




Enhancement of Infrared Images Using Super Resolution Techniques Based on Big Data Processing

Fathi E. Abd El-Samie¹  • Huda I. Ashiba² • H. Shendy¹ • Hala M. Mansour³ • Hossameldin M. Ahmed³ • Taha E. Taha¹ • Moawad I. Dessouky¹ • Mohamed F. Elkordy¹ • Mohammed Abd-Elnaby^{1,4} • Adel S. El-Fishawy¹

Received: 20 February 2018 / Revised: 2 April 2019 / Accepted: 11 April 2019 /
Published online: 9 December 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

This paper presents a super-resolution (SR) technique for enhancement of infrared (IR) images. The suggested technique relies on the image acquisition model, which benefits from the sparse representations of low-resolution (LR) and high-resolution (HR) patches of the IR images. It uses bicubic interpolation and minimum mean square error (MMSE) estimation in the prediction of the HR image with a scheme that can be interpreted as a feed-forward neural network. The suggested algorithm to overcome the problem of having only LR images due to hardware limitations is represented with a big data processing model. The performance of the suggested technique is compared with that of the standard regularized image interpolation technique as well as an adaptive block-by-block least-squares (LS) interpolation technique from the peak signal-to-noise ratio (PSNR) perspective. Numerical results reveal the superiority of the proposed SR technique.

✉ Fathi E. Abd El-Samie
fathi_sayed@yahoo.com

✉ Huda I. Ashiba
eng_h_2006@yahoo.com

H. Shendy
hrshendy@yahoo.com

Hala M. Mansour
hala.mansour@feng.bu.edu.eg

Taha E. Taha
taha117@hotmail.com

Moawad I. Dessouky
dr_moawad@yahoo.com

Mohammed Abd-Elnaby
moh_naby@yahoo.com

Adel S. El-Fishawy
aelfishawy@hotmail.com

Extended author information available on the last page of the article

Keywords Single-image super-resolution · Image interpolation · IR images · Neural networks · Big data processing

1 Introduction

Image interpolation is an essential tool for image processing scientists to acquire an HR image from an LR one. The motivation for the research in this field of image processing is the inability of most imaging sensors to obtain the required HR image with a moderate cost. The application of image interpolation in the field of IR image processing is a promising trend as IR images usually have low resolutions. Traditional kernel-based image interpolation techniques have been broadly studied. Most of the research in the kernel-based interpolation techniques was directed to obtaining the best interpolation basis functions [6, 12, 13, 22–24, 26, 33].

Splines, Keys' and optimal maximal order of minimum support (O-MOMS) interpolation techniques are the most common families utilized for image interpolation. These conventional techniques are space-invariant and they do not consider the spatial activities of the image to be interpolated. They also do not consider the mathematical model of the imaging process with a specific type of sensors. Spatially-adaptive kernel-based techniques depend on concepts such as the warped-distance concept. Although these adaptive techniques improve the quality of the interpolated image, especially, near edges, they still do not take into consideration the image capturing model. The entire kernel-based techniques and their adaptive variants can be considered as signal synthesis techniques [28, 33, 38].

This paper presents an efficient technique for obtaining IR images with high resolution. It is based on image SR concepts for resolution enhancement of LR images. The importance of the proposed SR technique is to overcome the limitations of having only LR images due to hardware limitations, because images are acquired from LR sensors.

The organization of this paper is as follows. Section 2 gives an explanation of the research motivations and the related work. Section 3 gives an explanation of the regularized image interpolation technique. Section 4 covers polynomial-based image interpolation. Section 5 gives an explanation of LS interpolation of IR images. Section 6 presents the concepts of SR applied to IR images. Section 7 gives a discussion of single-image SR applied to IR images. Section 8 presents the proposed technique. Section 9 gives the simulation results. Finally, section 10 gives the concluding remarks.

2 Motivations and related work

The rapid and massively increasing development of image processing technologies makes the acquisition of SR images a hot research topic with a very wide range of applications [32]. Moreover, applying SR techniques on low-quality images like IR images is very challenging. The main limitations to acquiring an HR image directly through an IR camera is the manufacturing difficulty, material properties, and imaging environment. Some researches offered designs to enhance the scanning section of the IR imaging system [2, 10]. These designs offered enhancement schemes using four different-angle plate refractors placed in parallel inside the device. These designs are still application-limited due to fabrication complexity, size, and cost.

The manufacturing challenges directed research towards the acquisition of HR infrared images from one or more LR images using SR techniques [3–5, 9, 11, 15, 21, 25, 27, 29, 35]. The SR acquisition techniques can be classified into three categories; interpolation techniques [3, 15, 27], reconstruction techniques [3, 9, 25] and learning techniques [5, 11, 21, 29, 35]. Interpolation techniques are the earliest and the most essential methodologies adopted for SR acquisition, but they give images with low quality. On the other hand, reconstruction techniques provide better-quality images by assuming that the LR image is a result of a degradation model comprising noise, distortion, blurring, and down-sampling, but there is still a need for prior knowledge to accurately obtain HR images from the LR ones. Learning techniques represent a new era in image processing resulting in satisfactory results without the need for prior knowledge about the LR degradation model. So, the concepts of learning are adopted in this paper.

Applying SR techniques on IR images is a very interesting and growing research topic. Researchers offered different methods for SR acquisition from IR images based on learning techniques [20, 30, 36, 37]. One of these methods [30] offers the learning stage through combining the information from visible and IR images as images from different sensors carrying complementary information for the same scene. Another method [36] presents the reconstruction of SR images based on their sparse representations by offering a pair of dictionaries, in which the HR and LR patches share the same sparse representations. Another trend adopts compressed sensing (CS) [20, 37] for SR acquisition of images [7] to solve the sparsity reconstruction problem.

3 Regularized image interpolation

Regularization theory, which was basically introduced by Tikhonov and Miller, provides a formal basis for the development of regularized solutions for ill-posed problems [8]. The stabilizing function approach is one of the basic methodologies for the development of regularized solutions. According to this approach, an ill-posed problem can be formulated as the constrained minimization of a certain function, called the stabilizing function [8]. The specific constraints imposed by the stabilizing function approach on the solution depend on the form and the properties of the function used.

From the nature of the problem, these constraints are necessarily related to the a priori information regarding the expected regularized solution. According to the regularization approach, the solution of Eq. (1) is obtained by the minimization of the cost function [8]:

$$\Psi(\hat{\mathbf{f}}) = \|\mathbf{g} - \mathbf{D}\hat{\mathbf{f}}\|^2 + \lambda \|\mathbf{Q}\hat{\mathbf{f}}\|^2 \quad (1)$$

where \mathbf{Q} is the regularization operator and λ is the regularization parameter.

This minimization is accomplished by taking the derivative of the cost function yielding:

$$\frac{\partial \Psi(\hat{\mathbf{f}})}{\partial \hat{\mathbf{f}}} = \mathbf{0} = 2\mathbf{D}^T(\mathbf{g} - \mathbf{D}\hat{\mathbf{f}}) - 2\lambda\mathbf{Q}^T\mathbf{Q}\hat{\mathbf{f}} \quad (2)$$

$$2\hat{\mathbf{f}}(\mathbf{D}^T\mathbf{D} + \lambda\mathbf{Q}^T\mathbf{Q}) - 2\mathbf{D}^T\mathbf{g} = \mathbf{0} \quad (3)$$

where T refers to matrix transpose.

Solving for that $\hat{\mathbf{f}}$ that provides the minimum of the cost function yields [26]:

$$\begin{aligned}\hat{\mathbf{f}} &= (\mathbf{D}^T \mathbf{D} + \lambda \mathbf{Q}^T \mathbf{Q})^{-1} \mathbf{D}^T \mathbf{g} \\ &= \mathbf{A}(\lambda) \mathbf{g}\end{aligned}\quad (4)$$

The rule of the regularization operator \mathbf{Q} is to move the small eigenvalues of \mathbf{D} away from zero, while leaving the large eigenvalues unchanged. It also incorporates prior knowledge about the required degree of smoothness of the estimated image into the interpolation process.

The generality of the linear operator \mathbf{Q} allows the development of a variety of constraints that can be incorporated into the interpolation operation. For instance [8]:

1. $\mathbf{Q} = \mathbf{I}$. In this case, the regularized solution reduces to the regularized inverse filter solution, which is named the pseudo-inverse filter solution, and it is represented as:

$$\hat{\mathbf{f}} = (\mathbf{D}^T \mathbf{D} + \lambda \mathbf{I})^{-1} \mathbf{D}^T \mathbf{g} \quad (5)$$

2. \mathbf{Q} = finite difference matrix. In this case, the operator \mathbf{Q} is chosen to minimize the second order (or higher order) difference energy of the estimated image.

The 2-D Laplacian is preferred for minimizing the second order difference energy. It is the most popular regularization operator. It is adopted in this paper. The regularization parameter λ controls the trade-off between fidelity of the data and the smoothness of the solution [8].

The solution of the regularized image interpolation problem is implemented by the segmentation of the LR image into overlapping segments and the interpolation of each segment separately using Eq. (7) as an inversion process. It is clear that if a global regularization parameter is used, a single matrix inversion process for a matrix of moderate dimensions is required, because the term $(\mathbf{D}^T \mathbf{D} + \lambda \mathbf{Q}^T \mathbf{Q})^{-1}$ is independent of the image to be interpolated. The interpolation formula can be written in the following form:

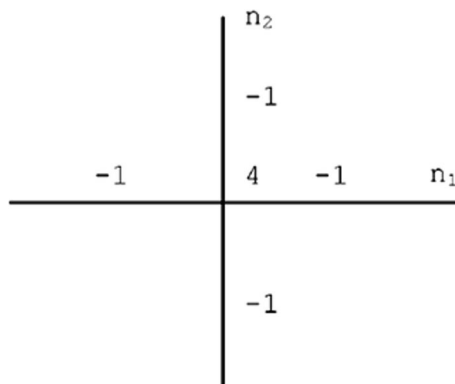


Fig. 1 2-D Laplacian operator

$$\hat{\mathbf{f}}_{i,j} = (\mathbf{D}^T \mathbf{D} + \lambda \mathbf{Q}^T \mathbf{Q})^{-1} \mathbf{D}^T \mathbf{g}_{i,j} \tag{6}$$

where $\mathbf{g}_{i,j}$ and $\hat{\mathbf{f}}_{i,j}$ are the lexicographically-ordered LR and the estimated HR blocks at position (i, j) , respectively.

4 Polynomial-based image interpolation

Traditional signal interpolation techniques such as the basic spline (B-spline) interpolation approximate a continuous function from the discrete signal values available and resample this continuous function again. This continuous function is represented as follows [22]:

$$\hat{f}(x) = \sum_k c_k \beta^n(x-x_k) \tag{7}$$

where $\beta^n(x)$ denotes the central B-spline of degree n that is given by [23]:

$$\beta^n(x) = \underbrace{\beta^0 * \beta^0 * \dots * \beta^0}_{(n+1) \text{ times}} \tag{8}$$

where the $*$ denotes convolution.

From the family of polynomial splines, the cubic spline tends to be the most popular. The closed-form approximation of the cubic spline basis function is given by [33]:

$$\beta^3(x) = \begin{cases} \frac{2}{3} - |x|^2 + \frac{|x|^3}{2} & 0 \leq |x| < 1 \\ \frac{(2-|x|)^3}{6} & 1 \leq |x| < 2 \\ 0 & 2 \leq |x| \end{cases} \tag{9}$$

For equally-spaced 1-D sampled data $f(x_k)$, we define the distances between x and x_k and between x_{k+1} and x as [28]:

$$s = x-x_k, \quad 1-s = x_{k+1}-x. \tag{10}$$

where $x_k \leq x \leq x_{k+1}$

Thus, the cubic spline interpolation process can be expressed as follows [28]:

$$\hat{f}(x) = c_{k-1} \left[(3+s)^3 - 4(2+s)^3 + 6(1+s)^3 - 4s^3 \right] / 6 + c_k \left[(2+s)^3 - 4(1+s)^3 + 6s^3 \right] / 6 + c_{k+1} \left[(1+s)^3 - 4s^3 \right] / 6 + c_{k+2} s^3 / 6 \tag{11}$$

The coefficients c_k for cubic spline interpolation are estimated using a pre-processing filtering step. For image interpolation, this process is applied along rows and then along columns [38].

5 Least-squares interpolation of IR images

In the application of the adaptive LS algorithm on IR images, the IR image to be interpolated is split into small overlapping blocks, and the objective is to obtain an interpolated version of each block. We suppose that the relationship between the available LR block and the estimated HR block is given by [3]:

$$\hat{\mathbf{X}}_{i,j} = \mathbf{W}\mathbf{Y}_{i,j} \tag{12}$$

where $\mathbf{Y}_{i,j}$ and $\hat{\mathbf{X}}_{i,j}$ are the lexicographically-ordered LR and estimated HR blocks at the block indices (i, j) , respectively. \mathbf{W} is the weight matrix required to obtain the HR vector from the LR vector. This matrix is required to be adaptive from block to block to accommodate for the local activity levels of each block.

The first look at Eq. (13) leads to the LS solution that can be obtained by minimizing the Mean Square Error (MSE) of the estimation as follows:

$$\Psi = \|\mathbf{X}_{i,j} - \hat{\mathbf{X}}_{i,j}\|^2 = \|\mathbf{X}_{i,j} - \mathbf{W}\mathbf{Y}_{i,j}\|^2 \tag{13}$$

Differentiating both sides of Eq.(14) with respect to \mathbf{W} gives:

$$\frac{\partial \Psi}{\partial \mathbf{W}} = -2(\mathbf{X}_{i,j} - \hat{\mathbf{X}}_{i,j})(\mathbf{Y}_{i,j})^T \tag{14}$$

This minimization leads directly to the following solution for \mathbf{W}

$$\mathbf{W}^{k+1} = \mathbf{W}^k - \eta \left[\frac{\partial \Psi}{\partial \mathbf{W}} \right]^k = \mathbf{W}^k + \mu (\mathbf{X}_{i,j} - \hat{\mathbf{X}}_{i,j}^k) (\mathbf{Y}_{i,j}^k)^T \tag{15}$$

where η is a constant, μ is the convergence parameter and k is the iteration number.

The utilization of the above equation in estimating the weight matrix \mathbf{W} requires the samples of the original HR block $\mathbf{X}_{i,j}$ to be known, which is not practical. This issue can be fixed by deducing the weights from another HR image and using these weights to interpolate the available LR image. This approach is expected to yield poor visual quality of the interpolated image.

An alternative to the above-mentioned algorithm is to consider the model that relates the available LR block to the original HR block, illustrated in Fig. 2. This model is offered by the following relation [16]:

$$\mathbf{Y}_{i,j} = \mathbf{H}\mathbf{X}_{i,j} \tag{16}$$

The matrix \mathbf{H} represents the filtering and down-sampling process that transforms the HR block to the LR block.

Thus, we can deduce the following cost function:

$$\Phi = \|\mathbf{H}(\mathbf{X}_{i,j} - \hat{\mathbf{X}}_{i,j})\|^2 \tag{17}$$

The above equation means reducing the MSE between the available LR block and a down-sampled version of the estimated HR block.

This leads to:

$$\Phi = \|\mathbf{Y}_{i,j} - \mathbf{H}\hat{\mathbf{X}}_{i,j}\|^2 = \|\mathbf{Y}_{i,j} - \mathbf{H}\mathbf{W}\mathbf{Y}_{i,j}\|^2 \tag{18}$$

Differentiating Eq. (18) with respect to \mathbf{W} :

$$\frac{\partial \Phi}{\partial \mathbf{W}} = -2\mathbf{H}^t (\mathbf{Y}_{i,j} - \mathbf{H}\hat{\mathbf{X}}_{i,j}) (\mathbf{Y}_{i,j})^T \tag{19}$$

Using Eq. (19), the weight matrix can be deduced with the following equation:

$$\mathbf{W}^{k+1} = \mathbf{W}^k - \eta \left[\frac{\partial \Phi}{\partial \mathbf{W}} \right]^k = \mathbf{W}^k + \mu \mathbf{H}^t (\mathbf{Y}_{i,j} - \mathbf{H}\hat{\mathbf{X}}_{i,j}^k) (\mathbf{Y}_{i,j}^k)^T \tag{20}$$

The adaptation of Eq. (20) can be easily performed, since it does not require the original HR block to be known a priori.

6 Super-resolution applied to IR images

The representations of the LR and HR images as vectors are $\mathbf{z}_l \in \mathbf{R}^{M_l}$ and $\mathbf{y}_h \in \mathbf{R}^{M_h}$, respectively, where $M_h = q^2 M_l$ and q is a scale-up factor of type integer that is higher than 1. We also refer to $\mathbf{B} \in \mathbf{R}^{M_h \times M_h}$ as the blur operator, and $\mathbf{H} \in \mathbf{R}^{M_l \times M_h}$ as the decimation operator with q in each axis. A well-known anti-aliasing low-pass filter is applied on the image to generate an LR image from an HR image [19],

$$\mathbf{z}_l = \mathbf{H}\mathbf{B}\mathbf{y}_h + \mathbf{A} \tag{21}$$

where \mathbf{A} is an additive noise vector or error. The relative problem of the reconstruction of \mathbf{y}_h from \mathbf{z}_l is denoted as zooming and deblurring. A bicubic low-pass filter and a Gaussian low-pass filter are possible blur kernels (blur operators \mathbf{B}).

As the work is based on the sparse representation of patch pairs, we have to learn the correspondence between LR and HR patches of the same dimensions. So, we obtain an image $\mathbf{y}_l \in \mathbf{R}^{M_h}$ by applying bicubic interpolation on the input LR image, and as we address the zooming and deblurring setup, we look forward at recovering the difference image $\hat{\mathbf{y}}_{hl} = \mathbf{y}_h - \mathbf{y}_l$ and then apply $\hat{\mathbf{y}}_h = \hat{\mathbf{y}}_{hl} + \mathbf{y}_l$ to get the final recovery. So, we keep the LR details and only predict the lost HR details. For the concept of image reconstruction based on patches, let an image patch that is centered at location Q with size $\sqrt{m} \times \sqrt{m}$ and that is extracted from the image vector \mathbf{y} of size M_h by the linear operator \mathbf{R}_Q to be $\mathbf{P}^Q = \mathbf{R}_Q \mathbf{y}$. A local model could be suggested to predict an HR patch $\mathbf{P}_h^Q = \mathbf{R}_Q \mathbf{y}_h$ from an LR one $\mathbf{P}_l^Q = \mathbf{R}_Q \mathbf{y}_l$. Once obtaining all HR patch predictions, the recovery of the HR image takes place by averaging of the overlapping recovered patches on their overlaps. Another factor that should be taken into consideration is the trade-off between the reconstructed image quality and the run time, while choosing the size of the overlap between adjacent patches. In order to achieve the best

reconstruction quality, work has to be performed with maximally-overlapping patches (overlapping between adjacent patches is of $\sqrt{m} \times \sqrt{m-1}$ pixels in the horizontal and vertical directions).

Finally, we briefly mention the sparsity-based synthesis model. The core idea of this model is that a signal $\mathbf{S} \in \mathbf{R}^m$ can be represented as a linear combination of a few atoms (signal prototypes) taken from a dictionary $\mathbf{D} \in \mathbf{R}^{m \times n}$, namely $\mathbf{S} = \mathbf{D}\boldsymbol{\alpha} + \boldsymbol{\eta}$, where $\boldsymbol{\alpha} \in \mathbf{R}^n$ is the sparse representation vector and $\boldsymbol{\eta}$ is the noise or model error. Similar to the previous approaches [34], we assume that each LR patch can be represented over a dictionary $\mathbf{D}_l \in \mathbf{R}^{m \times n_l}$ by a sparse vector $\boldsymbol{\alpha}_l \in \mathbf{R}^{n_l}$, and similarly an HR patch is represented over $\mathbf{D}_h \in \mathbf{R}^{m \times n_h}$ by $\boldsymbol{\alpha}_h \in \mathbf{R}^{n_h}$.

7 Single-image super-resolution applied to IR images

The concept of this model is to predict the missing HR details for each LR patch through a different number of atom pairs of LR and HR dictionaries. The difference in atom numbers between LR and HR dictionaries is natural as each dictionary characterizes a signal with a certain quality. So, for the low-quality one, the dictionary contains fewer atoms than those of the high-quality one. In addition, a small and orthogonal dictionary for the LR patches for complete and under-complete cases, which offers low complexity in sparse coding computations, is considered [14, 21].

First, we start with the low-cost pursuit stage to obtain $\boldsymbol{\alpha}_l$ which indicates the LR coefficients. Then, a suggested statistical parametric model is used for the prediction of $\boldsymbol{\alpha}_h$ which indicates the HR representation vector of each patch from its corresponding LR coefficients $\boldsymbol{\alpha}_l$. Finally, the single-image SR scheme is presented as a result of the suggested model.

7.1 Low-cost pursuit stage

To sparsely represent the patches, an under-complete dictionary ($n_l < m$) is sufficient enough. To allow a low-cost scale-up scheme, \mathbf{D}_l is assumed as an under-complete orthonormal dictionary. Therefore, the inner products of the LR patch with the dictionary atoms results in the LR coefficients.

$$\boldsymbol{\alpha}_l = (\mathbf{D}_l)^T \mathbf{P}_l \tag{22}$$

A convolutional network is then used to compute the LR coefficients for all overlapping patches $\{\mathbf{P}_l^O\}$. The sparsity pattern $\mathbf{x}_l \in \{-1, 1\}^{n_l}$ is computed as:

$$x_{l,j} = \begin{cases} 1, & |\alpha_{l,j}| > \delta \\ -1, & \text{Otherwise} \end{cases}, \quad \forall_j = 1, \dots, n_l, \tag{23}$$

where δ is the maximal threshold satisfying that set, adaptively, for each LR patch based on a residual error criterion.

$$\sum_{j=1}^{n_l} |a_{l,j}|^2 (|a_{l,j}| \leq \delta) \leq m\rho^2 \tag{24}$$

where ρ is a pre-specified parameter that indicates the targeted accuracy of the LR sparse representation.

7.2 The model

As explained in the previous sub-section, the LR dictionary \mathbf{D}_l is an under-complete orthogonal dictionary. On the other hand, the HR dictionary \mathbf{D}_h is assumed to be a complete or over-complete dictionary in order to allow a sufficient representation power. As there is a difference in the number of atoms between \mathbf{D}_l and \mathbf{D}_h , it is not valid to assume the LR and HR dictionaries to have the same sparsity pattern representations as in all previous works that consider dictionaries. So, a model is required to capture the relations between the two different sparsity patterns $-\mathbf{x}_l \in \{-1, 1\}^{n_l}$ for the LR patch and $\mathbf{x}_h \in \{-1, 1\}^{n_h}$ for the HR patch. It is necessary to consider the Boltzmann machine prior.

$$Pr(s) = \frac{1}{Z} \exp\left(\mathbf{b}^T \mathbf{x} + \frac{1}{2} \mathbf{x}^T \mathbf{V} \mathbf{x}\right) \tag{25}$$

where $\mathbf{b} \in \mathbf{R}^n$ is a bias vector and $\mathbf{V} \in \mathbf{R}^{n \times n}$ is an interaction matrix used within the sparsity pattern $\mathbf{x} \in \{1, -1\}^n$ of a single representation vector to capture statistical dependencies. We need to capture the dependencies between the sparsity patterns of the LR-HR pair. So, we use a variant of the Boltzmann machine, named restricted Boltzmann machine and given by the conditional probability.

$$\begin{aligned} Pr(\mathbf{x}_h | \mathbf{x}_l) &= \frac{1}{Z_1} \exp(\mathbf{b}_h^T \mathbf{x}_h + \mathbf{x}_h^T \mathbf{V}_{hl} \mathbf{x}_l) \\ &= \prod_{j=1}^{n_h} \Phi\left(\left(b_{h,j} + \mathbf{V}_{hl,j}^T \mathbf{x}_l\right) x_{h,j}\right) \end{aligned} \tag{26}$$

where $\mathbf{b}_h \in \mathbf{R}^{n_h}$ is a bias vector for the HR sparsity pattern, $\mathbf{V}_{hl} \in \mathbf{R}^{n_h \times n_l}$ is an interaction matrix connecting between the LR and HR sparsity patterns, and $\Phi(z) = (1 + \exp(-2z))^{-1}$ is the sigmoid function. The last equality in the previous equation holds since the entries of \mathbf{x}_h are statistically independent given \mathbf{x}_l .

This choice leads to a closed-form formula for the conditional marginal probability of each entry in \mathbf{x}_h given \mathbf{x}_l ,

$$Pr(x_{h,j} = 1 | \mathbf{x}_l) = \Phi\left(b_{h,j} + \mathbf{V}_{hl,j}^T \mathbf{x}_l\right), \quad \forall_j = 1, \dots, n_h, \tag{27}$$

that aligns with the sigmoid unit in neural networks. Then, HR coefficients α_h are addressed. Given the sparsity pattern \mathbf{s}_h and the LR coefficients α_l , the following model is suggested:

$$\alpha_{h,j} = \begin{cases} u_j & , \quad x_{h,j} = 1 \\ 0 & , \quad x_{h,j} = -1 \end{cases}, \quad \forall_j = 1, \dots, n_h, \tag{28}$$

where $\mathbf{u} \in \mathbf{R}^{n_h}$ is assumed to be Gaussian distributed given α_l , so that $\mathbf{u}|\alpha_l \sim \mathcal{N}(\mathbf{S}_{hl}\alpha_l, \Sigma_{hl})$ with $\mathbf{S}_{hl} \in \mathbf{R}^{n_h \times n_l}$ and $\Sigma_{hl} \in \mathbf{R}^{n_h \times n_h}$. Straightforward considerations lead to the following conditional expectation,

$$E[\alpha_{h,j}|x_{h,j} = 1, \alpha_l] = \mathbf{S}_{hl,j}^T \alpha_l, \quad \forall j = 1, \dots, n_h. \tag{29}$$

The last equations for each sparsity pattern \mathbf{x}_h perform a different mapping from α_l to α_h .

However, all 2^{n_h} possible mappings are described through the same matrix \mathbf{S}_{hl} . Notice that the prediction in this model is linear only, when the sparsity pattern \mathbf{x}_h is known, and as we will see in the next sub-section, the final estimator for α_l , j given α_l and \mathbf{x}_l is nonlinear.

7.3 Inference

An MMSE estimator is used for the prediction of each entry in α_h from \mathbf{x}_l and α_l [14, 21]

$$\begin{aligned} \hat{\alpha}_{h,j} &= E[\alpha_{h,j}|\mathbf{x}_l, \alpha_l] \\ &= \sum_{\mathbf{x}_h \in \Gamma_j} E[\alpha_{h,j}|\mathbf{x}_h, \mathbf{x}_l, \alpha_l] Pr(\mathbf{x}_h|\mathbf{x}_l, \alpha_l) \\ &= \sum_{\mathbf{x}_h \in \Gamma_j} E[\alpha_{h,j}|x_{h,j} = 1, \alpha_l] Pr(\mathbf{x}_h|\mathbf{x}_l) \\ &= E[\alpha_{h,j}|x_{h,j} = 1, \alpha_l] Pr(x_{h,j} = 1|\mathbf{x}_l) = (\mathbf{S}_{hl,j}^T \alpha_l) \Phi(b_{h,j} + \mathbf{V}_{hl,j}^T \mathbf{x}_l), \end{aligned} \tag{30}$$

where $\Gamma_j = \{\gamma \in \mathbf{R}^{n_h} : g_j = 1\}$.

7.4 Neural network model

The proposed model for single-image SR can be interpreted as a feed-forward neural network providing a highly fast and simple implementation. The objective of this proposed model is to find the network parameters to get the best prediction of the HR patches from the corresponding LR ones. The suggested network consists of the following parameters,

$$\Theta = \{\mathbf{D}_l, \mathbf{D}_h, \mathbf{S}_{hl}, \mathbf{b}_h, \mathbf{V}_{hl}\} \tag{31}$$

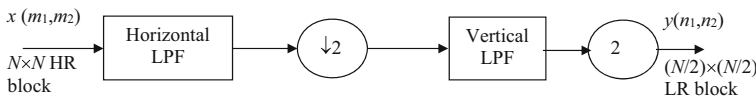


Fig. 2 Down-sampling process from the $N \times N$ HR block to the $N/2 \times N/2$ LR block



(a) The SR technique for scenario 1



(b) The SR technique for scenario 2



(c) The SR technique for scenario 3



(d) LR IR Image



(e) Regularized interpolation results



(f) Least square interpolation results

Fig. 3 Case 1 of Tree image, SNR = 25 dB. Scenario 1: Bicubic filtering followed by down-sampling by 2, Scenario 2: Bicubic filtering followed by down-sampling by 3, Scenario 3: Gaussian filtering of size 7×7 with standard deviation 1.6 followed by down-sampling by 3

We assume that the process of learning the model parameters is off-line, using a set of LR-HR image pairs. Patches are extracted from each image pair y_l, y_{hl} at the same locations, resulting in a training set consisting of N paired LR HR patches $\{P_l^Q, P_h^Q\}$.

The optimization problem formulates the training model parameters Θ

$$\underset{\Theta}{\text{Argmin}} \sum_{K=1}^N \|\mathbf{D}_h \left(\left[\Phi \left(\mathbf{b}_h + \mathbf{V}_h \mathbf{x}_l^Q \right) \right] \circ \left[\mathbf{S}_{hl} \left(\mathbf{D}_l \right)^T \mathbf{P}_l^Q \right] \right) - \mathbf{P}_h^Q\|_2^2 \quad (32)$$

The product above is the Hadamard product.



(a) The SR technique for scenario 1



(b) The SR technique for scenario 2



(c) The SR technique for scenario 3



(d) LR IR image

(e) Regularized
interpolation result(f) Least squares
interpolation result**Fig. 4** Case 2 of Plane image with SNR = 25 dB

To reduce the complexity of solving this joint optimization problem in order to allow learning of model parameters, an initial estimation of \mathbf{D}_l and \mathbf{D}_h dictionaries is set using directional PCAs [31] and K-SVD [1] as well-known approaches. Having the true sparsity patterns for each patch pair and given the \mathbf{D}_l and \mathbf{D}_h dictionaries estimates, we set an initial estimate for the covariance matrix \mathbf{S}_{hl} directly by solving an LS problem. After having the mentioned initial estimates, \mathbf{D}_h and \mathbf{S}_{hl} can be updated together to be well-tuned. Now, we reach the network innermost layer, where we update the restricted Boltzmann machine parameters \mathbf{V}_{hl} , \mathbf{b}_h , while the remaining parameters are kept fixed to the previous estimates. Finally, a last tuning of the \mathbf{D}_h dictionary takes place to enhance the prediction in terms of the HR patch error.

8 The proposed approach

This proposed approach is based on single-image super resolution applied to IR images.

The steps of the proposed approach can be summarized as follows:

1. Capture the LR IR image.

2. Apply bicubic interpolation with three scenarios on the LR image to generate the enhanced IR image.
3. Get pre-learned parameters by neural network.
4. Apply the estimation process for the HR image.
5. Extract overlapping patches of the HR image.
6. Compute the LR representation from the LR patches using Eq.(22).
7. Compute the LR sparsity pattern from the LR representation using Eq.(23).
8. Compute the MSE for the HR representation using Eq.(30).
9. Apply the recovery process for the HR patches.
10. Recover the LR-HR difference image from the patches.
11. Apply the recovery process for the HR image.



(a) The SR technique for scenario 1



(b) The SR technique for scenario 2



(c) The SR technique for scenario 3



(d) LR IR image



(e) Regularized
interpolation result



(f) Least squares
interpolation result

Fig. 5 Case 3 of Truck image with SNR = 25 dB

9 Simulation results

In this section, four IR images have been used to test the adaptive LS algorithm and the SR algorithm. Firstly, the model of image down-sampling given in Fig. 2 is applied to the original images to yield the LR images down-sampled by a factor of two in both directions.

The adaptive LS algorithm and the SR algorithm are then tested on the obtained LR images with $\text{SNR} = 25$ dB. The obtained results are given in Figs. 3, 4, 5 and 6. The values of the PSNR and the average number of iterations per block for the adaptive LS algorithm are given in the figures. The database properties of the IR images used are shown in Table 1. The PSNR results of the four cases for all scenarios are given in Table 2. It is clear that the results are good from the PSNR and computation time perspectives (Table 3).



(a) The SR technique for scenario 1



(b) The SR technique for scenario 2



(c) The SR technique for scenario 3



(d) LR IR image



(e) Regularized
interpolation result



(f) Least squares
interpolation result

Fig. 6 Case 4 of Man image with $\text{SNR} = 25$ dB

Table 1 Properties of test images

Image name	Image size
Truck	128 × 128
Plane	128 × 128
Tree	128 × 128
Man	128 × 128

10 Conclusion and future work

This paper investigated SR reconstruction for IR image enhancement with interpolation-based techniques, and learning-based techniques. A learning-based single-image SR technique was applied to IR images. The obtained results have shown good visual quality and superiority compared to those of other techniques. Simulation results revealed good quality of the obtained IR images. More enhancements could take place in the learning stage to provide better resolution IR images for further pattern recognition applications.

Table 2 Summary of the PSNR results

PSNR	Truck	Plane	Tree	Man
Regularized interpolation	14.63	17.05	17.03	16.01
LS interpolation	21.75	25.78	22.64	23.43
SR approach scenario 1	25.41	28.77	26.79	30.00
SR approach scenario 2	22.60	25.86	24.37	26.69
SR approach scenario 3	21.41	24.43	23.60	25.01
Recovery with sparsity scenario 2	23.42	27.10	24.68	28.66
Recovery with sparsity scenario 3	23.44	27.20	24.70	28.65
The proposed approach of sparsity scenario 1	26.98	30.57	27.76	32.85

Table 3 Summary of the computation time results

Computation Time	Truck	Plane	Tree	Man
Regularized interpolation	8.29	7.07	8.81	8.17
LS interpolation	8.64	5.47	21.55	4.99
SR approach scenario 1	4.49	0.0054	9.81	0.025
SR approach scenario 2	8.07	0.0055	18.85	0.0147
SR approach scenario 3	8.68	0.0058	22.15	0.0149
Recovery with sparsity scenario 2	7.56	7.69	18.51	103.97
Recovery with sparsity scenario 3	8.92	9.10	22.67	119.66
The proposed approach of sparsity scenario 1	4.20	4.08	10.11	51.14

References

1. Aharon M, Elad M, Bruckstein AM (2006) K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans Signal Processing* 54(11):4311–4322
2. Armstrong GR, Packard PD 1996 CMT and PtSi FLIR systems for EUCLID RTP 8.1, in: *Proc. SPIE*, 257–266
3. Ashiba HI, Awadalla KH, El-Halfawy SM, Abd El-Samie FE (2011) Adaptive Least Squares Interpolation of Infrared Images. *Springer, Journal of Circuits, Systems and Signal Processing* 30: 543–551
4. Bahy RM, Salama GI, Mahmoud TA (2014) Adaptive regularization based super resolution reconstruction technique for multi-focus low resolution images, *Signal Process.* 155–167
5. Baker S, Kannate T (2002) Limits on super resolution and how to break them. *IEEE Trans Pattern Anal Mach Intell* 24(9):1167–1183
6. Chen T, Wu HR, Qiu B (2001) Image Interpolation Using Across-Scale Pixel Correlation, *Proceedings of ICASSP*
7. Donoho DL (2006) Compressed sensing. *IEEE Transactions on Information Theory* 52:1289–1306
8. El-Khamy SE, Hadhoud MM, Dessouky MI, Salam BM, Abd El-Samie FE (2006) A new approach for regularized image interpolation. *J Braz Comput Soc* 11(3):65–79
9. Fattal R (2007) Image upsampling via imposed edge statistics, *ACM Transactions on Graphics (TOG)*, vol. 26(3), ACM
10. Fortin J, Chevrette P (1996) Realization of a fast micro-scanning device for infrared focal plane arrays, in: *Proc. SPIE* 2743, pp. 185
11. Freeman WT, Pasztor EC, Carmichael OT (2000) Learning low-level vision. *Int J Comput Vis* 40(1):25–47
12. Han JK, Kim HM (2001) Modified Cubic Convolution Scaler with Minimum Loss of Information. *Opt Eng* 40(4):540–546
13. Hou HS, Andrews HC (1978) Cubic Spline For Image Interpolation and Digital Filtering. *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. ASSP-26, 9:508–517
14. Huang J, Singh A, Ahuja N (2015) Single Image Super-resolution from Transformed Self-Exemplars. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5197–5206
15. Keys R (1981) Cubic convolution interpolation for digital image processing. *Acoustics, Speech and Signal Processing, IEEE Transactions on* 29(6):1153–1160
16. Liu Y, Nie L, Liu L, Rosenblum DS (2015) From action to activity: Sensor-based activity recognition. *Neurocomputing*
17. Liu Y, Nie L, Han L, Zhang L, Rosenblum DS (2015) Action2Activity: Recognizing Complex Activities from Sensor Data. *IJCAI*, PP.1617–1623
18. Liu Y, Zhang L, Nie L, Yan Y, Rosenblum DS (2016) Fortune Teller: Predicting Your Career Path The 30th AAAI Conference on Artificial Intelligence, PP.201–207
19. Mallat S, Yu G (2010) Super-resolution with sparse mixing estimators. *IEEE Trans Image Process* 19(11): 2889–2900
20. Mao Y, Wang Y, Zhou J, Jia H (2016) An infrared image super-resolution reconstruction method based on compressive sensing. *Infrared Phys Technol* 76:735–739
21. Peleg T, Elad M (2014) A Statistical Prediction Model Based on Sparse Representations for Single Image Super-Resolution. *IEEE Trans Image Process* 23:2569–2582
22. Shin JH, Jung JH, Paik JK (1998) Regularized Iterative Image Interpolation And Its Application To Spatially Scalable Coding. *IEEE Trans Consumer Electronics* 44(3):1042–1047 August
23. Sun J, Zhu J, Tappen MF (2010) Context-constrained hallucination for image super-resolution, in *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 1–8
24. Thevenaz P, Blu T, Unser M (2000) Interpolation Revisited, *IEEE Trans. Medical Imaging*, vol.19, 739–758
25. Tian J, Ma KK (2010) Stochastic super-resolution image reconstruction, *J. Vis. Commun. Image Represent.* 232–244
26. Unser M (1999) Splines A Perfect Fit For Signal and Image Processing, *IEEE Signal Processing Magazine*
27. Ur H, Gross D (1992) Improved resolution from sub-pixel shifted pictures, *CVGIP: Graph. Models Image Process.* 54: 181–186
28. Wang Z, Liu D, Yang J, Han W, Huang T (2015) Deep Networks for Image Super-Resolution with Sparse Prior, *IEEE International Conference on Computer Vision (ICCV)*, 370–378
29. Yang J, Wright J, Huang T, Ma Y (2010) Image super-resolution via sparse representation. *IEEE Trans Image Process* 19:2861–2873

30. Yang X, Wu W, Liu K, Zhou K, Yan B (2016) Fast multisensor infrared image super-resolution scheme with multiple regression models. *J Syst Archit* 64:11–25
31. Yu G, Sapiro G, Mallat S (2012) Solving inverse problems with piecewise linear estimators: from Gaussian mixture models to structured sparsity. *IEEE Trans Image Processing* 21(5):2481–2499
32. Yue L, Shen H, Li J, Yuan Q, Zhang H, Zhang L (2016) Image super-resolution: The techniques, applications, and future. *Signal Processing*. 128: 389–408
33. Zeyde R, Elad M, Protter, 2012 single image scale-up using Sparse-representations, *Curves and Surfaces*, 711–730
34. Zhang H, Zhang Y, Li H (2012) Generative Bayesian image super resolution with natural image prior. *IEEE Trans Image Process* 21(9):4054–4067
35. Zhang K, Tao D, Gao X, Li X, Xiong Z (2015) Learning multiple linear mappings for efficient single image super-resolution. *IEEE Trans Image Process* 24:846–861
36. Zhao Y, Chen Q, Sui X, Gu G (2015) A novel infrared image super-resolution method based on sparse representation. *Infrared Phys Technol* 71:506–513
37. Zhao Y, Sui X, Chen Q, Wu S (2016) Learning-based compressed sensing for infrared image super resolution. *Infrared Phys Technol* 76:139–147
38. Zhu Y, Zhang Y, Yuille AL (2014) Single Image Super-resolution using Deformable Patches . *Proc.IEEE Conf. Comput. Vision and Pattern Recognition*, 1–8

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Fathi E. Abd El-Samie received the B.Sc. (Honors), M.Sc., and Ph.D. degrees from Menoufia University, Menouf, Egypt, in 1998, 2001, and 2005, respectively. Since 2005, he has been a Teaching Staff Member with the Department of Electronics and Electrical Communications, Faculty of Electronic Engineering, Menoufia University. He worked as a researcher at KACST-TIC in Radio Frequency and Photonics for the e-Society (RFTONICs) from 2013 to 2015. His current research interests include image enhancement, image restoration, image interpolation, super-resolution reconstruction of images, data hiding, multimedia communications, medical image processing, optical signal processing, and digital communications. He was the recipient of the Most Cited Paper Award from the *Digital Signal Processing* journal in 2008.



Huda I. Ashiba received the B.Sc., and M.Sc., from the Faculty of Electronic Engineering, Menoufia University, Menouf, Egypt. She received the Ph.D. from the Faculty of Engineering in Shoubra, Benha University, 2017. She joined the teaching staff of the Department of Electronics and Electrical Communications, Bilbis Higher Institute of Engineering Bilbis, Sharqia, Egypt. She published 8 papers. Her current research area of interest includes image processing, image interpolation, super-resolution reconstruction of images, multimedia communications, and digital communications



H. Shendy B.Sc. of Electronics and Electrical Communication Engineering. He was born in Benha, Qalubia, Egypt in 1988. He received the B.Sc. degree from the Faculty of Electronic Engineering, Menoufia University, Egypt in 2010. He is a former Systems Engineer in DOCOMO InterTouch and he is now a Communication Engineer in Suez Canal Authority, Egypt. His research interests cover the fields of image processing, compression techniques and big data processing.



Hala M. Mansour received the B.Sc., M.Sc. and Ph.D. degrees from the Faculty of Engineering, [Cairo University](#), Egypt, in 1978, 1984, and 1991, respectively. She joined the teaching staff of the Department of Electrical Communications, Faculty of Engineering in Shoubra, Benha University, Egypt, in 1991. She has published more than 50 scientific papers in national and international conference proceedings and journals. Her current research areas of interest include signal and image processing and mobile communications.



Hossameldin M. Ahmed received the B.Sc., M.Sc. and Ph.D. degrees from the Faculty of Engineering, Shoubra, Benha University, Egypt. He joined the teaching staff of the Department of Electrical Communications, Faculty of Engineering, Shoubra, Benha University, Egypt, in 2009. He has published more than 10 scientific papers in national and international conference proceedings and journals. His current research areas of interest include signal and image processing and mobile communications.



Taha E. Taha was born in Tanta, Egypt, on October 11, 1946. He received the B.Sc. degree (with distinction) in communications engineering from Menoufia University, Egypt, in June 1969, the M.Sc. degree in communication engineering from Helwan University, Egypt, in April 1978, and the Ph.D. degree (very honorable) in electronic engineering from the National Polytechnic Institute, Toulouse, France, in June 1985. From September 1969 to July 1978, he was a demonstrator. In July 1978, he was an Assistant Lecturer. In November 1985, he was a Lecturer. In February 1990, he was an Assistant Professor. In September 1995, he was named Professor in the Faculty of Electronic Engineering, Menoufia University, Communications Department. He was appointed as Vice Dean from February 2002 to October 2005, and Head of the Communications Department, from November 2005 to July 2007. At present, he is a Professor Emeritus at the same department. His main research interests are surface acoustic wave devices, optical devices, superconductor devices, medical applications of ultrasound, medical imaging, and bio-informatics.



Moawad I. Dessouky received the B.Sc. (Honors) and M.Sc. degrees from the Faculty of Electronic Engineering, Menoufia University, Menouf, Egypt, in 1976 and 1981, respectively, and the Ph.D. from McMaster University, Canada, in 1986. He joined the teaching staff of the Department of Electronics and Electrical Communications, Faculty of Electronic Engineering, Menoufia University, Menouf, Egypt, in 1986. He has published more than 200 scientific papers in national and international conference proceedings and journals. He is currently a Professor Emeritus at the Faculty of Electronic Engineering, Menoufia University. He has received the most cited paper award from Digital Signal Processing journal for 2008. His current research areas of interest include spectral estimation techniques, image enhancement, image restoration, super-resolution reconstruction of images, satellite communications, and optical communications.



Mohamed F. Elkordy received the B.Sc. (Honors) M.Sc. and Ph. D degrees from the Faculty of Electronic Engineering, Menoufia University, Menouf, Egypt, in 1979, 1985, and 1991. He joined the teaching staff of the Department of Electronics and Electrical Communications, Faculty of Electronic Engineering, Menoufia University, Menouf, Egypt, in 1981. He has published more than 74 scientific papers in national and international conference proceedings and journals. He died in 2018.



Mohammed Abd-Elnaby received the B.S. (Hons.), M.S., and Ph.D. degrees in electronic engineering from Menoufia University, Menouf, Egypt in 2000, 2004 and 2010, respectively. Currently, he is working as an Associative Professor at the Department of Electronics and Electrical Communications, Faculty of Electronic Engineering, Menoufia University, Menouf, Egypt. His research interests include wireless networks, wireless resource management, MAC protocols, cognitive radio, and cooperative communications.



Adel S. El-Fishawy received the B.Sc in electronics and communications engineering, Menoufia university, 1974, and M.Sc. degree from Cairo University, Egypt, 1982, respectively. He received the Ph.D degree in signals and systems from Drexel University, ECE Dept., Phila., PA, USA, 1990. He has published several scientific papers in national, international conferences and journals. He is working as a tenured professor with the Department of Electronics and Electrical Communications, Faculty of Electronic Engineering, Menoufia University, Egypt. His areas of interest are electronics, modeling and simulations, digital communications, signal processing, speech processing, biomedical image processing, and digital communications.

Affiliations

Fathi E. Abd El-Samie¹ • Huda I. Ashiba² • H. Shendy¹ • Hala M. Mansour³ • Hossameldin M. Ahmed³ • Taha E. Taha¹ • Moawad I. Dessouky¹ • Mohamed F. Elkordy¹ • Mohammed Abd-Elnaby^{1,4} • Adel S. El-Fishawy¹

¹ Department of Electronics and Electrical Communications Engineering, Faculty of Electronic Engineering, Menoufia University, Menouf 32952, Egypt

² Department of Electronics and Electrical Communications, Bilbis Higher Institute of Engineering, Bilbis, Sharqia, Egypt

³ Department of Electronics and Electrical Communications, Shoubra Faculty of Engineering, Benha University, Benha, Egypt

⁴ Department of Computer Engineering, College of Computers and Information Technology, Taif University, Al-Hawiya 21974, Saudi Arabia