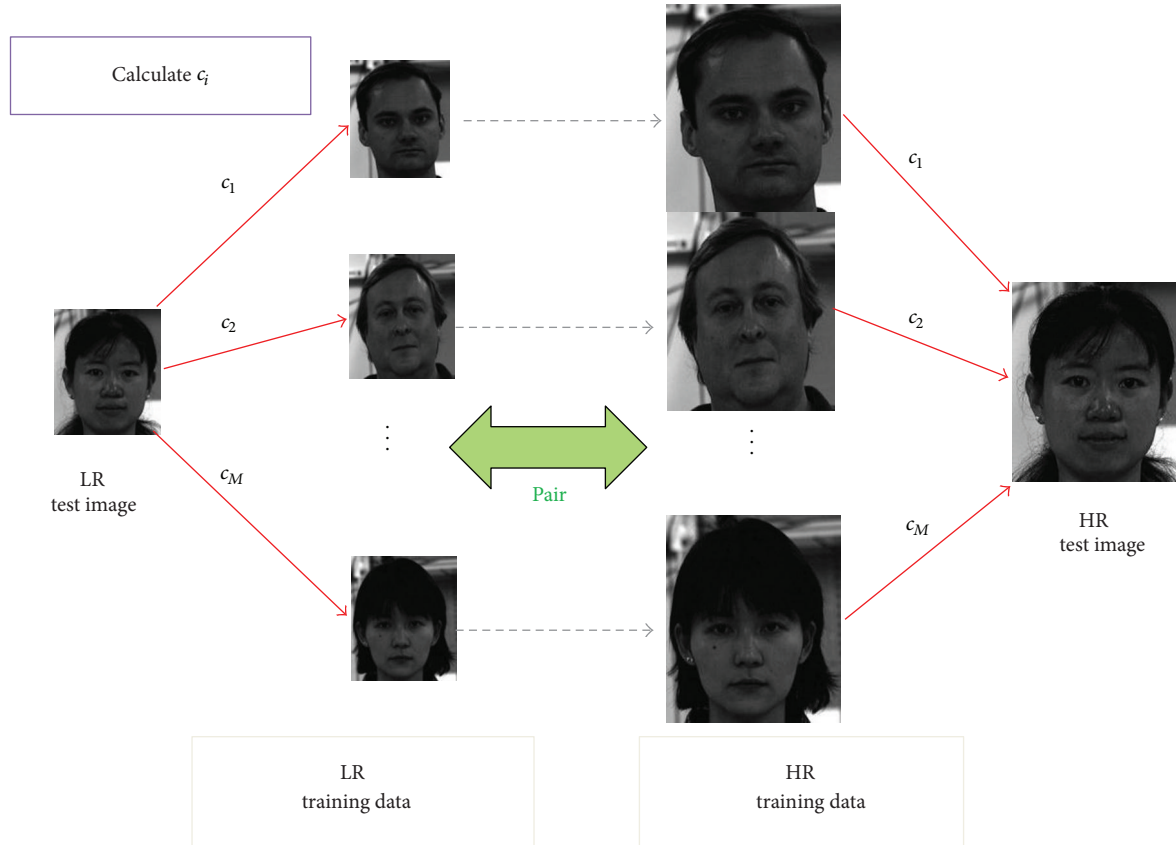


FIGURE 1: Schematic diagram of face hallucination.



FIGURE 2: Framework for recovering high-frequency components in face hallucination.

The paper is organized as follows. In Section 2, we describe the conventional face hallucination method. Our proposed residual image compensation methods are presented in Section 3. Our proposed patch-based alignment-free method is presented in Section 4. Section 5 presents experimental results and quantitative evaluation. Section 6 summarizes our conclusions.

2. Face Hallucination

The face hallucination method is one of learning-based SR methods, which is proposed for resolution enhancement of facial images [6–12]. In this section, we briefly introduce the basic concept of face hallucination, which is shown in Figure 1.

The basic idea of face hallucination is that a face image can be reconstructed from other face images by linear combination because all facial images have a similar structure. In face hallucination, an input LR image can be represented as a linear sum of the LR training images along with some learned coefficients. Due to the correlation between the LR and HR images in the training dataset, the output HR image can also be calculated by finding the linear sum of the corresponding HR images using the same coefficients.

We represent a two-dimensional face image using a column vector of all pixel values, and \mathbf{X}_i represents the input LR face image. HR training images are denoted by $\mathbf{H} = [\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_M]$, and the corresponding LR training images are $\mathbf{L} = [\mathbf{L}_1, \mathbf{L}_2, \dots, \mathbf{L}_M]$, where M is the number of training image pairs. First, we interpolate the LR training images and the input (test) LR image to the resolution space of the

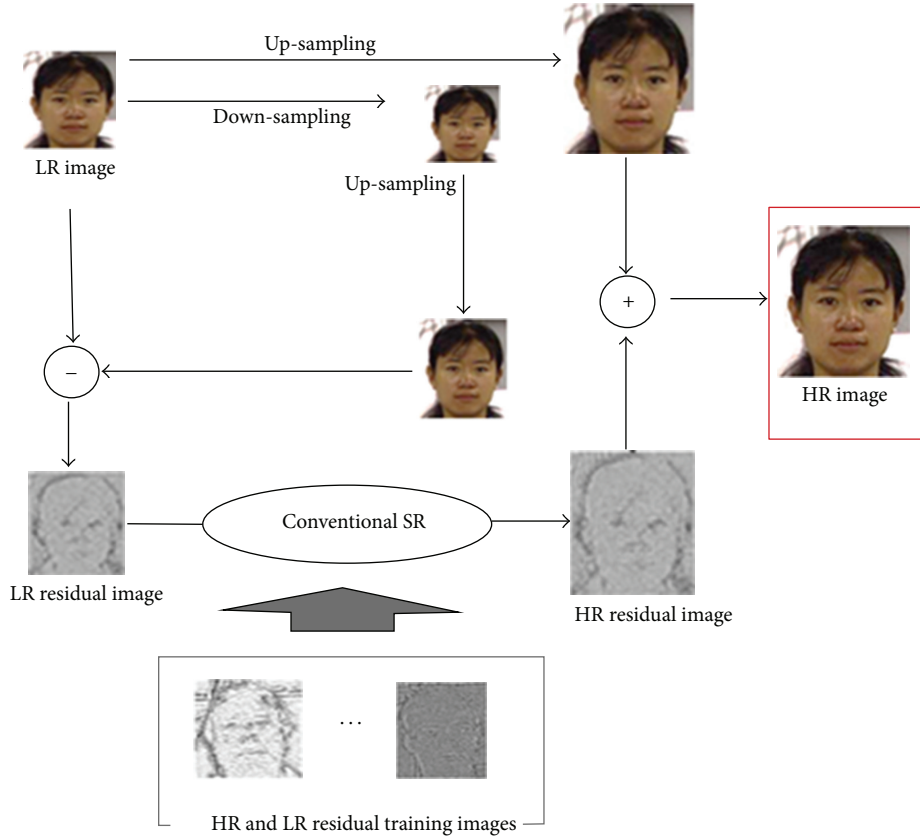


FIGURE 3: Proposed super resolution method 1.

HR training images, denoted by $\tilde{\mathbf{L}} = [\tilde{\mathbf{L}}_1, \tilde{\mathbf{L}}_2, \dots, \tilde{\mathbf{L}}_M]$ and $\tilde{\mathbf{X}}_i$, respectively. $\tilde{\mathbf{X}}_i$ may be represented by a linear sum of interpolated training LR images using

$$\tilde{\mathbf{X}}_i = c_1 \tilde{\mathbf{L}}_1 + c_2 \tilde{\mathbf{L}}_2 + \dots + c_M \tilde{\mathbf{L}}_M, \quad (1)$$

where $\mathbf{C} = [c_1, c_2, \dots, c_M]$ are the weight coefficients, satisfying the following constraint:

$$c_1 + c_2 + \dots + c_M = 1. \quad (2)$$

The optimal weights can be calculated by minimizing the error in reconstructing the input LR image $\tilde{\mathbf{X}}_i$ from training LR images. This error is defined in (3). After substitution of the constraints in (2) into (3), the weight vector may be obtained using covariance matrix \mathbf{S} in (4). So,

$$\varepsilon(\mathbf{C}) = \|\mathbf{X}_h - \tilde{\mathbf{X}}\|^2 = \left\| \mathbf{X}_h - \sum_{i=1}^M c_i \tilde{\mathbf{L}}_i \right\|^2, \quad (3)$$

$$\mathbf{C} = \frac{\mathbf{S}^{-1} \mathbf{1}}{\mathbf{1}^T \mathbf{S}^{-1} \mathbf{1}}, \quad (\mathbf{S} = (\tilde{\mathbf{X}}_i \mathbf{1}^T - \tilde{\mathbf{L}})^T (\tilde{\mathbf{X}}_i \mathbf{1}^T - \tilde{\mathbf{L}})). \quad (4)$$

After obtaining the coefficients for reconstructing the input LR image with LR training images as given in (1), we

replace $\tilde{\mathbf{L}}$ with \mathbf{H} using the same coefficients \mathbf{C} . Subsequently, the HR image \mathbf{X}_h can be obtained using

$$\mathbf{X}_h = c_1 \mathbf{H}_1 + c_2 \mathbf{H}_2 + \dots + c_M \mathbf{H}_M. \quad (5)$$

3. High Frequency Compensations Using Residual Images

Though face hallucination is a powerful and useful technique, some detailed high-frequency components cannot be recovered. In this paper, we propose a high-frequency compensation framework based on residual images for face hallucination method in order to improve the reconstruction performance. The basic idea of proposed framework is to reconstruct or estimate a residual image, which can be used to compensate the high-frequency components of the reconstructed high-resolution image as shown in Figure 2.

In order to estimate the residual image, we construct a new residual image database (pairs of LR and HR residual images) in addition to existing image database (pairs of LR and HR images) for training. The construction of the residual image database is shown in Figure 3. The HR and LR training image pairs $[\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_M]$ and $[\mathbf{L}_1, \mathbf{L}_2, \dots, \mathbf{L}_M]$ already exist in conventional face hallucination. With conventional face hallucination, for each LR training image \mathbf{L} , the other $M - 1$ training image pairs are used to obtain

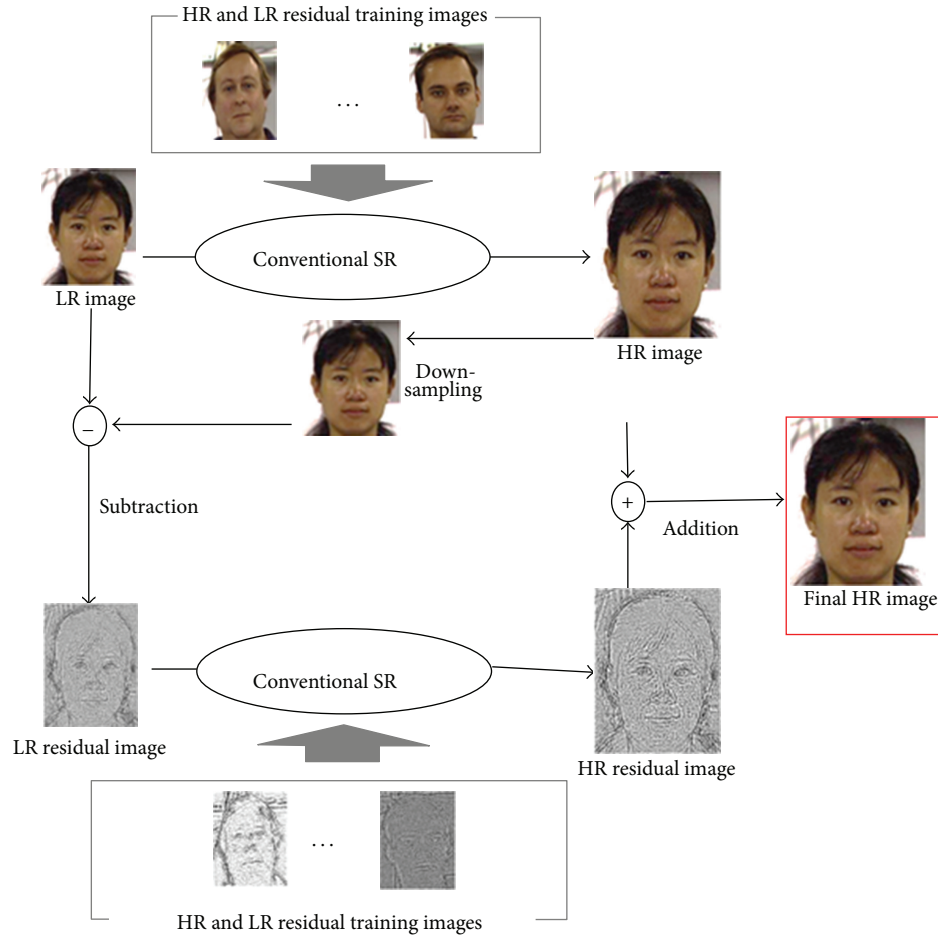


FIGURE 4: Proposed super resolution method 2.

the approximated HR image \hat{H} . The HR training residual image is the difference between the original HR image and the reconstructed HR image \hat{H} , while the LR residual image is the difference between the original LR image and the downsampled version of the reconstructed HR image.

With the two training pair databases, three approaches are proposed for high frequency compensation.

3.1. Proposed Method 1. The first approach is shown in Figure 3. We first use the conventional interpolation method to obtain an HR image and calculate the LR residual image between the input LR image and the downsampled reconstructed HR image. Then we reconstruct the HR residual image from the LR residual image using training residual image pairs. Finally we merge the HR residual and the interpolated HR images.

3.2. Proposed Method 2. The second approach is shown in Figure 4. We first use the conventional face hallucination method to obtain an HR image and calculate the LR residual image between the input LR image and the downsampled reconstructed HR image. Then we reconstruct the HR residual image from the LR residual image using training residual

image pairs. Finally we merge the HR residual and the reconstructed HR images.

3.3. Proposed Method 3. The third approach is shown in Figure 5. We first use our proposed SR method 1 to obtain an HR image and calculate the LR residual image between the input LR image and the downsampled reconstructed HR image. Then we reconstruct the HR residual image from the LR residual image using training residual image pairs. Finally we merge the HR residual and the reconstructed HR images.

4. Alignment-Free Patch-Based Face Hallucination

In the conventional face hallucination (Figure 6(a)), whole face image is used in training and SR reconstruction. Each pixel is reconstructed by the use of the corresponding pixel pairs at the same position. So the conventional face hallucination needs an accurate alignment of facial images in order to obtain good reconstructed HR images. With some imperfectly aligned facial datasets, it is difficult to reconstruct sharp HR images using face hallucination. In this paper, we propose a patch-based face hallucination. The basic

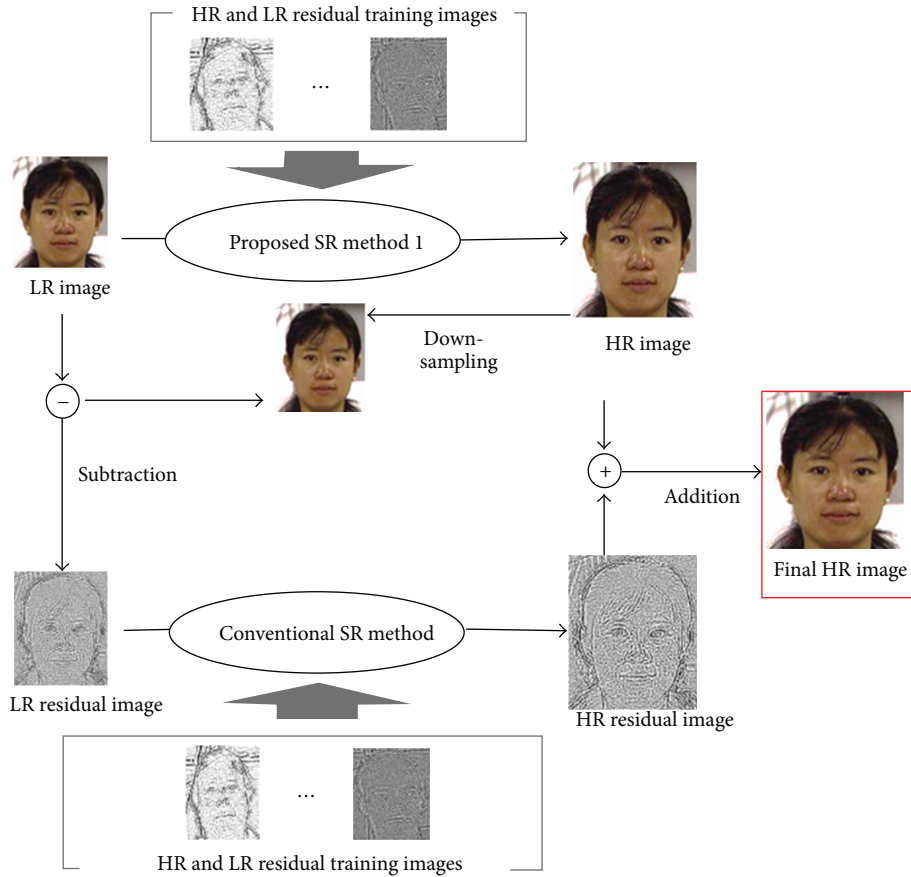


FIGURE 5: Proposed super resolution method 3.

idea of our proposed method is shown in Figure 6(b). In our proposed alignment-free patch-based face hallucination, we first segment facial images into overlapping patches and construct training patch pairs. For reconstruction of HR image, each overlapping patch in the input LR image is used as a template and its corresponding patches in LR training images are found by the use of an SSD (sum of squared difference) based template matching. Then its HR patches can be obtained by face hallucination. The whole HR image can be reconstructed by combining all of the HR patches.

5. Experimental Results

In order to validate the effectiveness of our proposed methods, we apply our proposed methods to two face databases. The first one is our developed MaVIC database (multiangle View, illumination and cosmetic facial image database) [13], which contains 99 aligned images of different persons and the size of each image is 320×400 . The second one is C&P database provided by Kanade et al. [14] and Pie [11], which contains 165 imperfectly aligned frontal face images, and each image size is 264×320 . We first generate the LR and HR image pairs by downsampling the original images. The size of LR images is 50×61 , while the size of HR images is 200×244 . The leave-one-out method is used in our experiments. In each database, we select one LR image

randomly as a test image and its HR image is used as a ground truth image for quantitative evaluation. Other image pairs are used for training. Our proposed three methods are used for HR reconstruction of the LR test image. In order to make a comparison, the conventional face hallucination method and the bicubic interpolation method are also used for reconstructions. For each method, a total of 20 experiments with a different test image are performed. The peak signal-to-noise ratio (PSNR) [dB] is used as a quantitative measure for evaluation of the HR reconstruction performance. For C&P's imperfectly aligned facial datasets, our proposed patch face hallucination method is used with a patch size of 3×3 and a 1×1 patch that overlaps with adjacent patches.

Firstly, we show experimental results with the aligned face database MaVIC. A typical example is shown in Figure 7. The test LR image, which is not included in the training samples, is shown in the upper left of Figure 7. The grand truth HR image is shown in the lower right of Figure 7. Others are reconstructed HR images with their PSNR by different methods. It can be seen that the reconstructed high-resolution images obtained using our proposed approaches are much better than those obtained by conventional face hallucination method and bicubic interpolation method and the proposed method 3 shows the best performance among three proposed methods. Similar results have also been obtained with other test images. PSNR evaluation results for

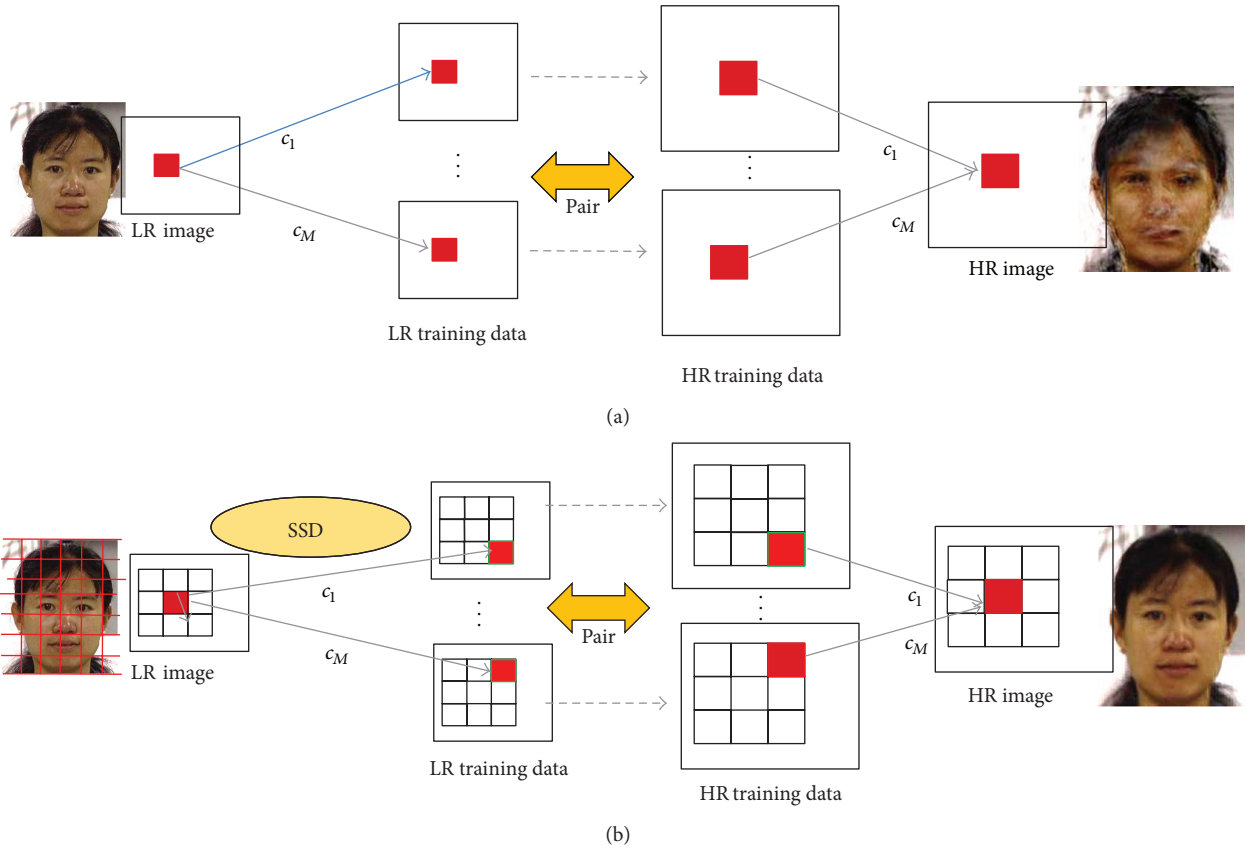


FIGURE 6: (a) Conventional face hallucination, (b) our proposed alignment-free patch-based face hallucination.

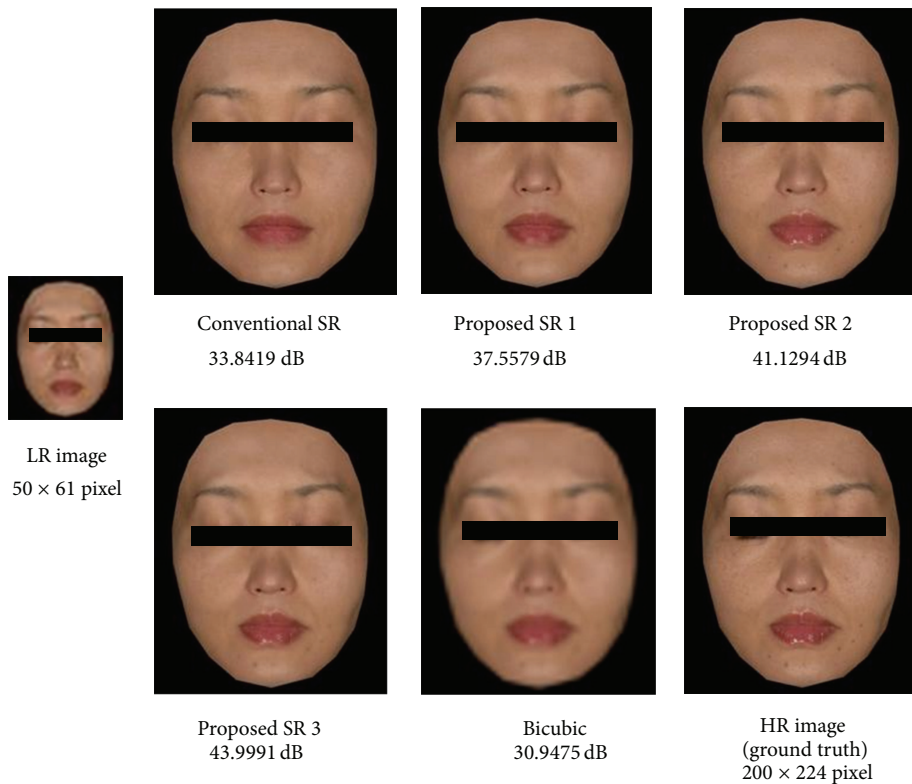


FIGURE 7: Typical reconstructed images for an aligned face image database (MaVIC).

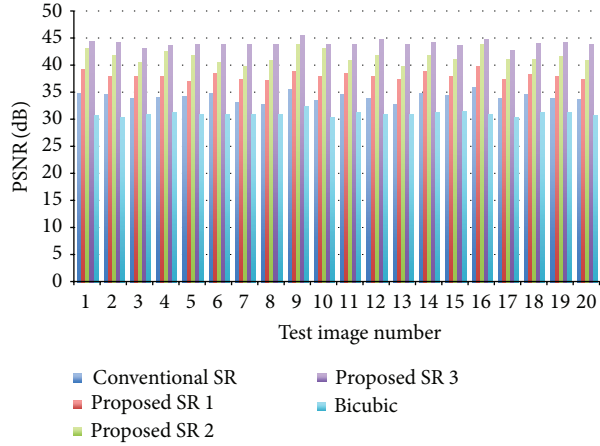


FIGURE 8: PSNR of 20 test images (MaVIC) by different methods.

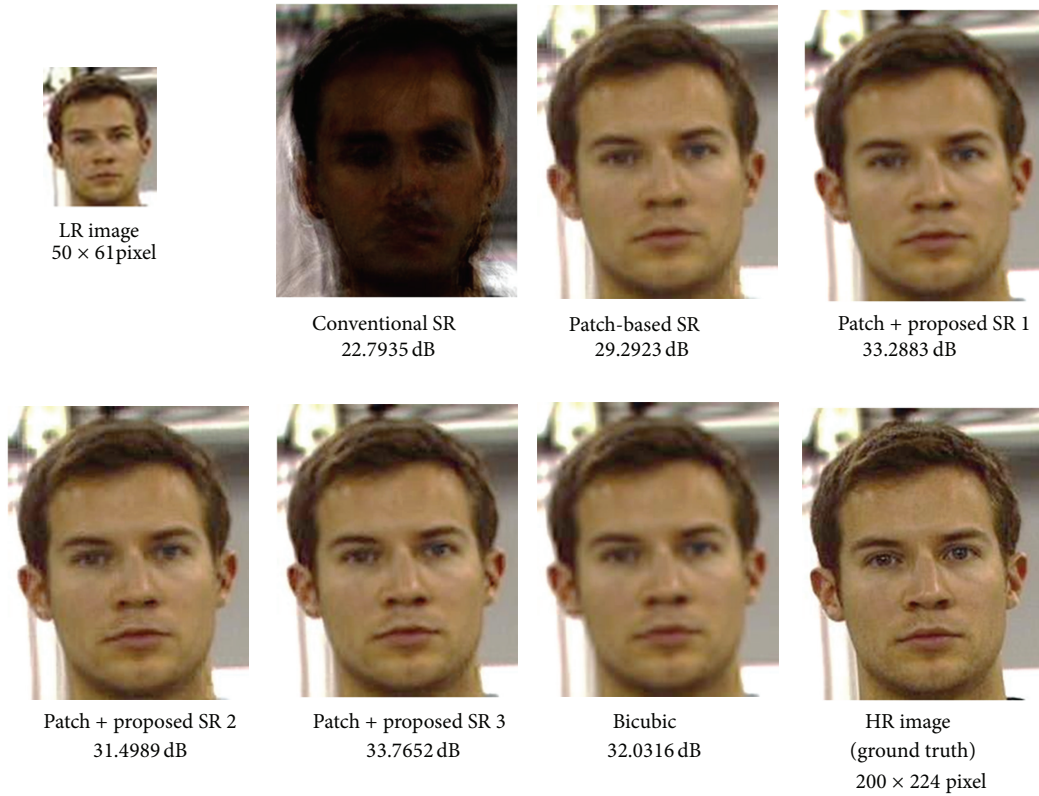


FIGURE 9: Typical reconstructed images for a nonaligned face image database (C&P).

all test images are shown in Figure 8. The image shown in Figure 7 is corresponding to image no. 20. The mean and standard deviation over 20 experiments for each method are summarized in Table 1. It can be seen that the averaged PSNR with bicubic interpolation is about 31.16. The averaged PSNR can be improved to 34.35 [dB] by the use of conventional face hallucination. Significant improvements can be achieved by our proposed high-frequency compensation methods. The proposed method 3 shows the best performance among three proposed methods. The averaged PSNR is improved to 44.17 [dB].

TABLE 1: Comparison of the averaged PSNR for an aligned face database (MaVIC).

Method	Conventional SR [10]	Proposed SR 1	Proposed SR 2	Proposed SR 3	Bicubic
Mean	34.35	38.18	41.74	44.17	31.16
Stv.	0.81	0.72	1.18	0.60	0.45

Next, we show experimental results with the unaligned face database (C&P). Unlike experiments with MaVIC, our proposed patch-based method is used for reconstruction of

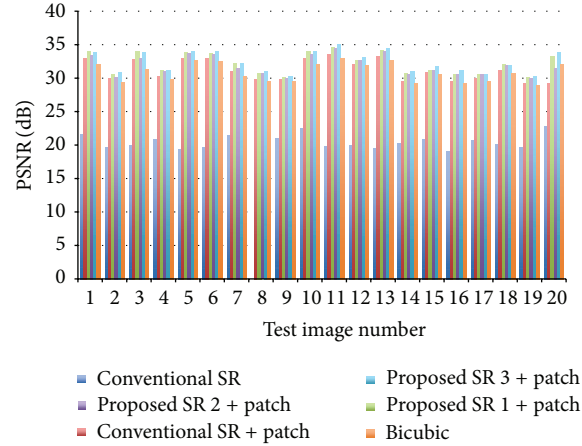


FIGURE 10: PSNR of 20 test images (C&P) by different methods.

TABLE 2: Comparison of the averaged PSNR for a nonaligned face image database (C&P).

Methods	Conventional SR [10]	Conventional SR + patch	Proposed SR 1 + patch	Proposed SR 2 + patch	Proposed SR 3 + patch	Bicubic
Mean	20.42	31.21	32.49	32.19	32.96	30.86
Stv.	1.03	1.54	1.60	1.50	1.58	1.37

HR images. A typical example is shown in Figure 9. The test LR image, which is not included in the training samples, is shown in the upper left of Figure 9. The grand truth HR image is shown in the lower right of Figure 9. Others are reconstructed HR images with their PSNR by different methods. PSNR evaluation results for all test images are shown in Figure 10. The image shown in Figure 9 is corresponding to image no. 20. The mean and standard deviation over 20 experiments for each method are summarized in Table 2. As we discussed in previous section, conventional face hallucination needs an accurate alignment of facial images in order to obtain good reconstructed HR images. For the unaligned face database C&P (usually existing face databases are not aligned), the conventional SR method, which uses whole face image for super resolution, gives a very poor result, while if we use our proposed patch-based framework, we can significantly improve the reconstructed HR image even with the conventional face hallucination method. The averaged PSNR can be improved from 20.42 [dB] to 31.21 [dB]. It means that our proposed method is an alignment-free SR method. Furthermore, the HR image with higher quality can be achieved by our proposed high-frequency compensation methods. As well as the results with MaVIC, the proposed method 3 shows the best performance among three proposed methods. The averaged PSNR is improved to 32.96 [dB]. Though smaller patch size may improve the quality of reconstructed HR images, it will increase the computation cost. Optimum patch size depends on the purpose and applications.

6. Conclusions

We proposed a residual image compensation framework together with a patch-based alignment-free method to improve the reconstruction quality for face hallucination. The basic idea of our proposed residual image compensation framework was to reconstruct or estimate a residual image, which can be used to compensate the high-frequency components of the reconstructed high-resolution image. Three approaches based on our proposed framework were proposed. In the patch-based alignment-free face hallucination, we first segmented facial images into overlapping patches and constructed training patch pairs. For an input LR image, the overlapping patches are also used to obtain the corresponding HR patches by face hallucination. The whole HR image can then be reconstructed by combining all of the HR patches. The effectiveness of our proposed methods has been demonstrated on both the aligned face database (MaVIC) and the unaligned face database (C&P). The reconstructed high-resolution images obtained using our proposed approaches are much better than those obtained by conventional face hallucination method and bicubic interpolation method. The averaged PSNR of reconstructed HR images was improved from 34.35 [db] to 44.17 [dB] for the aligned face database (MaVIC) and from 20.42 [dB] to 32.40 [dB] for the unaligned face database (C&P). The proposed high-frequency compensation method 3 shows the best performance among three proposed approaches.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is supported in part by the MEXT Support Program for the Strategic Research Foundation at Private Universities (2013–2017) and in part by the R-GIRO Research Fund from Ritsumeikan University. Images in MaVIC database are provided by Kao Co., Ltd. Authors would like to thank Mr T. Igarashi of Kao Co., Ltd., Mr. Y. Iwamoto, K. Taniguchi, and S. Nojima for their useful discussions and assistance in this paper and experiments.

References

- [1] S. Baker and T. Kanade, "Limits on super-resolution and how to break them," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 9, pp. 1167–1183, 2002.
- [2] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: a technical overview," *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 21–36, 2003.
- [3] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based super-resolution," *IEEE Computer Graphics and Applications*, vol. 22, no. 2, pp. 56–65, 2002.
- [4] H. Chang, D.-Y. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '04)*, vol. 1, pp. 1275–1282, July 2004.
- [5] J. Qiao, J. Liu, and Y.-W. Chen, "Joint blind super-resolution shadow removing," *IEICE Transactions on Information and Systems*, vol. E90-D, no. 12, pp. 2060–2069, 2007.
- [6] S. Baker and T. Kanade, "Hallucinating faces," in *Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition*, March 2000.
- [7] C. Liu, H.-Y. Shum, and W. T. Freeman, "Face hallucination: theory and practice," *International Journal of Computer Vision*, vol. 75, no. 1, pp. 115–134, 2007.
- [8] X. Ma, J. Zhang, and C. Qi, "An example-based two-step face Hallucination method through coefficient learning," in *Proceedings of the International Conference on Image Analysis and Recognition (ICLAR '09)*, vol. 5627 of *Lecture Notes in Computer Science*, pp. 471–480, 2009.
- [9] X. Wang and X. Tang, "Face hallucination and recognition," in *Proceedings of the Audio- and Video-Based Biometric Person Authentication (AVBPA' 03)*, vol. 2688 of *Lecture Notes in Computer Science*, p. 1054, 2003.
- [10] X. Wang and X. Tang, "Hallucinating face by eigentransformation," *IEEE Transactions on Systems, Man and Cybernetics C*, vol. 35, no. 3, pp. 425–434, 2005.
- [11] S. Sasatani, X.-H. Han, T. Igarashi, M. Ohashi, Y. Iwamoto, and Y.-W. Chen, "High frequency compensated face hallucination," in *Proceedings of the 18th IEEE International Conference on Image Processing (ICIP '11)*, pp. 1561–1564, September 2011.
- [12] Y. Liang, J.-H. Lai, W.-S. Zheng, and Z. Cai, "A survey of face hallucination," in *Proceedings of the 7th Chinese Conference on Biometric Recognition (CCBR '12)*, vol. 7701 of *Lecture Notes in Computer Science*, pp. 83–93, Springer.
- [13] Y. W. Chen, T. Fukui, X. Qiao, T. Igarashi, K. Nakao, and A. Kashimoto, "Multi-angle view, illumination and cosmetic facial image database (MaVIC) and quantitative analysis of facial appearance," in *Structural, Syntactic, and Statistical Pattern Recognition*, vol. 5342 of *Lecture Notes in Computer Science*, pp. 411–420, Springer, 2008.
- [14] T. Kanade, J. F. Cohn, and Y. Tian, "Comprehensive database for facial expression analysis," in *Proceedings of the IEEE International Conference on Automatic Face & Gesture Recognition*, pp. 46–53, 2000.