Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.18576/amis/100539

Estimation of Higher-order Regression via. Sparse Representation Model for Single Image Super-resolution Algorithm

*W. Jino Hans**, *N. Venkateswaran, Srinath Narayanan and Sandeep Ramachandran*

Department of ECE, SSN College of Engineering, Chennai, India - 603110.

Received: 25 Apr. 2016, Revised: 24 Jun. 2016, Accepted: 25 Jun. 2016
Published online: 1 Sep. 2016

Abstract: Super-resolution algorithms generate high-resolution (HR) imagery from single or multiple low-resolution (LR) degraded images. In this paper, an efficient single image super-resolution (SR) algorithm using higher-order regression is proposed. Image patches extracted from HR image will have self-similar example patches near its corresponding location in the LR image. A higher-order regression function is learned using these self-similar example patches via. sparse representation model. The regression function is based on local approximations and henceforth estimated from the localized image patches. Taylor series is used as local approximation of the regression function and hence the zeroth order regression co-efficient will yield the local estimate of the regression function and the higher-order regression co-efficient will provide the local estimate of the higher-order derivative of the regression function. The learned higher-order regression mapping function is applied to LR image patches to approximate its corresponding HR version. The proposed super-resolution approach is evaluated with standard test images and is compared against state-of-the-art SR algorithms. It is observed that the proposed technique preserves sharp high-frequency (HF) details and reconstructs visually appealing HR images without introducing andy artifacts.

Keywords: Super-resolution, Regression Model, Sparse Representation

1 Introduction

Super-resolution (SR) techniques generate high-resolution (HR) image from single or set of multiple low-resolution (LR) degraded images [1]. Due to deployment of high-definition (HD) displays in electronic gadgets, the need to super-resolute available LR images without introducing counterfeit details has increased to a greater extent. As the problem posed for SR recovery is extremely ill-posed, it requires sophisticated prior model to regularize it. Various prior model such as gradient profile [2], smooth edge [3], sparse and redundant priors [4],[5] etc has been used widely in SR algorithms. The demand for efficient SR algorithm has gained popularity among researchers recently which lead to numerous self-learning SR algorithms. [6],[7].

Conventional learning based SR approaches are classified into coding based approaches and regression based approaches based on the priori which maps the LR space with its HR counterpart. Traditional coding-based approaches include the neighbour embedding (NE) based learning methods and K-nearest neighbour (k-NN) learning methods. These methods requires exhaustive search over a vast training set to find similar textures to represent fine details in an image which makes these learning methods less efficient in practical application. Recently, sparse coding prior models are used effectively to model image patches and has witnessed quite a few state-of-the-art dictionary based SR algorithms. Sparse coding model efficiently represents image patches as a sparse linear combination of a few atoms selected from an over-complete dictionary. Compared with traditional coding-based approaches, sparse coding based algorithms are efficient as it requires to search over an over-complete dictionary rather than a large dataset. In contrary, regression based approaches [8],[9] will find a direct mapping between the LR image and its corresponding HR counterpart. However, most of the conventional regression-based SR algorithms estimate the lower order regression parameter to model the correspondence between LR-HR patch-pairs. Though the zeroth order

* Corresponding author e-mail: jinohansw@ssn.edu.in

regression parameter effectively estimates the local estimate of the regression function, it is often required to approximate the higher-order estimate of the regression function to effectively model the correspondence between distorted LR and HR patches.

This paper proposes a single image SR algorithm which effectively estimates a regression function using higher-order regression co-efficients generated by Taylor series. This regression serves as a mapping between the LR-HR dictionaries represented using sparse representation model. Image patches extracted from HR image will have self-similar example patches around its corresponding location in the LR image. A dictionary-based higher-order regression model is learned by using self-similar example patches which is then used to map the correspondence between LR image patch and its HR counterpart. The regression function is based on local approximations and henceforth estimated from the localized image patches. Taylor series is used as local approximation of the regression function and hence the zeroth order regression co-efficient will yield the local estimate of the regression function and the higher-order regression co-efficient will provide the local estimate of the higher-order derivative of the regression function. The learned higher-order regression correspondence is formulated as a mapping function which is applied to LR image patches thus estimating the HR image patches.

The rest of the paper is organized as follows. Section 2 provides an overview on local kernel regression models. Section 3 proposes the higher-order regression based SR methodology. Section 3.3 provides derivative estimation via. sparse representation model and section 4 highlights the experiments conducted and results obtained. Finally, Section 5 concludes this paper.

2 Local Kernel Regression Model

Kernel regression model [8],[9],[10] is used as a correspondence mapping between two dependent variables x_i, y_i through an unknown regression function $f(.)$, which is given by

$$y_i = f(x_i) + \epsilon, \qquad i = 1, 2, ...K \tag{1}$$

In Eq.(1), ϵ is the estimation error which is insignificant under smooth prior. Given the sample pairs (x_i, y_i), the aim of this regression model is to estimate the functional value $f(x)$ at each point x. This model without any prior assumptions can be solved using N-term Taylor series.

The observations y_i can be approximated using N-term Taylor series as follows:

$$f(x_i) \simeq f(x) + \frac{1}{1!}f'(x)(x_i - x) + \frac{1}{2!}f''(x)(x_i - x)^2 + ... + \frac{1}{N!}f^N(x)(x_i - x)^N \tag{2}$$

$$\simeq \beta_0 + \beta_1(x_i - x) + \beta_2(x_i - x)^2 + ... + \beta_N(x_i - x)^N \tag{3}$$

Where $\{\beta_n\}_{n=0}^N$ are the regression coefficients. The zeroth order regression coefficient β_0 will yield the local estimate of the regression function, if Taylor series is used as local approximation of the regression function. The higher-order regression coefficients $\{\beta_n\}_{n=1}^N$ will provide the local estimate of the higher-order derivative of the regression function. Since the regression function is based on local approximations, it is desired to estimate the regression parameters $\{\beta_n\}_{n=0}^N$ from the localized samples. This is achieved by giving higher weight for nearby samples and least weight for the samples farther away. Considering this, Eq.(3) can be solved as a weighted least square problem, given as

$$\min_{\{\beta_n\}} \sum_{i=1}^{k} [y_i - \beta_0 - \beta_1(x_i - x) - \beta_2(x_i - x)^2 - ... - \beta_N(x_i - x)^N]^2 K(.) \tag{4}$$

Where $K(.)$ is the kernel function. The kernel function K can be selected as Guassian, Exponential or any other function which satisfies the desired constraints. If Guassian kernel function is chosen, then the above equation becomes

$$\min_{\{\beta_n\}} \sum_{i=1}^{k} [y_i - \beta_0 - \beta_1(x_i - x) - \beta_2(x_i - x)^2 - ... - \beta_N(x_i - x)^N]^2 e^{-\left[(x_i - x)^2/h^2\right]} \tag{5}$$

In Eq.(5), $e^{-\left[(x_i - x)^2/h^2\right]}$ is the Guassian kernel function and h is the smoothing operator. The above equation can be represented in matrix form as follows

$$\min_b \|y - Xb\|^2 K \tag{6}$$

In Eq.(6), Let

$$y = \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ y_k \end{bmatrix} \qquad b = \begin{bmatrix} \beta_0 \\ \beta_1 \\ . \\ . \\ \beta_N \end{bmatrix}$$

$$K = DIAG \begin{bmatrix} e^{-\left[(x_1 - x)^2/h^2\right]} \\ e^{-\left[(x_2 - x)^2/h^2\right]} \\ . \\ . \\ . \\ e^{-\left[(x_k - x)^2/h^2\right]} \end{bmatrix}$$

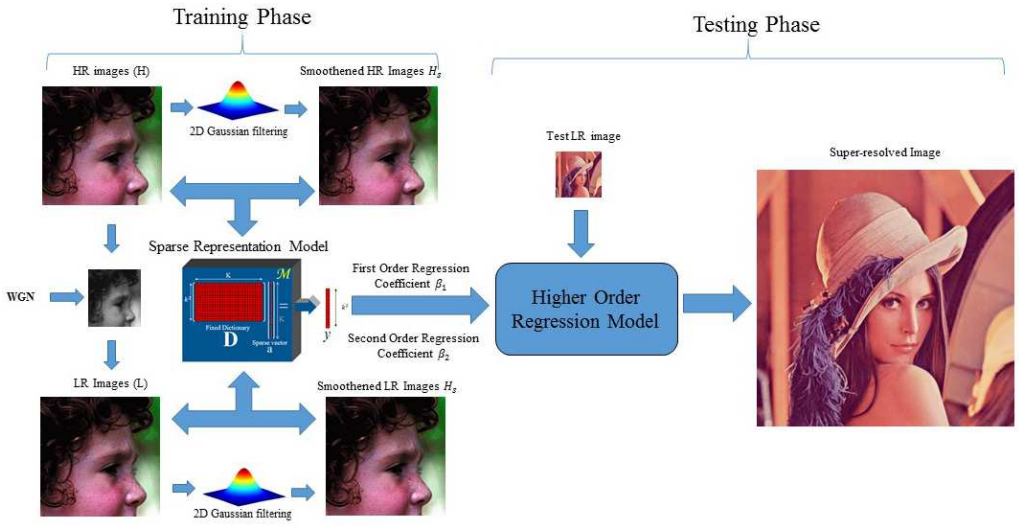


Fig. 1: Overview of the proposed super-resolution methodology

$$X = \begin{bmatrix} 1 & x_1 - x & . & . & \frac{1}{N!}(x_1 - x)^N \\ 1 & x_2 - x & & & . \\ . & . & . & & . \\ . & . & & . & . \\ 1 & x_k - x & & & \frac{1}{N!}(x_k - x)^N \end{bmatrix}$$

The solution to Eq.(6) is given as

$$\hat{b} = \left(X^T K X\right)^{-1} X^T K y \tag{7}$$

3 Single Image Super-resolution Methodology via. Dictionary based Second-order Regression Model

The proposed SR method is based on a dictionary-based second-order regression model. The overview of the proposed method is shown in Fig. 1. It uses the in-place self similarity [10] within the input image to construct a dictionary based on sparse representation. The trained dictionary is used to learn a robust non-linear second-order approximation function to map the LR image patch to its HR counterpart.

3.1 Notation

In this paper, the input LR image is denoted as $\mathbf{L} \in \mathbb{R}^{\mathbf{M}\times\mathbf{N}}$ and HR image is denoted as $\mathbf{H} \in \mathbb{R}^{\mathbf{kM}\times\mathbf{kN}}$, where k is the up-scale factor. The smoothened LR image is denoted as $\mathbf{L_s} \in \mathbb{R}^{\mathbf{M}\times\mathbf{N}}$. The smoothening is done using Guassian filtering the given LR image. The input LR image $\mathbf{L} \in \mathbb{R}^{\mathbf{M}\times\mathbf{N}}$ is up-scaled by a factor of k using bi-cubic interpolation and is denoted as $\mathbf{H_s} \in \mathbb{R}^{\mathbf{kM}\times\mathbf{kN}}$.

Image patches of size $p \times p$ extracted from $L \in \mathbb{R}^{M\times N}$ and $H \in \mathbb{R}^{kM\times kN}$ will be represented as column vectors and are denoted as $\mathbf{l}$ and $\mathbf{h}$ respectively. Similarly, image patches of size $p \times p$ extracted from $L_s \in \mathbb{R}^{M\times N}$ and $H_s \in \mathbb{R}^{kM\times kN}$ are denoted as $\mathbf{l_s}$ and $\mathbf{h_s}$ respectively. It is assumed that patches $\mathbf{l}, \mathbf{l_s}, \mathbf{h}$, and $\mathbf{h_s}$ are related with one another in terms of the center pixel. The center pixel co-ordinate for the patches $\{\mathbf{l}, \mathbf{l_s}\}$ and $\{\mathbf{h}, \mathbf{h_s}\}$ are aligned and remain same. As the input image $\mathbf{L}$ has more high frequency detail when compared with $\mathbf{L_s}$, the patches $\{\mathbf{l_s}, \mathbf{l}\}$ extracted from $\mathbf{L_s}$ and $\mathbf{L}$ respectively forms the self-exemplar LR-HR patch pairs. Similarly $\{\mathbf{h_s}, \mathbf{h}\}$ also forms LR-HR patch pairs as $\mathbf{H_s}$ lacks the HF details as in $\mathbf{H}$. Based on the LR-HR patch pairs, a higher-order regression model is learned to estimate the HR image patch $\mathbf{h}$ from its LR version $\mathbf{l}$. This is repeated iteratively for all overlapping patches of $\mathbf{L}$.

3.2 Higher-order regression model for local patch representation

Learning based SR methods aim to learn a relationship between LR-HR image patches. The proposed SR methodology establishes a non-linear mapping function f to map the given LR image patch with its corresponding HR patch. The LR-HR mapping problem is viewed as a regression problem [10] which is severely ill-posed inverse problem [11] and hence to find the mapping function, a good image prior and regularization [12] is required. The self-exemplar LR-HR image patches [13] $\{\mathbf{l}, \mathbf{l_s}\}$ serves as a good image prior to learn the HR image patch $\mathbf{h}$, from its LR counterpart $\mathbf{h_s}$. The mapping function f relates the LR-HR patch pair $\{\mathbf{h_s}, \mathbf{h}\}$ as

$\mathbf{h} = \mathbf{f}(\mathbf{h_s})$. The mapping function f can be estimated by Taylor series expansion as follows

$$h = f(h_s) \tag{8}$$

$$h = f(l_s + h_s - l_s) \tag{9}$$

$$h = f(l_s) + \nabla f^T(l_s)(h_s - l_s) + \frac{1}{2!}\nabla^2 f^T(l_s)(h_s - l_s)^2 + ... + \frac{1}{N!}\nabla^N f^T(l_s)(h_s - l_s)^N \tag{10}$$

As $\mathbf{l_s}$ is the smoothened version of $\mathbf{l}$, it is assumed $f(l_s) = l$. In Eq.10, the mapping gradient functions ∇f and $\nabla^2 f$ denote the first and second-order derivative of the mapping function f. Neglecting the higher-order derivatives, Eq.(10) becomes,

$$h = l + \nabla f^T(l_s)(h_s - l_s) + \frac{1}{2!}\nabla^2 f^T(l_s)(h_s - l_s)^2 + \eta \tag{11}$$

Where η is the residual error. As the difference between the patches $\mathbf{h_s}$-$\mathbf{l_s}$ will extract high frequency details [10], Eq.(11) will estimate the HF details from the LR-HR patch pairs which is used to extract the HR patches. As the LR-HR exemplar patches are taken from different image scale factors, it is more likely to have multiple approximate in-place self similar patches around the center pixel. This will introduce regression error, which is eliminated by taking regression for the n neighbourhood patches around the center pixel and averaging it with an weight function. This is given by

$$h = \sum_{i=1}^{n}\left(l_i + \nabla f^T(l_{s_i})(h_s - l_{s_i}) + \tfrac{1}{2!}\nabla^2 f^T(l_{s_i})(h_s - l_{s_i})^2\right)k_i \tag{12}$$

where $k_i = e^{-\left[(h_s - l_{s_i})^2/h^2\right]}$ is the Gaussian smoothing kernel.

In Eq.(12), the gradient functions ∇f and $\nabla^2 f$ has to be found to approximate the HR image patch $\mathbf{h}$. In many traditional SR approaches [10],[13] only the first order derivative is learned to approximate the HF details. Though it will approximate the HF details such as lines and arcs efficiently, it is insufficient to approximate complex structures such as curvature. In the preceding section an efficient method to approximate the first and second-order derivative via a dictionary trained with sparse representation is presented.

3.3 Estimation of regression co-efficients via. sparse representation model

The proposed method utilizes the theory of sparse representation model to estimates the derivative mapping functions ∇f and $\nabla^2 f$ by learning an over-complete dictionary. Inspired from the theory of sparse and redundant representation [5], two coupled over-complete dictionaries D_h and D_l of size $n \times K$, with n entries in K atoms, where atom represents the column vector in the dictionary, is built using the image patches extracted from HR and LR images respectively. The HR dictionary $D_h \in \mathbb{R}^{n\times K}$ with its sparse coefficient vector $a \in \mathbb{R}^{K\times 1}$ will represent the HR image patch $h \in \mathbb{R}^{n\times 1}$. As the LR dictionary is co-trained along with HR dictionary, the same sparse co-efficient vector a can be used to represent LR image. Any given HR image patch $\mathbf{h}$ can be well represented as a linear combination of the over-complete HR dictionary and sparse co-efficient vector as follows:

$$h \approx D_h a \tag{13}$$

Here $\|a\|_0 << C$, where C is the cardinality.

Any patch extracted from the observed LR image can be represented as a sparse linear combination of the co-trained LR dictionary D_l with the same sparse co-efficient vector a, such that $l \approx D_l a$.

3.3.1 *Dictionary Learning*

Let the HR image patch $\mathbf{h}$ is represented as

$$h^k = D_h a^k \tag{14}$$

where k is the number of sample patches. From this,

$$D_h = \underset{D_h}{argmin}\left\|h^k - D_h a^k\right\| \tag{15}$$

$$D_h = \underset{D_h}{argmin}\left\|H - D_h A\right\| \tag{16}$$

Where $H \rightarrow \left\{h^k\right\}_k$ and $A \rightarrow \left\{a^k\right\}_k$, Therefore,

$$H = D_h A \tag{17}$$

$$D_h = HA^{\dagger} \Rightarrow D_h = HA^T\left(AA^T\right)^{-1} \tag{18}$$

A modification of Yangs method is proposed to learn the first and second-order derivative mapping function in the following section.

3.3.2 *First and second-order derivative estimation*

The first order derivative mapping functions ∇f and $\nabla^2 f$ in Eq.12 is learned via sparse representation. The first order derivative of the HR image patch $\mathbf{h}$ is represented as

$$\nabla\mathbf{f}(\mathbf{h}) = \mathbf{Gh} \tag{19}$$

And the second-order derivative of the HR image patch $\mathbf{h}$ is represented as

$$\nabla^{\mathbf{2}}\mathbf{f}(\mathbf{h}) = \mathbf{Fh} \tag{20}$$

Here G is the gradient operator to extract the high frequency details in the image patch. And F is the Hessian operator, which estimates the second-order derivative of the image patch.

The following 2-D filter is used as gradient operator

$$G_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} \qquad G_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

Similarly the following 2-D filter is used as Hessian Operator

$$F_x = \begin{bmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix} \qquad F_y = \begin{bmatrix} 0 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

In Eq.19 & Eq.20, the HR image patch $\mathbf{h}$ can be represented as sparse linear combination of an over-complete dictionary and sparse co-efficient vector as follows

$$\mathbf{h} = \mathbf{D_h a} \tag{21}$$

Therefore Eq.19 & Eq.20 becomes,

$$\nabla \mathbf{f}(\mathbf{h}) = \mathbf{G D_h a} \tag{22}$$

And

$$\nabla^{\mathbf{2}} \mathbf{f}(\mathbf{h}) = \mathbf{F D_h a} \tag{23}$$

The LR image patch $\mathbf{l}$ is the degraded version of the HR image patch $\mathbf{h}$,

$$\mathbf{l} = \mathbf{S h}, \tag{24}$$

where $\mathbf{S} = \mathbf{QF}$. In this Q & F are the downsampling and filtering operator respectively.

$$\nabla \mathbf{f}(\mathbf{Sh}) = \mathbf{G S D_h a} \tag{25}$$

$$\nabla^{\mathbf{2}} \mathbf{f}(\mathbf{Sh}) = \mathbf{F S D_h a} \tag{26}$$

As the LR and the HR dictionary are co-trained with LR and HR image patches respectively, GSD_h and GSD_h will give D_l. Thus the above equation becomes

$$\nabla \mathbf{f}(\mathbf{l}) = \mathbf{G D_l a} \tag{27}$$

$$\nabla^{\mathbf{2}} \mathbf{f}(\mathbf{l}) = \mathbf{F D_l a} \tag{28}$$

In Eq.27 & Eq.28 , the sparse co-efficient vector a is common to both LR and HR patches respectively.The first and second-order derivative function approximated using Eq.27 & Eq.28 is applied on Eq.12 to estimate the HR image patch.

4 Results and Discussion

The proposed SR technique is evaluated both qualitatively and quantitatively by conducting experiments on various standard images taken from SR literature which includes Child, Chip, Girl etc. We also extend the efficacy of the algorithm by evaluating the same on two different type of images viz. Infrared (IR) thermal images and selfie images. Super-resolving IR thermal images primarily improves the quality of surveillance, infrared photography etc. and super-resolving selfie images will proclaim important background information in it. The IR thermal images used for evaluation are obtained from Terravic Facial IR Database and Terravic Weapon IR Database [14]. Acquisition of test selfie images which are used to evaluate the performance of the proposed algorithm are done using a state-of-the-art smartphone (Nexus 5 & IPhone 6) front camera. The test images are scaled-up by a factor of $s = 2$ and $s = 4$ for the proposed SR approach and the results are compared with other state-of-the-art SR approaches to validate its effectiveness. Few state-of-the-art SR approaches such as Yang et al.'s method [5], Kim et al. 's method [15], Dong et al. 's method [16], and He et al. 's method [17] are implemented using the open source code available on the author's webpage and are fairly compared with the results of proposed approach. Performance metrics such as root mean square error (RMSE), Peak-signal to noise ratio (PSNR) and Structural Similarity Index Measure (SSIM) [18] are used to evaluate the effectiveness of the proposed algorithm.

4.1 Experimental setup

The standard test images shown in Fig. 2 are LR images having limited pixel resolution of size 128×128. The effectiveness of the proposed algorithm is tested on these images. The performance of the proposed algorithm is also tested on a few IR thermal images as shown in Fig. 3. In Fig. 3, IR-1 and IR-2 are test IR thermal images obtained from Terravic Facial IR Database and IR-3 is obtained from Terravic Weapon IR Database. Fig. 4 shows test selfie images captured by different smartphone such as IPhone 4s, IPhone 6 and Nexus 5. In Fig. 4, #1 is a test selfie captured by Nexus 5 with a pixel resolution of 2 megapixel (MP), #2 is captured by IPhone 4s with pixel density of 1 MP and #3 is captured by IPhone 6 with 1.2 MP pixel resolution. In all experiments, patch size of 3×3 is used and the patches are lexicographically arranged. The LR test images are up-scaled by a factor of $s = 2$ and $s = 4$ using standard bi-cubic interpolation technique to avoid resolution disparity. The up-scaled image serve as the low frequency component of the target HR image. A standard deviation of 0.4 is used in the low-pass Gaussian filtering to obtain the low-frequency component L_s of the input LR image L. For clear images we use the in-place self similar samples for regression,

Fig. 2: Standard test images of size 128×128

Fig. 3: Test IR thermal images of size 64×64

Fig. 4: Test selfie images of size 256×256

whereas in the case of noisy images, we average all the 9 in-place examples for robust estimation. The dictionaries are constructed with 512 atoms using sparse representation.

4.2 Qualitative Analysis

The subjective assessment for SR approaches relies on upon few attributes of the recreated image, such as, sharpness and visual appeal. The reproduced HF elements guarantees the sharpness in the reconstructed image. It is noteworthy that the SR algorithm ought to protect the original HF details of interest. Furthermore, it should not produce any false HF details. The reproduced image is required to be original and visually appealing, being free from visual artifacts like jagging, staircase and ringing artifacts.

To objectively examine the effectiveness of the proposed technique, test images shown in Fig. 2,3 and 4 are up-scaled with a scale-up factor of $k = 3$ and $k = 4$ and are compared against other state-of-the-art SR approaches. Fig. 5 showcase the $3\times$ zooming of child, chip and girl image of size 128×128 with other SR strategies. The upper left corner of the picture is highlighted in square boxes and it contains the region of interest (ROI) in which we concentrate. For the child image in Fig.5, the sought ROI is the eye lash area. It is up-scaled by a factor $3\times$ and is considered for comparison. It is seen from Fig.5, that the proposed strategy viably protects the sharp HF points of interest. Similarily, for chip image it is observed that the proposed strategy reconstructs the sharp edge details in the characters. For instance, the subtle edge elements inside of the number 7 in the ROI is very much preserved by the proposed strategy. Also, for girl image, the ROI near the eye area is considered for comparison. It appears that the magnified ROI shown in Fig. 5(d) has rich surfaces and its sharp points of interest are reconstructed near the eye temples. Also, it is observed that the eye ball however look clear and sharp in Fig. 5(e), it present excessively sharp edges which makes the eye look unnatural. It is observed that Timotfe et al.'s outcome is visually appealing and gives sharp elements. It is observed from

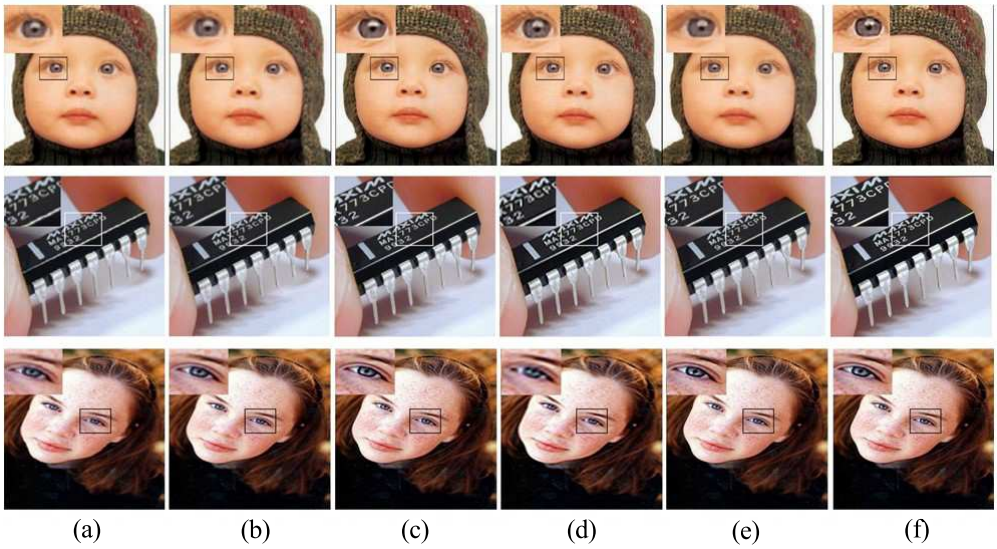


Fig. 5: Visual comparison for standard test images with state-of-the-art SR algorithms for $3\times$ magnification (a) standard test images, SR images obtained by (b) Yang et al.'method (c) Kim et al. 's method (d) Dong et al. 's method (e) He et al. 's method (f) proposed method

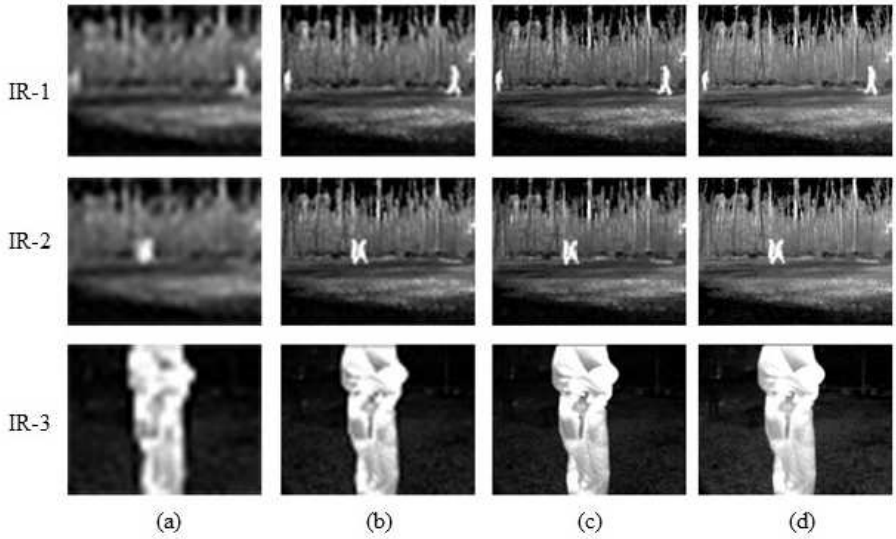


Fig. 6: Visual comparison for IR thermal images with state-of-the-art SR algorithms for $4\times$ magnification (a) Input IR thermal images, SR images obtained by (b) Yang et al.'method (c) Dong et al. 's method (d) proposed method

Fig. 5(f) that the proposed methodology not just preserves the sharp subtle elements but also the rich surfaces as well, thus yielding a high quality output for all the standard test images considered.

The application of the proposed algorithm to thermal imagery is shown in Fig. 6. The input LR IR thermal image in Fig. 6(a) shows three different input IR images. In IR-1, two persons enter the scene in opposite direction. The LR image is down-sampled such that the person in the image is not visually clear. The person in the image is characterized as a blurred white region. The image is up-scaled by a factor of $4\times$ by Yang et al.'method, Dong et al. 's method and the proposed method. The results in Fig. 6(b), shows the SR image obtained by Yang et al.'s method and is observed that it is possible to identify the persons but it does not give more information and are

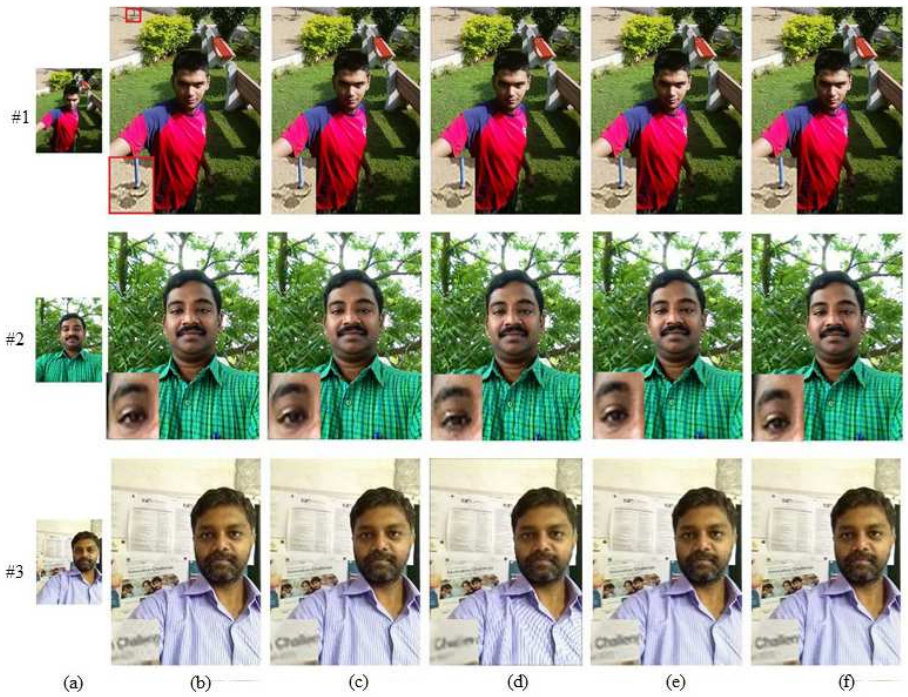


Fig. 7: Visual comparison for selfie images with state-of-the-art SR algorithms for $4\times$ magnification (a) test selfie images, SR images obtained by SR images obtained by (b) Yang et al.'method (c) Kim et al. 's method (d) Dong et al. 's method (e) He et al. 's method (f) proposed method

suffering from jaggy artifacts. Fig. 6(c) shows the image up-scaled by Dong et.al's method and Fig. 6(d) shows the image up-scaled with the proposed method. Compared with other methods, it is evident that the proposed method gives superior results which are visually pleasing. In IR-2 shown in Fig.6, two people walk very close to each other and it is well characterized in the proposed method. Similarly, in IR-3 shown in Fig.6, a person carrying weapon is shown. The input LR image is very much blurred such that the weapon is not visible. The proposed method shows better details about the weapon and it can be easily traced. This improvement in identification of separate entities in an image is of major use to weapon detection and safety.

Fig.7 portrays the visual comparison of SR results obtained by various algorithms for selfe images. It is a fact that selfie images are typical social images mainly consisting of a foreground of facial information with a distinct background. The higher-order regression provides a significant improvement in the contour development of the faces and its edges faithfully reconstructing HF details of the background information in it. This is clearly observed in Fig. 7, were we can clearly see that the facial information reconstruction is stronger in the proposed method results whereas both Yang et. al's method and Dong et al.'s results shown in Fig. 7(b) and (d) induce non-linear artifacts due to linear mapping and over-training respectively. We can also verify that beyond the second-order regression, the coefficients do not contribute greatly to the results as the visually distinct contours and edges are all significantly improved by second-order regression coefficients itself.

4.3 Quantitative Analysis

To evaluate the performance of the proposed SR algorithm, detailed comparison is performed based on its root mean square error (RMSE), peak signal to noise ratio (PSNR) and structural similarity index measure (SSIM). The results are computed and are tabulated in Table 1.

From Table 1, it is observed that the proposed SR algorithm has the minimal RMSE value across various applications. It is a result of the combined linear higher-order regression and the in-place self-similarity using correspondence of patches using sparse representation based dictionary learning. Table 1 summarizes the quantitative comparison of the proposed method with various SR algorithms on standard images. Three standard test images were used for implementing the algorithm with $k = 2$ and $k = 4$. In all the experiments the size of the test image is chosen as 128×128. For child image, the proposed method achieves better PSNR and SSIM value than Dong et. als

Table 1: Summary of RMSE, PSNR and SSIM for standard test images with scale factor $k = 2$ and $k = 4$ with other state-of-art SR approaches

Test Image	Algorithm	$k = 2$			$k = 4$		
		RMSE	**PSNR**	**SSIM**	**RMSE**	**PSNR**	**SSIM**
Chip	Yang et al.	4.7429	34.61	0.8956	6.2739	32.18	0.8145
	Kim et al.	4.7483	34.60	0.8958	6.2956	32.15	0.8148
	Dong et al.	4.7702	34.56	0.8978	6.3393	32.09	0.8144
	He et al.	4.6886	34.71	0.8971	6.2667	32.19	0.8165
	Proposed method	4.5994	34.89	0.8969	6.2523	32.21	0.8192
Girl	Yang et al.	6.2523	32.21	0.9537	7.6479	30.46	0.8921
	Kim et al.	6.2595	32.20	0.9536	7.6655	30.44	0.8918
	Dong et al.	6.4349	31.96	0.9528	7.7364	30.36	0.8914
	He et al.	6.2667	32.19	0.9536	7.6215	30.49	0.8924
	Proposed method	6.2339	32.41	0.9532	7.6040	30.51	0.8952
Child	Yang et al.	9.9092	28.21	0.8962	11.5356	26.89	0.8319
	Kim et al.	9.9321	28.19	0.8965	11.5622	26.87	0.8321
	Dong et al.	10.0587	28.08	0.8974	11.6692	26.79	0.8325
	He et al.	9.8978	28.22	0.8971	11.4958	26.92	0.8334
	Proposed method	9.8764	28.34	0.8963	11.4430	26.96	0.8358

Table 2: Summary of RMSE, PSNR and SSIM for IR thermal images with scale factor $k = 4$ with other state-of-art SR approaches

Test Image	Algorithm	$k = 2$		
		RMSE	**PSNR**	**SSIM**
IR-1	Yang et al.	5.8350	32.81	0.9128
	Kim et al.	5.8619	32.77	0.9122
	Dong et al.	5.7285	32.97	0.9133
	He et al.	5.8350	32.81	0.9127
	Proposed method	5.6956	33.02	0.9136
IR-2	Yang et al.	11.4167	26.98	0.9287
	Kim et al.	11.4430	26.96	0.9280
	Dong et al.	11.9961	26.55	0.9293
	He et al.	12.3464	26.30	0.9281
	Proposed method	11.0673	27.25	0.9312

Table 3: Summary of RMSE, PSNR and SSIM for Selfie images with scale factor $k = 4$ with other state-of-art SR approaches

Test Image	Algorithm	$k = 2$		
		RMSE	**PSNR**	**SSIM**
#1	Yang et al.	8.8316	29.21	0.9935
	Kim et al.	7.9899	30.08	0.9962
	Dong et al.	10.2457	27.92	0.9784
	He et al.	9.5728	28.51	0.9917
	Proposed method	7.9623	30.11	0.9960
#2	Yang et al.	9.9207	28.20	0.9885
	Kim et al.	9.2585	28.80	0.9922
	Dong et al.	11.5622	26.87	0.9558
	He et al.	10.8903	27.39	0.9825
	Proposed method	7.3712	30.78	0.9928

method. For the Girl image, Yang et al.'s method and He et al.'s method have almost the same PSNR values. It can be seen that for $k = 4$, the proposed method achieved the highest PSNR value and SSIM index for all the standard test images. It proves that the proposed method results in an image with highest similarity structures as in the ground truth image with minimum distortions. It is also be seen that the PSNR value and SSIM value is similar to Dong et al.s method. However, Dong et al.'s method suffer from much larger training time making it computationally complex and unviable in real-time. The computation time of the proposed algorithm is very minimal due to the efficient training of the dictionary using sparse representation.

Table 2 provides the performance details of the state-of-the-art algorithms compared with the proposed methodology for IR thermal images. For IR-1, it can be clearly seen that the PSNR value enters the 33dB range whereas all other algorithms lie in the high 32dB range. As specified previously, this improvement does not reflect the structural similarity in the IR image, as thermal images are strongly characterized by structural congruity and its application is essentially not affected by errors such as shot noise etc. Similar improvement can be seen for IR-2, where two people cross each other. The LR

image fails to recognize the distinct separation of the crossing people and shows one overlapped blob. But the SR output clearly demarcates when the outline of one person ends and the next begins, thus providing visual clarity and improving congruency in an image. An improvement in PSNR of about 0.25dB (4 times approx.) is shown in the IR-2 image from the seminal Yang et.als implementation which produces a PSNR of 26.98. Table 3 shows the performance metrics on the second application, namely selfie images. Selfie images contain strong information content and PSNR best characterizes the image, as even a mild noise affects the visual appeal of the image. The proposed methodology provides the best PSNR of 30.11 & 30.78 for selfie images I and II respectively. For selfie image 1, Kim et. als results closely reflects that of proposed algorithm due to better modelling. For image 2, the proposed results are higher than the others by at least 1dB showing an improvement in visual quality and appeal. Thus the proposed algorithm showcases strong improvement in results for different, unique but also widely used applications.

5 Conclusion

In this paper, a new single image SR algorithm using higher-order regression model is presented. The correspondence across the LR-HR image patches are learned by an higher-order regression model without having any external training samples. The first and second order regression coefficeints are obtained by learning a dictionary constructed based on sparse representation. The training patches which are extracted from the test image make sure that no counterfeit HF details are induced. Extensive experiments performed on benchmark test images confirm both the qualitative improvement and quantitative effectiveness of the proposed algorithm when compared to state-of-the-art SR approaches.

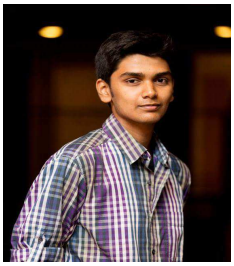
References

[1] Park, Sung C. and Park, Min K. and Kang, Moon G, Super-resolution image reconstruction: a technical overview, IEEE Signal Processing Magazine, Vol. 20, pp.21–36 (2003)
[2] Fattal and Raanan, Image Upsampling via Imposed Edge Statistics, ACM Trans. Graph., Vol. 26, (2007)
[3] Shengyang Dai and Mei Han and Wei Xu and Ying Wu and Yihong Gong, Soft edge smoothness prior for alpha channel super resolution, CVPR, (2007)
[4] Weisheng Dong and Lei Zhang and Guangming Shi and Xiaolin Wu, Image Deblurring and Super-Resolution by Adaptive Sparse Domain Selection and Adaptive Regularization, Trans. Img. Proc., Vol. 20, pp. 1838–1857, (2011)
[5] Yang, Jianchao and Wright, John and Huang, Thomas S. and Ma, Yi, Image Super-resolution via Sparse Representation, Trans. Img. Proc., Vol. 19, pp. 2861–2873, (2010)
[6] Kaibing Zhang, Xinbo Gao, Dacheng Tao and Xuelong Li, Single Image Super-Resolution With Non-Local Means and Steering Kernel Regression, IEEE Transactions on Image Processing, Vol. 21, pp. 4544–4556, (2012)
[7] Suetake, Noriaki and Sakano, Morihiko and Uchino, Eiji, Image super-resolution based on local self-similarity, Journal on Optical Review, Vol. 15, pp. 26–30 (2008)
[8] Jing Hu and Yupin Luo , Single-image superresolution based on local regression and nonlocal self-similarity, Journal of Electronic Imaging, Vol. 23, (2014)
[9] Ni, Karl S. and Nguyen, Truong Q., Image Superresolution Using Support Vector Regression., IEEE Transactions on Image Processing, Vol. 16, pp. 1596–1610 (2007)
[10] Yang, Jianchao, Lin, Zhe, Cohen, Scott, Fast Image Super-Resolution Based on In-Place Example Regression., CVPR, pp. 1059-1066, (2013)
[11] Sina Farsiu and Dirk Robinson and Michael Elad and Peyman Milanfar, Fast and Robust Multi-Frame Super-Resolution, IEEE Transactions on Image Processing, Vol. 13, pp. 1327–1344 (2003)
[12] Gao, Xinbo and Zhang, Kaibing and Tao, Dacheng and Li, Xuelong, Joint Learning for Single-Image Super-Resolution via a Coupled Constraint., IEEE Transactions on Image Processing, Vol. 21, pp. 469-480, (2012)
[13] Tang, Y., Li, L., Li, X., Learning Similarity With Multikernel Method, IEEE Transactions on Systems, Man, and Cybernetics, Vol. 41, No. 1, pp. 131–138, (2011)
[14] Zheng Wu, Nathan Fuller, Diane Theriault, Margrit Betke, ”A Thermal Infrared Video Benchmark for Visual Analysis”, in Proceeding of 10th IEEE Workshop on Perception Beyond the Visible Spectrum (PBVS), Columbus, Ohio, USA, 2014.
[15] Kim, Kwang In and Kwon, Younghee, Single-Image Super-Resolution Using Sparse Regression and Natural Image Prior, IEEE transactions on Pattern Analysis and Machine Intelligence, Vol. 32, No. 6, pp. 1127–1133, (2010)
[16] Weisheng Dong, Lei Zhang, Rastislav Lukac, and Guangming Shi, Sparse Representation Based Image Interpolation With Nonlocal Autoregressive Modeling, IEEE Transactions on Image Processing, Vol. 22, pp. 1382–1394, (2013)
[17] H. He and W.-C. Siu, Single image super-resolution using Gaussian process regression, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 11), pp. 449–456, IEEE, Providence, RI, USA, June 2011.
[18] Zhou Wang and Alan C. Bovik and Hamid R. Sheikh and Eero P. Simoncelli, Image Quality Assessment: From Error Visibility to Structural Similarity, IEEE Transactions on Image Processing, Vol.13, No. 4, pp. 600–612 (2004)

W. Jino Hans received the B.E. degree in Electronics and Communication Engineering from Madurai Kamaraj University, Tamil Nadu, India in 2004, M.Tech in Embedded systems and Technology from SRM University, India in 2006 and currently pursuing Ph.D. from Anna University Chennai, India. He is presently an Assistant professor in the department of Electronics and Communication Engineering, SSN College of Engineering, Chennai, India. His research interest include Inverse problems in Image processing, Single image Super-resolution Algorithms, Matrix regression Analysis.

N. Venkateswaran received his PhD from Anna University, India in Signal and Image Processing. He received his M.Tech degree in Communication Systems from Pondicherry Engineering College, Pondicherry University in 1995. Currently, Dr.N.Venkateswaran is a professor in the Department of Electronics and Communication Engineering at SSN college of Engineering, India. Prior to June 2010, he was Professor at Sri Venkateswara College of Engineering, India. He has more than 25 years of teaching, research and industry experience. Dr.N.Venkateswaran has published more than 60 technical papers in various journals and conferences. His research interest include Signal and Image processing, Sparse signal representation, Wireless Communications and its applications. He is a member of IEEE and Fellow of IETE (India).

Srinath Narayanan received his B.E degree in Electronics and Communication Engineering from SSN College of Engineering, under Anna University in 2016. He is currently admitted as a graduate student into Jacobs School of Engineering, at University of California, San Diego. He will be pursuing his Masters in Electrical and Computer Engineering with Intelligent Systems, Robotics and Control specialization. His main research interests include the confluence of image processing and machine learning, computer vision and data science.



Sandeep Ramachandran received his B.E degree in Electronics and Communication Engineering from SSN College of Engineering, under Anna University in 2016. He is currently working as a Software Engineer in Zoho Corporation, Chennai. His main research interests include image processing and machine learning.