ISSN No. 0976-5697

Volume 8, No. 3, March – April 2017

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

A Novel Facial Image Resolution Enhancement Algorithm using Pixels Homogeneity

Vikram Mutneja
Ph.D. Research Scholar
I.K. Gujral Punjab Technical University
Kapurthala, Punjab (India)

Dr. Satvir Singh
Associate Professor
I.K. Gujral Punjab Technical University
Kapurthala, Punjab (India)

Abstract: Facial image processing is an important step in the surveillance cameras based biometric identification systems. The resolution of the facial images from surveillance videos is usually very less owing to the factors such as hardware constraints, distance between the subjects and CCTVs etc. In this work, a novel spatial domain based face resolution enhancement technique using single LR image has been proposed. It is based on the homogeneity levels of the pixels in respective neighboring four quadrants in 3×3 scanning window, used as LR patch which is replaced by interpolated 5×5 HR patch. The kernel matrix is generated for all the pixels in the source image which encodes the homogeneity of all pixels in image. The performance analysis of proposed technique has been done for different values of resolution enhancement factors, in terms of quality metrics Structural Similarity Index Measure (SSIM), Peak Signal to Noise Ratio (PSNR), Signal to Noise Ratio (SNR) and Mean Square Error (MSE). From the results, we contend that the proposed algorithm is very effective in generating good quality HR facial image from single LR facial image.

Keywords: Face Resolution Enhancement, Low Resolution Surveillance Videos, Spatial Domain Processing, Neighbourhood Processing, Image Quality Enhancement

I. INTRODUCTION

Most of the times images of people received from the surveillance videos are of low resolution because of different factors such as distance between the camera and place of footage, environment factors, wide coverage area, installation problems, out of focus and bandwidth issue etc. There are hardware constraints such as storage space limitations because of which the frames need to be compressed or converted to lower resolution for storage. Therefore it is always required to preprocess the frames from surveillance videos before passing them to the biometric identification system. The performance of the biometric systems largely depends upon the quality of the input facial images in terms of image as well as face quality metrics. Out of the various properties to judge the quality of the input facial images, resolution of faces is very important factor.

The face biometric systems can be broadly categorized into two types, local and global. Local methods are based upon using position and/or size of facial landmarks such as eyes, nose tip, lip center, chin center etc. [19, 13], and while global are independent of measurement of facial landmarks. Supposedly the facial images, in which the facial landmarks are localizable are of better quality. Secondly the accuracy of the system also depends upon the accuracy in locating the facial landmarks. The chances of accurately locating the facial landmarks is directly related the face resolution. It is easier to locate the facial landmarks in high resolution images, which in turn increases the hit rate of the system. Face resolution enhancement aims at enhancing the overall signal strength of the facial image by applying the resolution enhancement technique which are classified in two broad categories. Firstly reconstruction based method such as works done by [1, 9, 19]. Secondly learning based techniques such as [21, 18]. Some of the works have done hybridization of both techniques such as by [20]. Reconstruction based models are based on the generalized smoothness priors while the learning based techniques select recognition based priors. Most of the spatial domain based techniques are based on neighbor embedding [2, 5, 11].

Face resolution enhancement techniques are also classified based on number of input facial images. In super-resolution, the final enhanced HR image is computed from multiple images which may be captured from consecutive frames of video or other source. Other are the techniques which use single LR facial input image using relative intensity level of the pixels in surrounding quadrants within N×N LR patch, N>=2. Normally learning based techniques, also known as face hallucination, are based on single facial image as input and use the trained system or dictionary to estimate the missing HR details.

This work is based on neighborhood embedding using single LR facial input image. This paper is structured as follows: Section II provides Literature Survey, Section III describes experimental setup. The proposed technique based on pixels homogeneity has been explained in Section IV, Section V discusses experimental results, and finally Section VI gives conclusions and future scope.

II. LITERATURE SURVEY

[3] Authors proposed a novel super resolution technique of noisy facial images by using the contour features to suppress noise and used the standard deviation prior to enhance the low quality contour feature. Authors demonstrated the superiority of their method as compared to existing techniques using simulations as well as real-world scenario based experiments.

[14] Proposed a new method of face resolution enhancement using a recursive error back projection of example based learning. Authors exploited the use of shape and texture information from low resolution facial image in spatial domain to estimate optimal coefficients for a linear

combination of prototypes of shape and those of texture by solving least square minimization to reconstruct its high resolution version.

[10] Proposed facial image super-resolution method using missing image interpolation based on smooth regression with a novel local structure prior (SRLSP). However the proposed technique is based on assumption of uniformity in local structures of face image patches at same position so as to enable smooth regression function to learn the relationship between LR pixels and missing HR pixels of respective patches. The proposed method was tested on

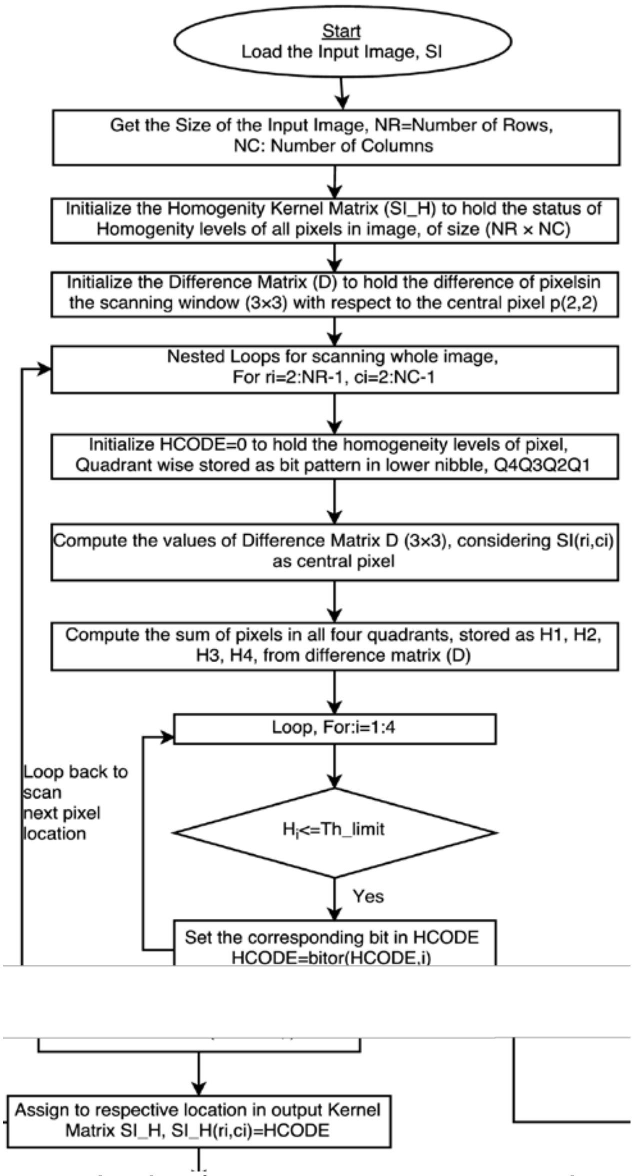

Figure 2: Flow Chart of Function to Generate Homogeneity Kernel Matrix

Yale-B database and reported the results superior to state-of-the-art face super-resolution methods in face recognition.

[6] Proposed a novel sparse coding based face hallucination method by incorporating intrinsic geometric structure of training samples for dictionary learning to minimize artificial effects in reconstructed HR face image. The method is based on graph construction in HR manifold and K-selection mean constraints for finding optimal weight HR face reconstruction. [17] Proposed single image super-resolution using sparse representation and neighbor embedding in learning based approaches using partitioned feature space and statistical prediction model. [18] Emphasized on reducing the dependence of proposed face hallucination technique on size and number of training samples using direct combined approach. [8] Used single LR input and used the concepts such appearance and geometrical features, optical flow model to estimate the local structure in resultant HR image. [4] is also based on

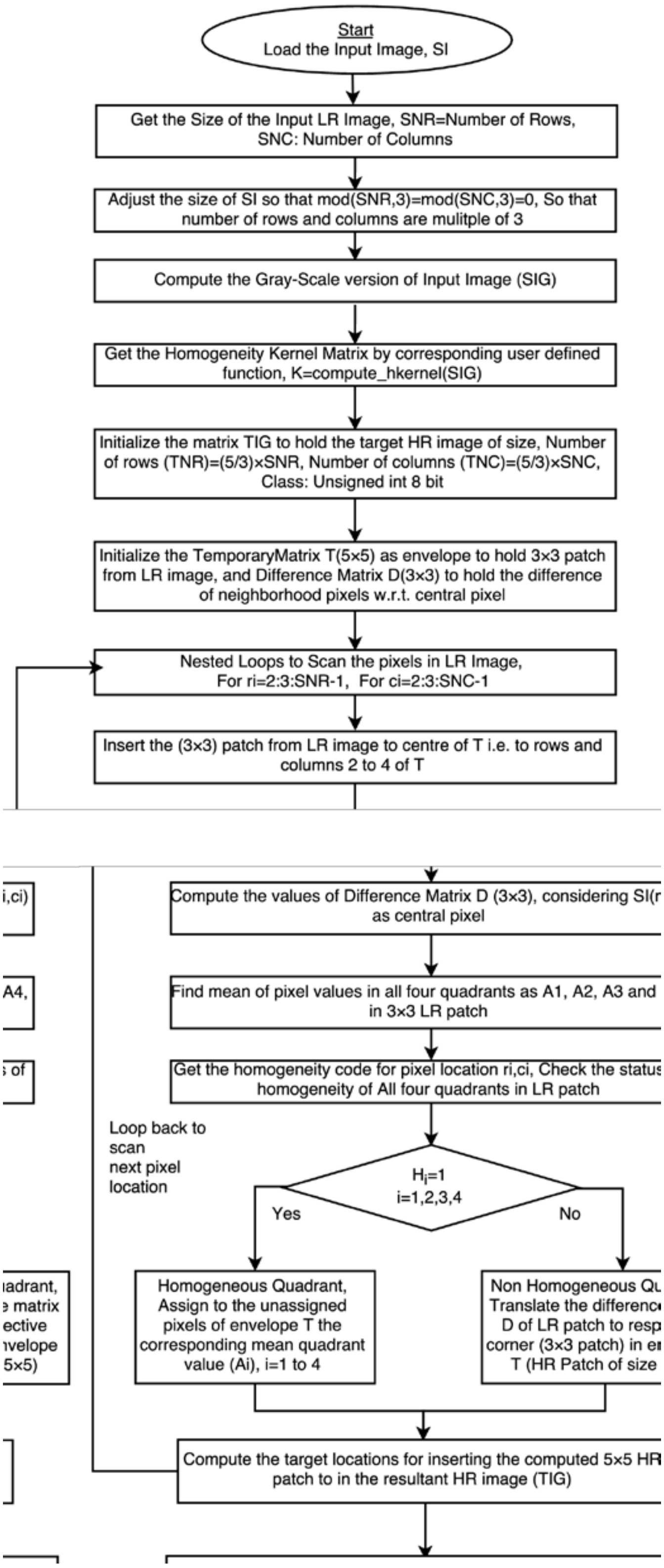

Figure 3: Flow Chart Proposed Technique

single image face hallucination which uses similarity measurements between each input LR patch and all local LR and HR neighborhood patches of training images.

III. EXPERIMENTAL SETUP

The machine running windows 7 (64 Bit) on Intel core i3 1.9 GHz and having NVIDIA graphics processing unit GeForce GT 740M has been used to develop and test the proposed method. The work has been done in MATLAB version 8.5.0.197613 (R2015a).

IV. PROPOSED TECHNIQUE

The proposed technique works on the basis of creating homogeneity kernel matrix for all the pixels of input LR image. The scanning is done in 3×3 local window size, which is slid over whole input LR image pixel by pixel. The homogeneity level corresponding to four quadrants (refer Figure 3) is encoded based on criteria as per Equation 1.

$$\sum_{i=1}^{4} [p_{ij}] < th\_limit \tag{1}$$

Where j denotes the particular quadrant (j=1,2,3,4) and i iterates for all the pixels in respective quadrant. The parameter *th_limit* has been set empirically to value 15. Figure 2 shows the working of function compute_hkernel(SI) for generating the kernel matrix and figure 2 the flow chart of proposed technique to generate the HR image by applying the homogeneity kernel matrix to input LR image in gray-scale version. To apply the proposed technique on the color image, the algorithm depicted in figure 3 is applied recursively to channels R, G and B to generate the resultant color image.

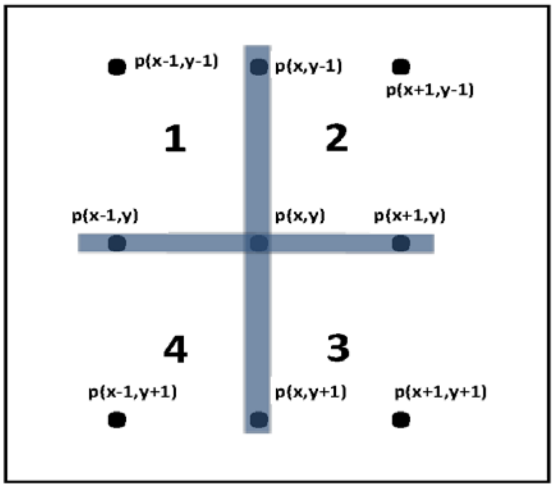

Figure 1: 3×3 Pixels Scanning Window Showing Quadrants (1-4)

Table 1 shows the comparison of homogeneity levels of pixels in the facial images at various stages of application of proposed resolution enhancement technique. It includes input low resolution images for the tests corresponding to single iteration and double iterations, original HR Image and generated HR images. In the types of the homogeneity column, Q refers to quadrant, numeric value as number of quadrants, H as horizontal, V as vertical, L as lower, U as upper, LE as left, R as right and C as cross. e.g. 3QRU can be read as 3 quadrants, right upper. The numeric values attached with H in first column specifies the position number of quadrant e.g. H23 means, homogeneity has been found in quadrants 2 and 3. (refer Figure 1). Figure 4 shows the diagrammatic representation of impact of resolution enhancement on variations of homogeneity levels in image.

From Table 1 and Figure 4 it is inferred that there is rise in percentage of pixels of different homogeneity levels (except for H12, H34 and H1234) with increase in image resolution, being maximum in original HR image.

Table 1: Comparison Number of Pixels Having Different Homogeneity Levels in Input LR, Intermediate and Output HR Facial Images

Homogenity Levels	***Type of Homogeneity***	***Input LR Image (1 Iteration Test)***	***Input LR Image (2 Iterations Test)***	***Intermediate Image (2 Iterations Test)***	***Generated HR Image (1 Iteration Test)***	***Generated HR Image (2 Iterations Test)***	***Original HR Image***
	Image Size	***241×269***	***145×161***	***245×270***	***405×450***	***410×455***	***401×449***
	Type of Homogeneity						
H0	0Q	155	80	80	143	183	708
H1	1Q	183	118	155	248	264	475
H2	1Q	316	178	284	428	431	604
H12	2QHU	701	458	1044	1518	1939	1661
H3	1Q	89	49	100	120	147	258
H13	2QCN	24	8	5	13	1	89
H23	2QVR	1033	505	1172	2306	2403	2691
H123	3QRU	1040	625	1049	2094	1834	2562
H4	1Q	99	63	111	130	128	298
H14	1QVLE	907	492	1020	1973	2215	2166
H24	2QCP	6	1	4	4	1	67
H124	3QLEU	1044	616	1073	1741	1742	2393
H34	2QHL	826	581	1143	1503	1874	1678
H134	3QLEL	740	377	796	1602	1444	2146
H234	3QRL	664	287	702	1325	1391	1812
H1234	4Q	55986	18299	56386	165396	168827	158745

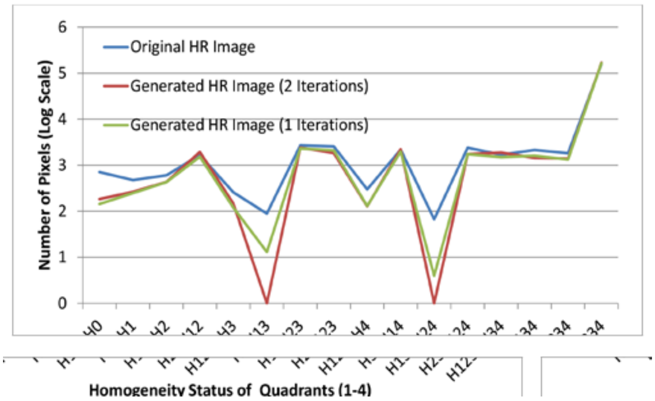

Figure 4: Comparison Homogeneity Levels of Pixels

V. Results and Discussion

The proposed technique was tested on the test images from the database Color FERET [15,16]. Portions of the research in this paper use the FERET database of facial images collected under the FERET program, sponsored by the DOD Counterdrug Technology Development Program Office FERET [16, 15]. The quality metrics used to evaluate the performance of the proposed technique are Structural Similarity Index Measure (*SSIM*), Peak Signal to Noise Ratio (*PSNR*), Signal to Noise Ratio (*SNR*) and Mean Square Error (*MSE*), which have been measured of generated HR image with respect to the original HR image. *PSNR* gives the measure of peak signal to noise power ratio, while *SNR* uses average signal power.

The proposed method has been tested on color as well as gray-scale images. Figures 5 and 6 show the results of application of proposed technique for scaling factor 2.7779 on gray-scale images corresponding to iteration 1 and 2 respectively while figure 7 shows the results for scaling factor 1.6667. Table 2 shows the corresponding measured values of quality metrics. Figures 8, 9 10 show the corresponding results of application of proposed technique on original color images and Table 3 the measured values of quality metrics. By comparison of results in Table 2 and Table 3, it is inferred that application of proposed technique with color based processing (i.e. of channels R,G and B) gives better results as compared to gray-scale only based processing.

Table 2: Quality Metrics for Resolution Enhancement (Gray-Scale)

Quality Metric	*Output HR Image (Single Iteration) Size(450×405), SF=1.667*	*Intermediate Image (Double Iteration) Size(245×270), SF=1.667*	*Output HR Image (Double Iteration) Size(450×405), SF=2.7789*
SSIM	0.919	0.924	0.856
PSNR	34.959	32.854	29.434
SNR	28.096	25.995	22.571
MSE	20.759	33.702	74.074

To study the efficiency of the proposed technique for different resolution enhancement factors and different resolutions of input facial image (for same HR output), the test images from FERET are down-scaled to the factors 0.6000, 0.3600, 0.2160, 0.1296 and 0.0778. The corresponding results are shown in table 4. We have been able to achieve good results in terms of various quality metrics.

To compare the performance of the proposed algorithm with the previous work, a test dataset of hundred images taken from color FERET was created. The resolution enhancement factor was set as 2.7889. Table 5 shows the comparison of the proposed algorithm with some of the previous works in terms of SSIM and PSNR (db), computed as mean value by application of proposed algorithm on this test dataset.

Figure 5: Result Homogeneity Technique on Gray Scale Image for resolution enhancement factor 1.667×1.667, First Stage

Figure 6: Result Homogeneity Technique on Gray Scale Image for resolution enhancement factor 1.667×1.667, Second Stage

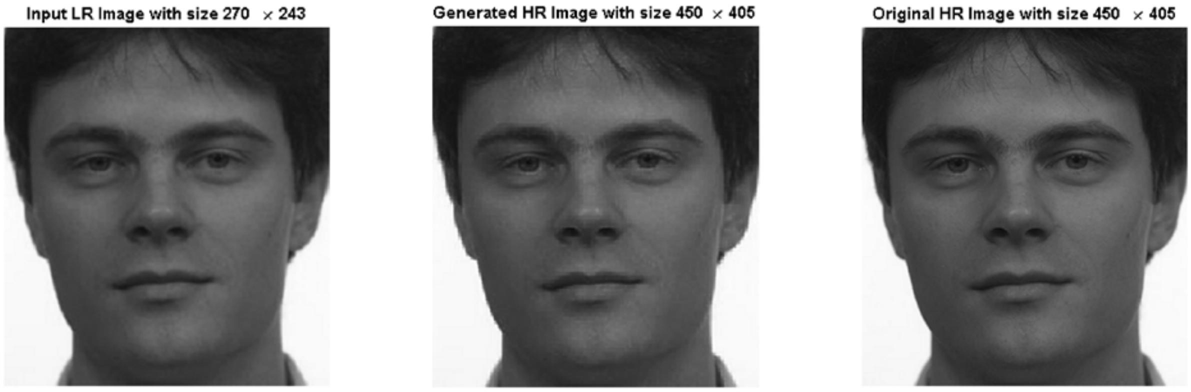

Figure 7: Result Homogeneity Technique on Gray Scale Image for resolution enhancement factor 1.667

Figure 8: Result Homogeneity Technique on Color Image for resolution enhancement factor 1.667×1.667, First Stage

Figure 9: Result Homogeneity Technique on Color Image for resolution enhancement factor 1.667×1.667, Second Stage

Figure 10: Result Homogeneity Technique on Color Image for resolution enhancement factor 1.667

Table 3: Quality Metrics for Resolution Enhancement (Color Image)

Quality Metric	***Output HR Image (Single Iteration) Size(450×405), SF=1.667***	***Intermediate Image (Double Iteration) Size(245×270), SF=1.667***	***Output HR Image (Double Iteration) Size(450×405), SF=2.7789***
SSIM	0.963	0.959	0.926
PSNR	34.343	32.263	28.955
SNR	27.389	25.312	22.000
MSE	23.921	38.613	82.709

Table 4: Results For Various Image Resolution Enhancement Factors

Resolution Enhancement Factor	***1.67 by 1.67***	***3.10 by 2.53***	***5.22 by 4.34***	***8.89 by 7.25***	***15.45 by 12.64***
Initial Resolution	270×243	147×162	90×99	54×60	33×36
Final Resolution	450×405	455×410	470×430	480×435	510×455
SSIM	0.963	0.926	0.894	0.848	0.781
PSNR	34.343	28.955	26.958	24.099	20.866
SNR	27.389	22	20.003	17.144	13.911
MSE	23.921	42.709	91.017	151.016	205.706

Table 5: Comparison of Mean SSIM and Mean PSNR with Previous work

Method	***Mean SSIM***	***Mean PSNR(db)***
[7]	0.8692	28.5
[12]	0.8444	28.5072
[3]	0.7067	22.3266
Proposed	0.925	30.89

VI. CONCLUSIONS AND FUTURE SCOPE

In this work, we have developed improved face resolution enhancement technique based on homogeneity of pixels in 3×3 LR patches, scanned over whole input facial image. The homogeneity kernel matrix is computed which is used to generate resultant HR image. 3×3 LR patches are slid to the nucleus of 5×5 HR patches and remaining pixels in HR patch are computed by application of homogeneity kernel matrix. We have tested the proposed technique for different resolution enhancement factors at minimum (1.67 by 1.67) to maximum (15.45 by 12.64) and low resolution facial images with minimum size 33×36. The accuracy of the proposed technique has been measured in terms of quality metrics *SSIM*, *PSNR*, *SNR* and *MSE*. *SSIM* being most representative measure to reflect the quality of resultant HR image with respect to original HR image, we have been able to achieve *SSIM* value 0.963 corresponding to lowest resolution enhancement factor 1.67*by*1.67 which falls down to 0.781 for highest resolution enhancement factor (15.45 by 12.64) and lowest size of LR facial image (33×36). From the results, we conclude that the proposed system is very effective in generating good quality HR facial image from single input LR facial image.

In near future, we intend to do improvements in the proposed technique by introducing adaptability in selection of LR patch size and number of iterations so as to achieve desired resolution enhancement factor. The proposed technique can be integrated with the video based search space reduction techniques such as frame difference, motion segmentation in processing surveillance videos. Further we intend to make the improvements in time efficiency of the proposed technique by using GPU acceleration.

VII. REFERENCES

[1] S. Baker and T. Kanade. Limits on super-resolution and how to break them. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(9):1167–1183, 2002.

[2] H. Chang, D.-Y. Yeung, and Y. Xiong. Super-resolution through neighbor embedding. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, volume 1, pages I–I. IEEE, 2004.

[3] L. Chen, R. Hu, C. Liang, Q. Li, and Z. Han. A novel face super resolution approach for noisy images using contour feature and standard deviation prior. *Multimedia Tools and Applications*, 76(2):2467–2493, 2017.

[4] M. Gong and K. Wang. Single face hallucination via local neighbor patches. *AEU-International Journal of Electronics and Communications*, 74:88–93, 2017.

[5] K. Huang, R. Hu, J. Jiang, Z. Han, and F. Wang. Face image super-resolution through improved neighbor embedding. In *International Conference on Multimedia Modeling*, pages 409–420. Springer, 2016.

[6] K. Huang, R. Hu, J. Jiang, Z. Han, and F. Wang. Hrm graph constrained dictionary learning for face image super-resolution. *Multimedia Tools and Applications*, 76(2):3139–3162, 2017.

[7] Z. Hui and K.-M. Lam. Eigentransformation-based face super-resolution in the wavelet domain. *Pattern Recognition Letters*, 33(6):718–727, 2012.

[8] Z. Hui, W. Liu, and K.-M. Lam. A novel correspondence-based face-hallucination method. *Image and Vision Computing*, 2017.

[9] J. Jiang, C. Chen, K. Huang, Z. Cai, and R. Hu. Noise robust position-patch based face super-resolution via tikhonov regularized neighbor representation. *Information Sciences*, 367:354–372, 2016.

[10] J. Jiang, C. Chen, J. Ma, Z. Wang, Z. Wang, and R. Hu. Srlsp: A face image super-resolution algorithm using smooth regression with local structure prior. *IEEE Transactions on Multimedia*, 19(1):27–40, 2017.

[11] J. Jiang, R. Hu, Z. Wang, Z. Han, and J. Ma. Facial image hallucination through coupled-layer neighbor embedding. *IEEE Transactions on Circuits and Systems for Video Technology*, 26(9):1674–1684, 2016.

[12] H. Li, L. Xu, and G. Liu. Face hallucination via similarity constraints. *IEEE signal processing letters*, 20(1):19–22, 2013.

[13] S. Liu and M.-H. Yang. Compressed face hallucination. In *Image Processing (ICIP), 2014 IEEE International Conference on*, pages 4032–4036. IEEE, 2014.

[14] J.-S. Park and S.-W. Lee. Resolution enhancement of facial image using an error back-projection of example-based learning. In *Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on*, pages 831–836. IEEE, 2004.

[15] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss. The feret evaluation methodology for face-recognition algorithms. *IEEE Transactions on pattern analysis and machine intelligence*, 22(10):1090–1104, 2000.

[16] P. J. Phillips, H. Wechsler, J. Huang, and P. J. Rauss. The feret database and evaluation procedure for face-recognition algorithms. *Image and vision computing*, 16(5):295–306, 1998.

[17] V. A. Rahiman and S. N. George. Single image super resolution using neighbor embedding and statistical prediction model. *Computers & Electrical Engineering*, 2017.

[18] C.-T. Tu, M.-C. Ho, and M.-Y. Lin. A new approach for face hallucination based on a two-dimensional direct combined model. *Pattern Recognition*, 62:1–20, 2017.

[19] J. Yang, J. Wright, T. S. Huang, and Y. Ma. Image super-resolution via sparse representation. *IEEE transactions on image processing*, 19(11):2861–2873, 2010.

[20] F. Yeganli, M. Nazzal, M. Unal, and H. Ozkaramanli. Image super-resolution via sparse representation over multiple learned dictionaries based on edge sharpness. *Signal, Image and Video Processing*, 10(3):535–542, 2016.

[21] Q. Yin, Z. Cao, and E. Zhou. Face hallucination using convolutional neural networks, Aug. 2 2016. US Patent 9,405,960.

Biographies

Mr. Vikram Mutneja is Ph.D. research scholar in I.K. Gujral Punjab Technical University, Kapurthala (Punajb, India) in the discipline Electronics Engineering. He received his Bachelor's degree (B.Tech.) from Guru Nanak Dev University, Amritsar, Punjab (India) with specialization in Electronics & Communication Engineering in year 1998, Master's degree (M.Tech.) from Giani Zail Singh College of Engineering & Technology, Bathinda (Punjab) (India) under Punjab Technical University (PTU, Jalandhar) with first division in Electronics & Communication Engineering in year 2008. He is a Life Member of ISTE. He has industrial experience of around 5 years (1998-2003) in which he worked in diverse fields such as Hardware and Networking Technical Support, Software Development for Web and Embedded Systems. During his teaching experience of around 13 years (2003-2016) he worked mainly in the areas of embedded systems, VLSI and Digital Signal Processing. He is working presently as Assistant Professor in department of Electronics & Communication Engineering in Shaheed Bhagat Singh State Technical Campus, Ferozepur, Punjab (India). His current field of research includes areas of image processing for detection and extraction of facial features from low resolution surveillance videos, and acceleration using GPU computing.

Dr. Satvir Singh is Associate Professor in I.K. Gujral Punjab Technical University, Kapurthala (Punajb, India). He received his Bachelor's degree (B.Tech.) from Dr. B. R. Ambedkar National Institute of Technology, Jalandhar, Punjab (India) with specialization in Electronics & Communication Engineering in year 1998, Master's degree (M.E.) from Delhi Technological University (Formerly, Delhi College of Engineering), Delhi (India) with distinction in Electronics & Communication Engineering in year 2000 and Doctoral degree (Ph.D.) from Maharshi Dayanand University, Rohtak, Haryana (India) in year 2011. He is a Life Member of ISTE and Corporate Member of IETE. During his 15 years of teaching experience he served as Assistant Professor and Head, Department of Electronics & Communication Engineering at BRCM College of Engineering & Technology, Bahal, (Bhiwani) Haryana, India and as Professor & Head, Department of Electronics & Communication Engineering at Shaheed Bhagat Singh State Technical Campus (Formerly, SBS College of Engineering & Technology), Ferozepur Punjab, India. His fields of special interest include Evolutionary Algorithms, High Performance Computing, Type-1 & Type-2 Fuzzy Logic Systems, Wireless Sensor Networks and Artificial Neural Networks for solving engineering problems.

