Apparent Resolution Enhancement: Structural Similarity Perspective

Rehanullah Khan
IT Department
Qassim University
Saudi Arabia

ABSTRACT

This paper presents a detailed investigation and comparative analysis of the two well-known classes of Super Resolution (SR) i.e. Reconstruction Based Super Resolution (RBSR) and the Example Based Super Resolution (EBSR), taking into account the seven variables: Resolution Factor, Aliasing, Stills, Motion, Compression, Noise and Detection scenario. The EBSR uses high and low frequency relationship and the RBSR is based on the frames sequence information. The EBSR and RBSR are tested on number of images to investigate which of the two classes of SR algorithms are best suited for preserving structural similarity to the original image and for visual analysis of the SR image. Experimental results show that over-all SSIM index for the EBSR is higher than RBSR and thus preserves the image quality better compared to the RBSR. In an evaluation of EBSR and RBSR for feature based detection scenario, it is observed that face detection in the EBSR resultant image has the same performance compared to that of the Viola-Jones approach [21] without SR; therefore, we do not gain any improvement in EBSR. In the case of RBSR, due to the registration errors, the over-all face detection performance after SR by the Viol-Jones algorithm is reduced by 3%. As such, it is an extension of the author's conference work [12].

Keywords
EBSR, RBSR, SR, RSR, SSIM Index,

1. INTRODUCTION

Image processing applications regarding different fields such as medicine, defense etc. mostly requires high resolution images/videos for interpretation purposes. Being an important aspect of image processing, high resolution image reproduction has been a keen area for researchers. Super Resolution approach is one among their efforts in this regard. For obtaining super resolved image, generally three techniques are used: Sharpening, sequence aggregation and frequency correspondences. Sharpening technique is used for amplifying the high frequencies in an image by changing the frequency spectrum of the image. One problem with sharpening technique is that noise is amplified too, with the high frequencies, if it is not the case; it results in an increased resolution of an image. Sequence aggregation technique is the representative of Reconstruction Based Super Resolution approaches, which produces a high resolution image from a sequence of Low Resolution (LR) images. In frequency correspondence approach, high resolution images are obtained through learning the correspondences from low resolution images. Such approaches are termed as Example Based Super-Resolution (EBSR) approaches [7]. In general, the SR methods can be classified into two classes; Reconstruction based methods and Example based methods [1]. The Reconstruction Based Super Resolution methods reconstruct the LR images to obtain a high resolution image. A number of SR algorithms are based on the RBSR approach. Frequency domain [20], Maximum Likelihood, Non-uniform interpolation (consisting of three stages including registration or motion estimation, interpolation and noise reduction), Regularized Super Resolution and Projection onto Convex Sets [3, 8, 9] are considered the most valuable algorithms in this regards. Recently, the EBSR methods for Super Resolution got the attention of the researchers by producing remarkable results, for example, starting with the work of [7,16]. The EBSR methods rely on the training data for the prediction of details in LR images, for the SR output. Markov network and one pass algorithm are the best representatives of this type of SR methods. EBSR methods are very effective in specific domain e.g. faces, flowers etc. [7].

This article aims at a structural comparison between generated SR images using the two classes of SR (namely Reconstruction Based Super Resolution and Example Based Super Resolution). As such, it is an extension of the author's own conference work [12]. The comparison and evaluation of both approaches is based on two factors: SSIM index and the visual appearance, taking into consideration the seven variables including: Resolution factor, aliasing, stills, motion, compression, noise and detection scenario. Based on the experimental results, it is concluded that compared to Reconstruction Based Super Resolution, Example Based Super Resolution approaches are best suited for SR in terms of preserving the structure of the input image. Such structural similarity plays a very important role in different applications of computer vision such as face and human body recognition and other textured based recognition approaches. In this work, we used the Robust Super
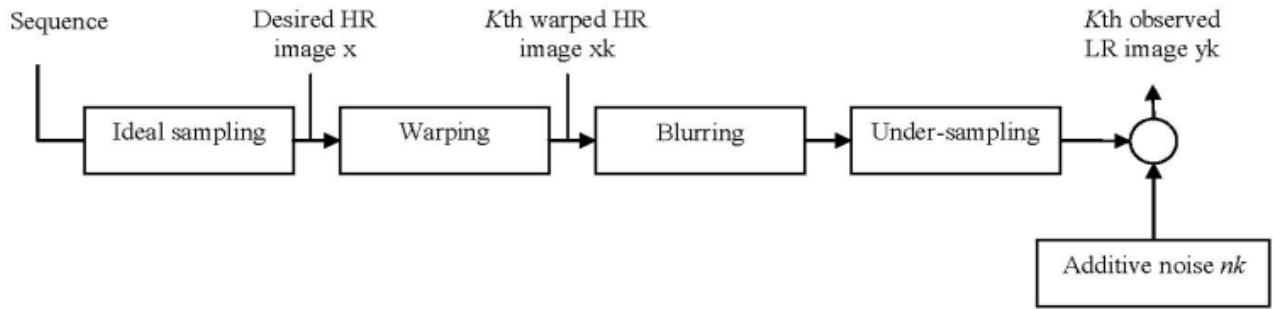
Figure 1: Block diagram of obtaining LR observed image from the HR image: sampling phase normally results the lost of high frequency components of images. Image quality can be further descreased due to rotation of imaging device and camera bluring. Finally, the spatial sampling noise degrades the resolution of the HR scene/image. In process of going back to the HR image ffrom the LR image, all the steps need to be undone that resulted in the LR image.

Resolution [24] for the reconstruction stage of SR and the approach proposed by Keran et al. [11] for the Motion estimation [2] in the RBSR. For EBSR, Freeman et al. [7] approach is adopted.

RBSR are based on the assumption that high resolution signal can be well predicted from the low resolution signal (image) [4,5,6]. Work done by Tsai et al. [20], known as frequency domain method and is considered the pioneering approach in this regards. Iterative backward projection algorithm by Irani et al. [9] is another RBSR approach for obtaining HR images. [5] Represents a unifying approach based on the matrix vector notation. The super resolution images can be considered as the maximum A-posterior solution to stochastic optimization in regularized SR reconstruction approach and the prior knowledge of the HR images and noise for the reduction of inconsistent measurements effects [3,9,10,16].

Projection onto Convex-sets algorithm uses the prior knowledge with constrains, for the reduction of noise in LR images [9]. Generally, high frequency components are suppressed by the enforced smoothness constraint which results in a blurred image. Simultaneous SR in time and space and the Robustness issues related to SR are discussed in greater details in [6] and [19] respectively. Motion segmentation being a crucial factor in SR for the moving objects, especially in the presence of noise, is nicely handled in the said approaches. Zomet et al. [24] proposed an iterative algorithm, which is based on the median estimator. The said algorithm results in super resolved background only, in the image. Freeman's [7] algorithm which is used in this work for the EBSR uses the generative model learned from HR images. A high frequency detailed image is obtained by using such model. Due to this learning characteristic, EBSR approaches are applicable only to specific domains [1], especially in dynamic scenes, EBSR approaches failed to produce sufficient results and require huge amount of training data for handling real world dynamic video sequences. The authors in [17,18,21,22,23,25] discuss different approaches for RBSR.

Experi`Mental Analysis
The comparisons and evaluations of both SR methods i.e. EBSR and RBSR are carried out in different scenarios that are favorable for both the methods and with seven variables including: Resolution factor, aliasing, stills, motion, compression, noise and detection scenario. As EBSR approaches rely on training, thus, the training set used in experimentation contains the images close in complexity to that of testing set. The selected dataset for the evaluation has the images with motion in consecutive frames that is the pre-requisite of RBSR.

The comparison of both approaches is based on the SSIM index value [22], ranging from 0 to 1. The higher value of SSIM Index indicates the higher structural similarity between the original image and the obtained HR images. SSIM index based comparison is a better choice for this comparative study, as it has shown better performance over its other competitors like MSE and PSNR measures [22]. Figure 2 shows an example LR image (2(a)) and the resultant EBSR SR image (2(b)) and RBSR SR image (2(c)). Figure 2(b) shows the results of EBSR having an SSIM index of 0.92, while Figure 2(c) shows the super resolved image through RBSR/RSR approach, having an SSIM index of 0.83.

(a) (b) (c)
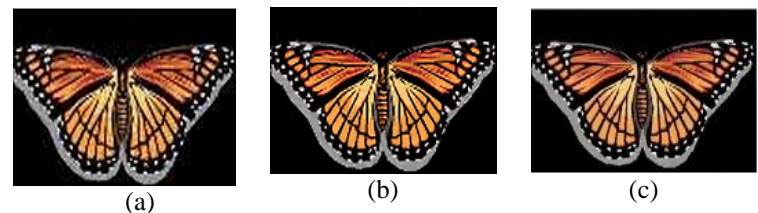
Figure 2: Comparison of EBSR and RBSR approaches in terms of SSIM index: (a) Original image (b) Resultant SR image of EBSR with SSIM index of 0.92. (c) Resultant SR image of Robust Super Resolution (RSR) (RBSR approach) with SSIM index of 0.83.

A zoomed version of a portion of these images, showing the high frequency components details is shown in Figure 3. The results of visual comparison (shown in Figure 3) and the high SSIM score (Figure 2) indicate that EBSR outperforms the RBSR.

(a) **(b)** **(c)**
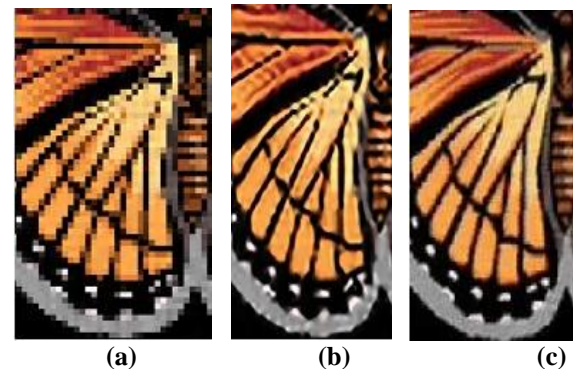
Figure 3: Visual comparision of Example Based Super Resolution and Reconstruction Based Super Resolution approches: (a) A zoomed portion of the original image in Figure 2(a). (b) A zoomed portion of the resultant image of Example-based Super-Resolution in Figure 2(b). (c) A zoomed portion of the resultant image of Robust Super Resolution (RBSR approach) in Figure 2(c).

Other parameters regarding the SR, along with the experimental results of EBSR and RBSR comparisons regarding these parameters, are discussed as follows. These parameters include: Resolution factor, aliasing, stills, motion, compression, noise and detection.

2.1 Resolution Factor

In our experimentally setup, the RBSR provides better performance (in terms of SSIM index score) up to the twice of the original image resolution, however, the EBSR produces better performance even up to four times of the original resolution as can be judged from the Figure 2 and Figure 3.

2.2 Aliasing

If the low-resolution images i.e. the input images have under sampling and/or aliasing issues, these are tackled during the 1st step (registration) of SR approaches. RBSR approaches nicely tackle these aliasing artifacts, however, the EBSR techniques enhances these artifacts along with high frequency texture, as illustrated in Figures 4 and 5. In the experimentation setup, the SSIM for EBSR is higher than the RBSR approach. Due to the fact that such aliasing artifacts are visually deceptive to the original image, the SSIM score is also deceptive in this case. The SSIM index scores in this case are 0.96 and 0.71 for EBSR and RBSR approaches respectively (Figure 4(b) and (c)).

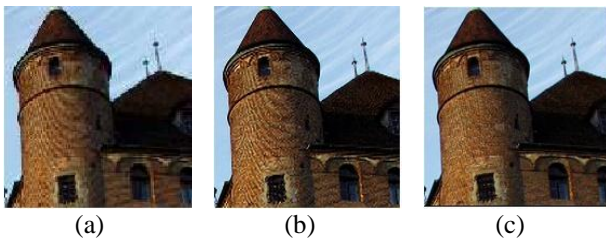
(a)   (b)   (c)

Figure 4: Comparison of EBSR and RBSR approaches in the presence of aliasing artifacts on a building images: (a) Original image. (b) Resultant SR image of EBSR with SSIM index of 0.96 (c) Resultant SR image (RSR) (RBSR approach) with SSIM index of 0.71.
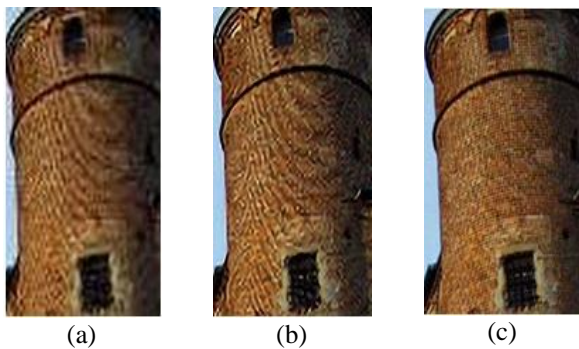

(a)   (b)   (c)

Figure 5: Comparision of EBSR and RBSR with aliasing artifacts for building images: (a) A zoomed portion of the original image in Figure 4(a). (b) A zoomed portion of the resultant SR image of Example-based Super-Resolution in Figure 4(b). (c) A zoomed portion of the resultant SR image of Reconstruction Based Super Resolution in Figure 4(c)

2.3 Stills

RBSR based algorithms combine the details present in the image sequence. If all the images of a sequence do not have any change or motion, so in this case the result of RBSR approach is just the interpolation and no enhanced details. However, on the other side, the EBSR approaches can

produce better SR taking benefits of the training samples; see for example, Figure 6 and Figure 7. Figure 6(a) shows the original image. EBSR resultant super resolved image is shown in 6(b) and its SSIM score is 0.99. RBSR resultant SR image in this particular case is shown in Figure 6 (c) with 0.98 SSIM index score. The higher value of SSIM index of EBSR as compared to RBSR approach verifies its expediency over RBSR. To highlight the high frequency components, portions of images in Figure 6 are zoomed-in in Figure 7.

(a)   (b)   (c)

Figure 6: Comparison of Example Based Super Resolution and Reconstruction Based Super Resolution in the presence of no motion in consecutive frames. (a) Original image (b) Result of an Example-based Super-Resolution with SSIM index of 0.99. (c) Result of Robust Super-Resolution (RSR) with SSIM index of 0.98.
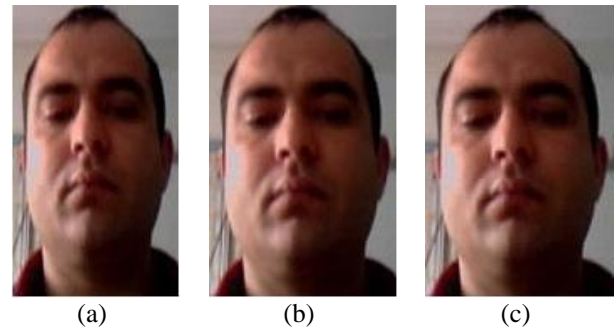

(a)   (b)   (c)

Figure 7: Comparison of Example Based Super Resolution and Reconstruction Based Super Resolution in the presence of no motion. (a) A zoomed portion of original image in Figure 6(a). (b) A zoomed portion of the result of Example-based Super-Resolution in Figure 6(b). (c) A zoomed portion of the result of Reconstruction based Super Resolution in Figure 6(c).

2.4 Motion

In images having the presence of large motions (Figure 8), the RBSR approaches will produce ghosting effects, as shown in Figure 8(c). In such cases, the EBSR approaches produce better performance as compared to RBSR approaches, as clear from the SSIM scores for the two approaches. In this particular experiment, the SSIM index score for EBSR is 0.78, which is much better than RBSR score of 0.60. The visual comparison of resultant SR images for both the approaches is shown in Figure 8.

(a) (b) (c)

Figure 8: Comparison of Example Based Super Resolution and Reconstruction Based Super Resolution approches in presence of large motion between the frames: (a) Original image. (b) Resultant SR image of an Example Based Super Resolution with SSIM index of 0.78. (c) Resultant image of Reconstruction Based Super Resolution with SSIM index of 0.60. (Image Source: Prof. Robert Sablatnig, CVL, Vienna, Austria)

2.5 Compression

Compression badly affects both RBSR and EBSR approaches. In both cases the compression artifacts are amplified in the resultant SR images. In case of compression, the SSIM score of EBSR (0.99) is higher than the SSIM score 0.93 of RBSR as shown in Figure 9. The high frequency details in Figure 9 are zoomed in Figure 10.

(a) (b) (c)

Figure 9: Comparition and evaluation results of Example Based Super Resolution and Reconstruction Based Super Resolution on images with compression artifacts. (a) Original image (b) Resultant SR image of an EBSR with SSIM index of 0.99. (c) Resultant SR image of Robust Super-Resolution (RSR) (a RBSR approch) with SSIM index of 0.93.

(a) (b) (c)

Figure 10: Visual Comparison of both Example Based Super Resolution and Reconstruction Based Super Resolution in the presence of strong compression artifacts. (a) A zoomed portion of the original image of Figure 9(a). (b) A zoomed portion of the resultant SR image of Example Based Super Resolution of Figure 9(b). (c) A zoomed portion of the resultant SR image of Reconstruction Based Super Resolution of Figure 9(c).

2.6 Noise

In presence of noise, the RBSR approaches algorithms perform better than EBSR approaches. The reason of the failure of EBSR approaches is that these approaches learn the correspondence inferences in the spatial domain. Therefore, do not provide robustness against noise and also amplify the noise signals in SR images.

2.7 Detection Scenario

For the comparison and the effects of the RBSR and EBSR algorithms on the detection scenarios, we opt for the rectangular features based face detection. For a face detection system, we use the well know approach of Viola-Jones [21]. In a comprehensive evaluation of EBSR and RBSR on face images, it is observed that face detection in the EBSR resultant image has the same performance of that of the Viola-Jones approach; therefore, we do not gain any improvement in EBSR. In the case of RBSR, due to the registration errors, the over-all face detection performance after SR by the Viol-Jones algorithm was reduced by 3%.

3. CONCLUSION

Comparison and evaluation of EBSR and RBSR is carried out in different scenarios that are favorable for both the methods and with seven variables including: Resolution factor, aliasing, stills, motion, compression, noise and detection scenario. Given the high over-all SSIM index, it can be concluded that compared to RBSR, EBSR performs better in preserving the structure of the original image. In terms of visual appearance, EBSR is preferred except when there are aliasing artifacts or when the images are strongly compressed. In an evaluation of EBSR and RBSR in feature based detection scenario, it is observed that face detection in the EBSR resultant image has the same performance compared to that of the Viola-Jones approach; therefore, we do not gain any improvement in EBSR.

4. REFERENCES

[1] S. Baker and T. Kanade 2002, "Limits on super-resolution and how to break them," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 9, pp. 1167 – 1183.

[2] J. R. Bergen, P. Anandan, K. J. Hanna, and R. Hingorani 1992, "Hierarchical model-based motion estimation," in ECCV '92: Proceedings of the Second European Conference on Computer Vision. London, UK: Springer-Verlag, pp. 237–252.

[3] D. Capel and A. Zisserman 2000, "Super-resolution enhancement of text image sequences," in in Proc. International Conference on Pattern Recognition. IEEE Computer Society Press, pp. 600–605.

[4] M. chao Chiang and T. E. Boult 1996, "Efficient image warping and super resolution," in In IEEE Workshop on Applications of Computer Vision (WACV96. IEEE Computer Society, pp. 56–61. M.

[5] Elad and A. Feuer 1997, "Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images," Image Processing, IEEE Transactions on, vol. 6, no. 12, pp. 1646–1658.

[6] P. E. Eren, M. I. Sezan, and A. M. Tekalp 1997, "Robust, object-based high-resolution image reconstruction from

low-resolution video," IEEE Transactions on Image Processing, vol. 6, no. 10, pp. 1446–1451.

[7] W. T. Freeman, T. R. Jones, and E. C. Pasztor 2002, "Example-based superresolution," IEEE Computer Graphics and Applications, vol. 22, no. 2, pp. 56–65.

[8] R. Hardie, K. Barnard, and E. Armstrong 1997, "Joint map registration and high resolution image estimation using a sequence of undersampled images," vol. 6, no. 12, pp. 1621–1633.

[9] M. Irani and S. Peleg 1991, "Improving resolution by image registration," CVGIP: Graph. Models Image Process., vol. 53, no. 3, pp. 231–239.

[10] Z. Jiang, T.-T. Wong, and H. Bao 2003, "Practical super-resolution from dynamic video sequences," in CVPR , pp. 549–554.

[11] D. Keren, S. Peleg, and R. Brada 1988, "Image sequence enhancement using sub-pixel displacements," pp. 742–746.

[12] R. Khan, R. Sablatnig, A. Bais, and Y. M. Khawaja 2011, "Comparison of reconstruction and example-based super-resolution," in IEEE ICET, pp. 1–6.

[13] F. Lin, S. Denman, V. Chandran, and S. Sridharan 2007, "Automatic tracking, super-resolution and recognition of human faces from surveillance video," in MVA, pp. 37–40.

[14] F. C. Lin, S. Denman, V. Chandran, and S. Sridharan 2007, "Automatic tracking, super-resolution and recognition of human faces from surveillance video," in IAPR Conference on Machine Vision Applications. The International Association for Pattern Recognition, pp. 37–40.

[15] F. Liu, J. Wang, S. Zhu, M. Gleicher, and Y. Gong 2008, "Noisy video superresolution," in MM '08: Proceeding of the 16th ACM international conference on Multimedia. New York, NY, USA: ACM, pp. 713–716.

[16] A. J. Patti, M. I. Sezan, and A. M. Tekalp 1997, "Superresolution video reconstruction with arbitrary sampling lattices and nonzero aperture time," IEEE Transactions on Image Processing, vol. 6, pp. 1064–1076.

[17] T. Q. Pham, L. J. van Vliet, and K. Schutte 2006, "Resolution enhancement of low-quality videos using a high-resolution frame," J. G. Apostolopoulos and A. Said, Eds., vol. 6077, no. 1. SPIE, pp. 607–708.

[18] R. R. Schultz and R. L. Stevenson 1996, "Extraction of high-resolution frames from video sequences," IEEE Transactions on Image Processing, vol. 5, pp. 996–1011,.

[19] E. Shechtman, Y. Caspi, and M. Irani 2002, "Increasing space-time resolution in video," in ECCV '02: Proceedings of the 7th European Conference on Computer Vision-Part I. London, UK: Springer-Verlag, pp. 753–768.

[20] R. Y. Tsai and T. S. Huang 1984, "Multi-frame image restoration and registration," in Advances in Computer Vision and Image Processing, Greenwich, CT, pp. 317–339.

[21] P. Viola and M. J. Jones 2004, "Robust real-time face detection," IJCV, vol. 57, no. 2, pp. 137–154.

[22] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli 2004, "Image quality assessment: From error visibility to structural similarity," IEEE Transactions on Image Processing, vol. 13, pp. 600–612.

[23] G. Wolberg 1990, Digital Image Warping. IEEE Computer Society Press, Los Alamitos, CA.

[24] A. Zomet, A. Rav-acha, and S. Peleg 2001, "Robust super resolution," In Proc. of the IEEE Workshop on Applications of Computer Vision, pp. 645–650.

[25] H. Chang, D. Y. Yeung, and Y. Xiong 2004, Super-resolution through neighbor embedding. In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, volume 1, pages 275–282.