Authenticating super-resolved image and Enhancing its PSNR using watermark

Mehul S. Raval*, Vaibhav B. Joshi*, Dhruv Gupta* and Shubhalaxmi J. Kher†
*Institute of Engineering and Technology
Ahmedabad University
Ahmedabad, INDIA
Email: mehul.raval@ahduni.edu.in
†Department of Electrical Engineering
Arkansas State University
Jonesboro, Arkansas, USA
Email: skher@astate.edu

***Abstract*—Most watermarking methods reduce peak signal-to-noise ratio (PSNR) of a host signal like images. However in this paper, we propose a novel watermarking scheme which increases PSNR of the super-resolved image while ensuring its authenticity. The idea is to capture the information missed by a super-resolution algorithm and fuse it with the super-resolved image. This information when fused increases the PSNR and can also be used as a watermark. The singular values (SV's) derived from edges of the low resolution(LR) image forms the image dependent watermark. Next, compressive sensing (CS) framework is used to generate linear measurements of the watermark. These are then embedded in the super-resolved image using arithmetic coding. Watermark is extracted at the receiver using L_1 minimization to authenticate the super-resolved image. Only after positive authentication, the edge information is regenerated using the SV's (watermark). These edges are added to the super-resolved image, increasing its PSNR. To best of our knowledge, it is the first algorithm that authenticates super-resolved image while increasing signal fidelity.**

I. Introduction

Super-resolution (SR) [1] is an algorithmic approach to construct high resolution (HR) images from single or multiple low resolution images. The performance of inverse ill-posed problem like super-resolution is usually evaluated with a measure like peak signal-to-noise ratio (PSNR). It is a ratio between the maximum possible signal power to the noise power as described in Eq. 2. For the image with dimensions $M \times N$, mean square error (MSE) is given by,

$$MSE = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [I(i,j) - K(i,j)]^2 \qquad (1)$$

and peak signal to noise ratio (PSNR) is,

$$PSNR_{dB} = 10log_{10}\Big(\frac{MAX_I}{MSE}\Big) \qquad (2)$$

where, I represents the original image, K represents the processed image and MAX_I represents the maximum value of a single image pixels.

MSE [2] between the super-resolved and ground truth image (if available) is used to compute PSNR. For a gray scale image with 8 bits per pixel representation, the maximum possible pixel value is 255. The higher PSNR indicates better signal fidelity. The state of the art SR algorithm improves the signal power [3]. The PSNR is an accepted norm of figure of merit for signal fidelity measurements in the research community [4]. Super-resolution find applications in domains of medical imaging, remote sensing [5], biometrics, image fusion, image in-painting, and many more.

In medical applications, many imaging modalities like; fluoroscopy, magnetic resonance imaging (MRI), positron emission tomography(PET), computed axial tomography (CAT) exist. While some of these techniques reveal anatomical information, others are used to measure the functional activities of the body. Each of these imaging systems has a typical resolution governed by the physics of the instrumentation, signal-to-noise ratio (SNR) and system timing. A most common goal is to increase the image resolution without changing the hardware and obtain the true 3D imaging. Therefore, the increased resolution leads to better diagnosis as clinicians can detect the weak bio-markers during the early stages of the disease.

The super-resolved medical image plays a very important role in tele-medicine; which is about providing health care at distance. It has been developed with an aim to overcome the geographical barriers in the health care and provides an effective mechanism of communication between the patients and the clinicians. The most common medical data includes images, health informatics, and electronic patient records (EPR). However, there is a potential risk that data may be compromised during storage and transmission over insecure networks. Given that health record of an individual is an extremely sensitive information, it should be secured [6] [7] [8]. One should check whether the data has been modified during its passage through a communication channel or not. This important attribute of information security is known as data authentication. Hence, for perfect remote diagnosis, authenticity of the data like; super-resolved medical images should be preserved over a channel.

Watermarking [9], [10] is one of the mechanism which provides an excellent authentication mechanism. Specifically, fragile [11] or semi-fragile watermarks are mainly applied to content authentication and integrity attestation as they are

978-1-4799-7560-0/15 $31.00 © 2015 IEEE
DOI 10.1109/SSCI.2015.77

sensitive to modifications. Improvement in the robustness due to watermark embedding, directly into a super resolved image has been reported in [12] [12]. However, the method has several shortcomings and one of the shortcoming is degradation of PSNR of the host image due to watermark insertion. Researchers also proposed the copy right protection scheme for the super-resolution [3]. The goal was to minimize PSNR degradation in super-resolved image due to watermarking. The method employed fuzzy-neural network to provide copy right protection for SR images with minimal degradation in PSNR.

Any watermarking scheme can be seen as an interplay of three important attributes namely; capacity, fidelity, and security. The change in one of the attribute affects the other e.g. increase in watermark capacity degrades the host signal fidelity. The distortion of host signal fidelity is often measured by the PSNR of the watermarked image. Even though invisible watermark does not affect perceptual quality but the host PSNR degrades due to embedding operations as most watermarking schemes tend to fill the "holes" into the host signal.

The researchers are actively involved in solving two problems namely; 1. improving resolution of signals like; medical images 2. providing security to a data over the communication channel. However, there is very little overlap between these communities. The proposed work is an attempt to bridge this gap. We observe that medical imaging needs high resolution, increased PSNR, and security over an insecure channel.

Super-resolution is an inverse ill-posed problem, therefore it cannot perfectly reinstate HR image from the LR image. If one can capture the difference between the HR image and its SR version (reconstructed from LR); it can be advantageous as follows: 1. PSNR is improved by adding difference(Diff) to the SR version i.e. $enhanced\ SR(ESR) = Diff + SR$; 2. $Diff$ can also be used as a watermark, in an authentication framework.

Therefore, in this work new algorithm is proposed which uses watermark derived from $Diff$ to authenticate the super-resolved image. After successful authentication, difference is regenerated from the watermark and added to the super-resolved image, increasing its PSNR. Thus, a watermarking algorithm authenticates the super-resolved image and at the same time increases its PSNR. The idea is to find edges/$Diff$ which cannot be constructed by the SR algorithm. Also, these edges are dependent on the SR image, hence they are used as a watermark. The watermark embedding rate is minimized by computing singular value decomposition (SVD) of these edges, taking their linear measurements using compressive sensing and adding them to the super-resolved image. At the receiver, these SV's are estimated using CS frame work, then they authenticate the SR image and finally difference is regenerated. These edges or $Diff$ when added to the SR image improves the PSNR.

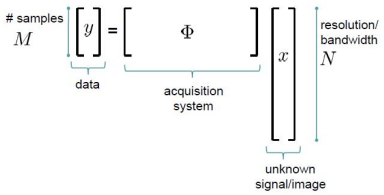
Our assessment results show that PSNR is increased when tested using number of super-resolution algorithms [13]. However, in this paper, results using robust super-resolution [14] are presented as it generated the best PSNR for SR images after difference addition. One must carefully note that proposed watermarking algorithm does not provide copyright protection to the super-resolved image but is meant for providing authentication. It has been designed to increase the PSNR of the super-resolved image.

A. Compressive sensing (CS)

Compressive sensing (CS) [15] is a framework of reconstructing the sparse signal from its linear measurements. If the signal is non-sparse, it can be sparsified by de-correlating it with the known basis. Conventionally, the linear measurement has lower dimensionality than sparse signal. However, in this work compressive sensing is used to recover signal corrupted with the noise. As discussed later in Section II, the linear measurements are rounded to the nearest integers, so that they can be embedded into the space created due to loss-less compression of the SR image.

B. Basic L_1 minimization

This is one of the most popular framework in sparse signal recovery. Its power stems from the fact that it converts the search into convex problem and provides the accurate recovery. Any under-determined system is given as $y = \Phi \times x$:



Now, for the given y an attempt to estimate the x is made by applying pseudo inverse of Φ to y. In another words, the basic noise-less recovery formulation is defined as follows:

$min_x \parallel x \parallel_{L1}$ subject to $\Phi x = y$

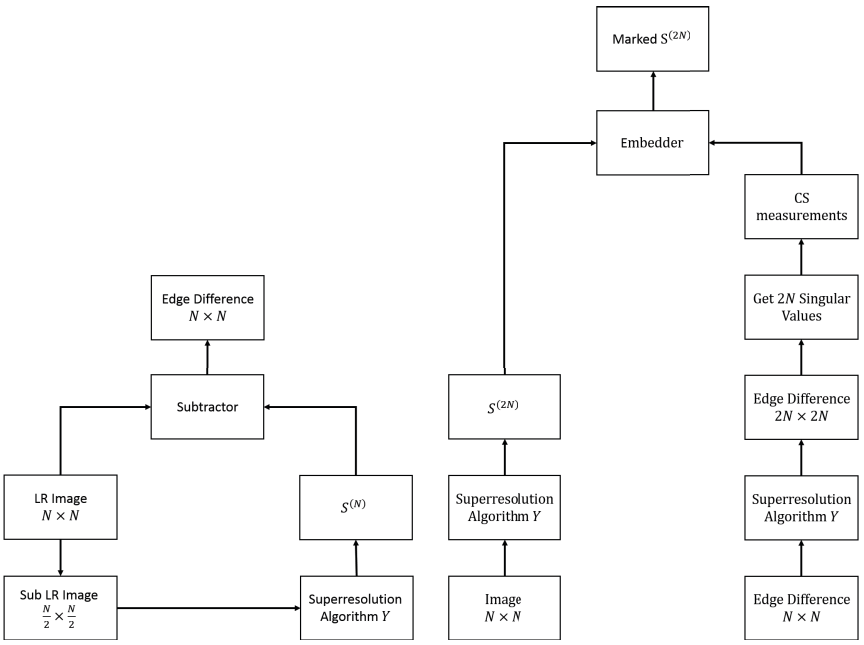
In L_1 minimization process, the final goal is to formalize the recovery condition under which the procedure for estimation of x is effective. This basic formulation is also formally defined as the basis pursuit (BP) and the term reveals that L_1 norm has a tendency to locate the sparse solution if it exists.

However, the CS framework in the present work has been used as follows. Like in a channel code, we generate the oversampled linear measurements of x, i.e. $y = \phi x$. The measurement matrix (ϕ) is square and full rank. These measurements are rounded and thus suffer from the rounding errors (e), corrupting them i.e. $y = \phi x + e$. Now given y and ϕ, the recovery algorithm find the signal x such that error $(y - \phi x)$ has minimum L_1 norm.

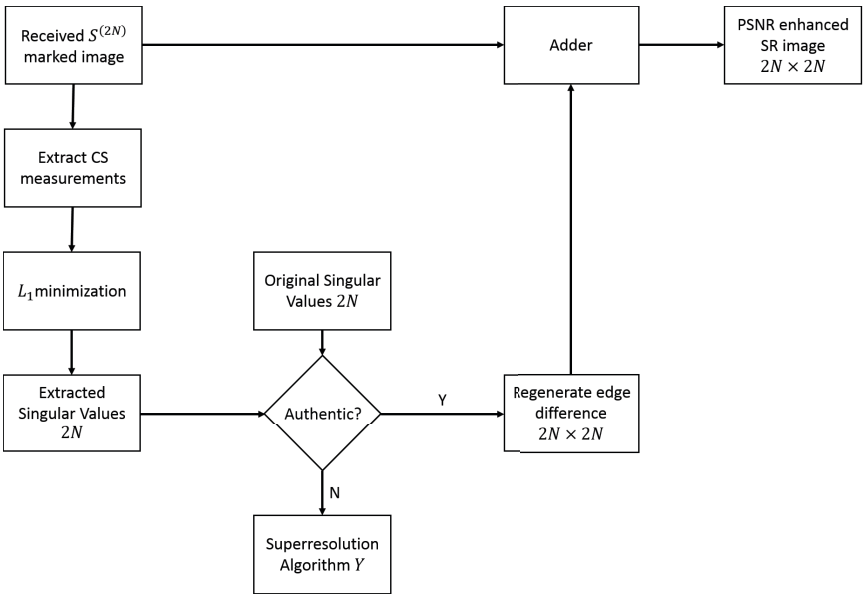
The rest of the paper is organized as follows. Section II covers the proposed watermarking algorithm featuring increase in PSNR of the super-resolved image. The experimental validation of the proposed hypothesis is given in Section III. Section IV presents conclusion.

II. WATERMARKING ALGORITHM WITH PSNR IMPROVEMENT

The algorithm is developed principally for the gray-scale images but can be extended for color images. The process of the proposed watermarking algorithm is exemplified in Fig. 1. Given the low resolution image ($N \times N$), the embedding procedure includes the following steps.



(a)



(b)

Fig. 1: Block diagram of the proposed algorithm for (a) watermark embedding (b) watermark extraction

A. Generating edge difference

The idea of this part is to locate the edge details in an LR image which cannot be produced with the super-resolution algorithm. Given the low resolution image L, procedure to generate edge difference includes the following steps.

1) The $N \times N$ image L is decimated by a factor of 2 to get sub LR image D of size $\frac{N}{2} \times \frac{N}{2}$.
2) Apply the super-resolution algorithm (Y) on the decimated image to generate the super-resolved $N \times N$ image $S^{(N)}$.
3) Generate the difference image $Diff = S^{(N)} - L$ at the spatial resolution $N \times N$.

The watermark derived from the difference image is embedded in the SR image $S^{(2N)}$ at $2N \times 2N$ spatial resolution. The

watermark embedding is discussed in the following subsection.

B. Watermark embedding in the $2N \times 2N$ SR image

The embedding process for generating the watermarked image has the following steps:

1) The images L and $Diff$ are super-resolved using super-resolution algorithm Y to get $S^{(2N)}$ and $Diff'$ of size $2N \times 2N$.
2) Next apply singular value decomposition (SVD) on $Diff'$ to get singular values $V \in \mathbb{R}^{2N}$ and corresponding Eigen vectors. The V's forms the watermark. The zeros are appended to V such that $V \in \mathbb{R}^{8N}$
3) Generate the over sampled linear CS measurements $y \in \mathbb{R}^{8N} (y = \Phi V)$ of the V's using the measurement matrix $\Phi \in \mathbb{R}^{8N \times 8N}$. The Φ is a full rank matrix.
4) The $S^{(2N)}$ is loss-lessly compressed using arithmetic coding.
5) The y's are rounded to the nearest integer and embedded into the space created due to loss-less compression of the $S^{(2N)}$. This will generate the final marked image.

Embedding of the oversampled y's are based on the context of channel coding [16].

C. Watermark extraction and generation of PSNR enhanced SR image

The watermark extraction and generation of the PSNR enhanced SR image include the following steps:

1) Extract the lossy CS measurements y' from the received $S^{(2N)}$ SR image. The image is then decompressed.
2) Apply the L_1 minimization framework and solve the following optimization problem to recover the singular values V'. The watermark extraction algorithm requires the measurement matrix Φ.

$$\min_{V'} \|y - \Phi V'\|_1 \tag{3}$$

The above optimization problem finds the vector V' such that error $y - \Phi V'$ has minimum L_1 norm. It extracts first $2N$ values of V'.
3) Apply SVD to the $Diff'$ image available at the receiver to get V and the corresponding Eigen vectors.
4) Compare V' and V using the normalized correlation (NC). The NC is defined as

$$NC(V, V') = \frac{V^T \times V'}{\| V \| \times \| V' \|} \tag{4}$$

Here, a high value of normalized correlation coefficient indicates that the SR image is authentic and low value of NC means that image is fake and should be rejected.
5) Combine V' and Eigen vectors obtained in step 3 to regenerate difference image $Diff''$.
6) Add the scaled $Diff''$ and the genuine SR image to generate the PSNR enhanced SR image at $2N \times 2N$ resolution. The scale factor is dependent on the gradient of the SR image and thus computed automatically for every SR image.

TABLE II: Hamming Distance between Original SR and edge enhanced SR images

Image	Hamming Distance between Original and SR images	Hamming Distance between Original and enhanced SR images
Teacher	6.4894×10^4	6.4396×10^4
Kid	6.3724×10^4	6.3302×10^4
Castle	6.5072×10^4	6.4922×10^4

The L_1 minimization framework solves the norm approximation problem using a primal-dual algorithm [16]. If the error $y - \Phi V$ is sparse i.e. the error has minimum L_1 norm and the decoder recovers the signal V from the oversampled y's.

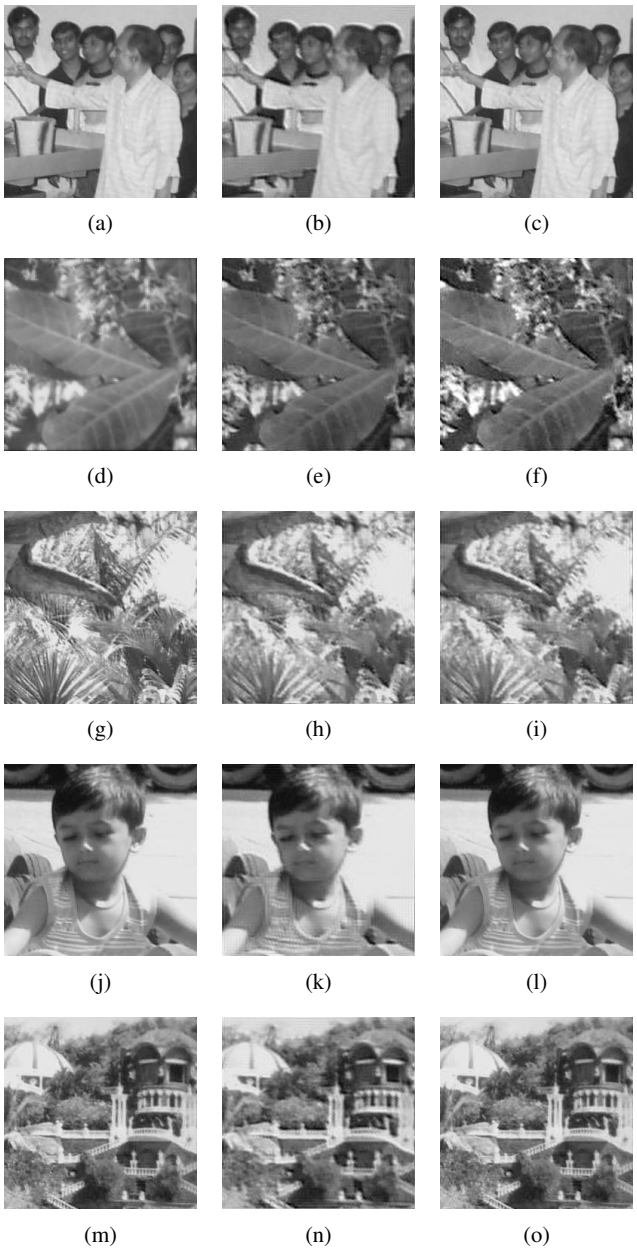
III. EXPERIMENTAL RESULTS

In the experiments, images with the size of 128×128 ($N \times N$) were used as the LR images. The grayscale images with the size of 256×256 ($2N \times 2N$) are used as the ground truth images. Robust super-resolution algorithm [14] is applied to D, difference image $Diff$ and the LR image L. SR algorithm takes the image at resolution X and outputs the image at resolution $2X$. The various dimensions are $V \in \mathbb{R}^{256}$, $y \in \mathbb{R}^{1024}$, $\Phi \in \mathbb{R}^{1024 \times 1024}$. The measurement matrix is generated from the standard normal distribution. The hiding rate of 0.015 bpp on an average is obtained during embedding the CS measurements y's in the $S^{(2N)}$. Some of the ground truth(GT), SR and enhanced SR (ESR) images are shown in Fig. 2. As one may note images with various textures has been used to test the proposed algorithm. Due to limitations of the printing resolution, these set of images may look similar to untrained eyes in the paper. Therefore, a zoomed portion of the images is shown in Fig. 4 to showcase improvement in the structural details and signal fidelity of the super-resolved image after the difference addition. Also, to quantify that ESR image is "closer" (in some distance sense) than SR image to the ground truth image, a hamming distance between ground truth images and ESR i.e D(ESR, GT) / SR images i.e D(SR, GT) is computed. TABLE II indicates the hamming distances for the three cases. It can be observed that for each image and also on the average, Hamming distance between GT and ESR images is smaller than the distance between GT and SR i.e D(ESR, GT) $<$ D(SR, GT). This establishes the fact that ESR has better signal fidelity than SR image

The authenticity of the SR image is computed by checking the normalized correlation (NC) between the extracted and original singular values. The NC is compared with a preset threshold τ. A value of $NC > \tau$ indicates that the SR image is authentic and than $Diff''$ is added to it. The choice of τ is crucial in deciding about the authenticity of the SR image. It can be varied for 0 to 1, which can result in every image passing the thresholding test (large number of false positives) to none of them clearing the test (large number of false negatives). The suitable τ can be decided based on threat perception, i.e. its value can be kept high, when threat perception for a system is high and vice-versa. In the above experimentation, very high threat perception is assumed with

TABLE I: PSNR, VIF and SSIM for SR and edge enhanced SR images

Image	PSNR (dB) SR	PSNR (dB) enhanced SR	VIF SR	VIF enhanced SR	SSIM SR	SSIM enhanced SR
Teacher	29.65	31.75	0.4079	0.4095	0.7849	0.7849
Leaves	24.6	24.98	0.4157	0.4169	0.7282	0.7287
Trees	31.35	33.35	0.3679	0.3690	0.6399	0.6392
Kid	32.99	34.91	0.4882	0.4893	0.8174	0.8167
Castle	30.34	32.38	0.4001	0.4019	0.7087	0.7093



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

(j) (k) (l)

(m) (n) (o)

Fig. 2: First column ground truth images a) 'Teacher', d) 'leaves', g) 'Trees', j) 'Kid', m) 'Castle'; second column (b–n) SR images; third column (c–o) PSNR enhanced SR images

(a) (b) (c)

Fig. 3: Difference images for 'Teacher', 'Trees' and 'Kid'

aim of false positives tending to zero. NC between original and extracted SV's for 100+ SR images is always more than 0.9 was observed. Thus in the present work, a threshold of $\tau = 0.9$ has been selected so as to keep the false positives to a minimum. The high value of threshold may look conservative and can result in many false negatives because every image for which $NC < \tau$ will be rejected. The negatives of difference images are shown in Fig. 3. One can note the increase in PSNR upto 2 dB due to the addition of the $Diff''$ and the authenticated SR image. Thus a single scheme serves dual purpose of authentication and fidelity improvement. The results are reported in Table I. The information fidelity and structural similarity in the SR image should not change significantly due to addition of the $Diff''$. In order to showcase this invariability, visual information fidelity (VIF) and structural similarity index measure (SSIM) between the ground truth images and its SR / PSNR enhanced SR versions was computed. The numbers enumerated in Table I indicate the structure and information fidelity is preserved in the PSNR enhanced SR image, infact it slightly improves. Moreover, to validate this fact perceptually, we show a zoomed portion of SR images with zoomed portion in Fig. 4a to Fig. 4c and its PSNR enhanced versions in Fig. 4d to Fig. 4f is shown. It can be seen that structure preserved PSNR enhanced SR image. In this paper, the images shown in Fig. 2 are super resolved using robust super-resolution [14] with the online implementation available at [13].

IV. Conclusion

This paper presents a single scheme which provides authentication and PSNR improvement for a super-resolved image. It links super-resolution and watermarking domain such that they complement each other. Results shown in Table I show that proposed scheme improve PSNR by fusing information which SR algorithm could not capture. The enhanced SR fidelity is quantified by calculating hamming distance between ground truth and enhanced SR /SR images. The hamming distance is lower in the former as compared to the latter.

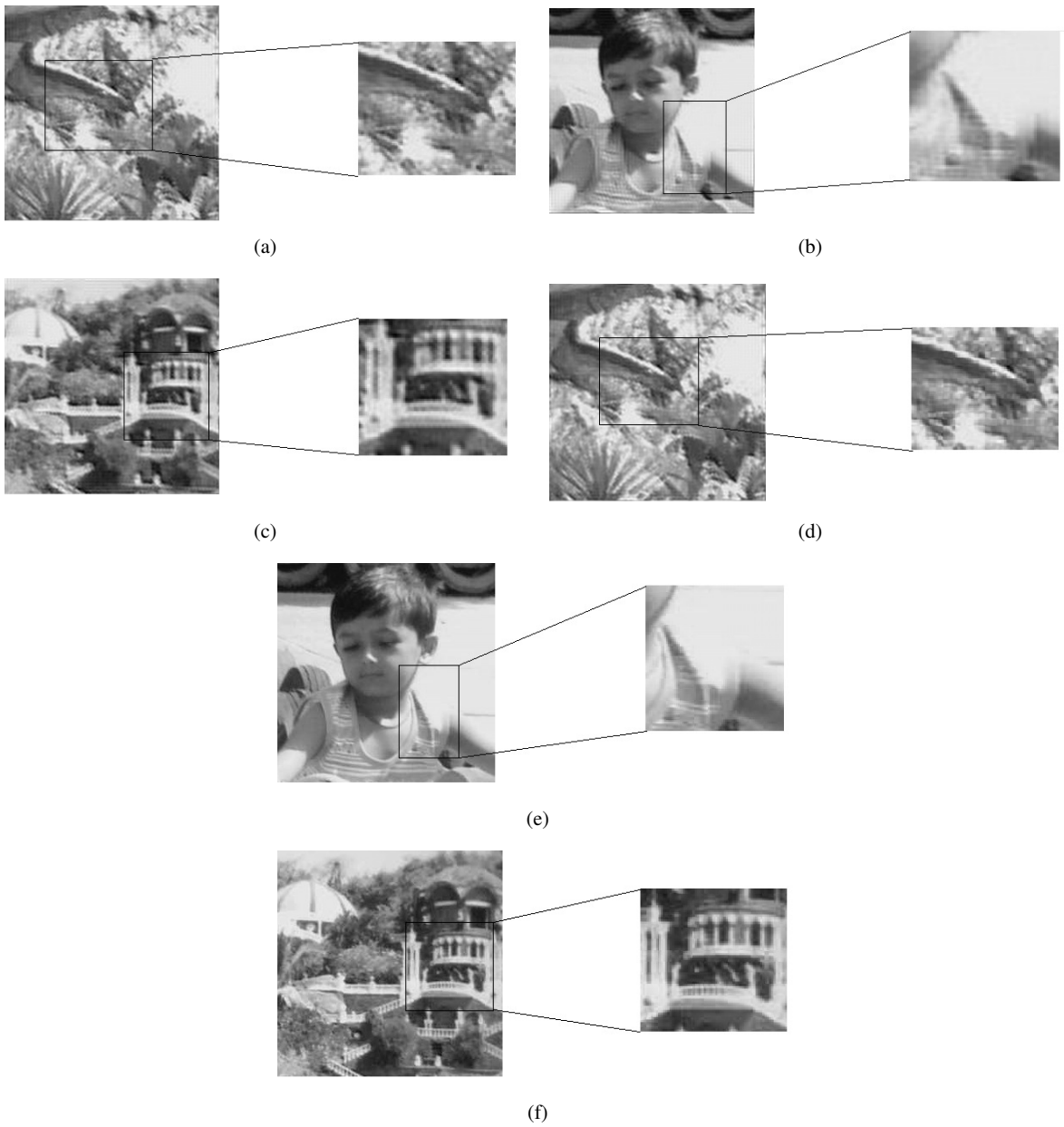


Fig. 4: a), b) and c) SR images with zoomed portion, d), e) and f) Enhanced SR images with zoomed portion

REFERENCES

[1] N. Khatri and M. V. Joshi, "Image super-resolution: use of self-learning and gabor prior," in *Computer Vision–ACCV 2012*. Springer, 2013, pp. 413–424.

[2] Z. Wang and A. C. Bovik, "Mean squared error: love it or leave it? a new look at signal fidelity measures," *Signal Processing Magazine, IEEE*, vol. 26, no. 1, pp. 98–117, 2009.

[3] M. S. Raval, M. V. Joshi, and S. Kher, "Fuzzy neural based copyright protection scheme for superresolution," in *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2013, pp. 328–332.

[4] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *Image Processing, IEEE Transactions on*, vol. 13, no. 4, pp. 600–612, 2004.

[5] V. Harikumar, M. V. Joshi, M. S. Raval, and P. P. Gajjar, "Multiresolution image fusion using compressive sensing and graph cuts," in *SPIE Remote Sensing*. International Society for Optics and Photonics, 2012, pp. 853 705–853 705.

[6] A. Burns and M. E. Johnson, "Securing health information," *IT Professional*, vol. 17, no. 1, pp. 23–29, 2015.

[7] N. Hjelm, "Benefits and drawbacks of telemedicine," *Journal of telemedicine and telecare*, vol. 11, no. 2, pp. 60–70, 2005.

[8] M. Simplicio, L. H. Iwaya, B. M. Barros, T. C. Carvalho, M. Naslund *et al.*, "Secourhealth: a delay-tolerant security framework for mobile health data collection," *Biomedical and Health Informatics, IEEE Journal of*, vol. 19, no. 2, pp. 761–772, 2015.

[9] M. Raval, P. Rege, and S. Parulkar, "Secure and robust watermarking

technique," in *TENCON 2012, IEEE Region 10 Conference*, Nov 2012, pp. 1–6.

[10] M. S. Raval and P. P. Rege, "Scalar quantization based multiple patterns data hiding technique for gray scale images," *GVIP Journal*, vol. 5, no. 9, pp. 55–61, 2005.

[11] B. Joshi, Vaibhav, M. S. Raval, P. P. Rege, and S. K. Parulkar, "Multi-stage VQ based exact authentication for biometric images," *Journal of Computing*, vol. 2, no. 1-2, pp. R3–25, 2013.

[12] X. Feng and Y. Chen, "Digital image watermarking based on super-resolution image reconstruction," in *Fuzzy Systems and Knowledge Discovery (FSKD), 2012 9th International Conference on*. IEEE, 2012, pp. 1778–1782.

[13] E. polytechnique fédérale de Lausanne. (2012, June) Super-resolution. http://lcav.epfl.ch/software/superresolution.

[14] A. Zomet, A. Rav-Acha, and S. Peleg, "Robust super-resolution," in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, vol. 1. IEEE, 2001, pp. I–645.

[15] M. A. Joshi, M. S. Raval, Y. H. Dandawate, K. R. Joshi, and S. P. Metkar, *Image and Video Compression: Fundamentals, Techniques, and Applications*. CRC Press, 2014.

[16] E. J. Candes and T. Tao, "Decoding by linear programming," *Information Theory, IEEE Transactions on*, vol. 51, no. 12, pp. 4203–4215, 2005.

[17] H. Wu, J. Dugelay, and Y. Shi, "Reversible image data hiding with contrast enhancement," *IEEE Signal Processing Letters*, vol. 22, no. 1, pp. 81–85, Jan. 2015.