http://dx.doi.org/10.21611/qirt.2015.0092

An Efficient Super-resolution Algorithm for IR Thermal Images Based On Sparse Representation

by W.Jino Hans* and N.Venkateswaran*

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

Abstract

In this paper, an efficient learning strategy to super-resolve the down-sampled IR thermal image is presented. Though the resolution offered by state-of-the-art non-coolant based Focal Path Array (FPA) IR thermal imaging device is significantly high, the images captured by these devices has to be down-scaled for storing and transmitting it over a network. IR thermal images captured by non-coolant based FPA thermal imaging device is down-scaled by a scale factor s and is represented as LR image. It is up-scaled by a simple interpolator such as bilinear interpolator to the desired magnification size to form the pseudo-HR image. Image patches are extracted from same location (i, j), from both LR and pseudo-HR image to form the self-example patch-pairs. Image patches which carry HF details are efficiently selected based on a threshold on sample maximum mean square error (SMMSE). Two effective dictionaries (LR and HR) constructed from patch-pairs are trained with state-of-the-art K-SVD algorithm. It is used within the sparse representation framework to reconstruct the up-scaled IR thermal Image. Quantitative and qualitative results claim that the proposed method super-resolve the existing LR images without introducing false HF details.

1. Introduction

The recent advances in sophisticated non coolant based Focal Path Array (FPA) IR thermal imaging devices has gained popularity in diversified applications such as medical, industrial and remote sensing [1]. Though the resolution offered by state-of-the-art thermal IR imaging device is significantly high, there are still limitations which has to be addressed. Also, IR thermal images captured by these devices has to be down-scaled for storing and transmitting it over a network. For further analysis and interpretations, such down-scaled IR images require efficient super-resolution (SR) algorithms to up-scale without any counterfeit details. It is vital to visualize sharp thermal discontinuities in thermal IR images for efficient diagnosis in various field of medical and industrial applications. The growing demand in this field has gained popularity among the researchers in image processing community to post process IR thermal images to improve its resolution with efficient SR algorithm. SR algorithms aims at generating high resolution (HR) image from single or ensemble of low resolution (LR) images [2]. SR algorithms can be classified under two main categories: the classical SR from multiple images of the same scene with sub-pixel shifts [2] and SR from single image [3].

In the classical approach, multiple LR frames of the same scene are combined to from the HR image. This method is expensive as it requires multiple LR frames of the same scene with sub-pixel shifts. Single image SR algorithms can be classified under three main categories: interpolation based methods, reconstruction based methods and learning based methods. Interpolation based SR approach [4] uses simple interpolators such as bilinear or bicubic interpolator to upscale the existing LR image to the size of the HR image. Though this method is very simple to implement, it tend to produce artifacts such as ringing, jagging etc. in HF details.

Reconstruction based SR approach [5] handle single image SR problem as an inverse problem. Restoration of HF details in an HR image is severely ill-posed and requires sophisticated prior models for regularization. Many state-of-the-art prior models such as gradient profile, smooth edge prior and sparse prior model [6] have been proposed for efficient reconstruction of HR images Sparse representation model (SRM) efficiently models image patches as a sparse linear combination of a few atoms selected from an over-complete dictionary [6]. SRM is used extensively in solving many ill-posed image processing tasks such as image super-resolution (SR), image de-noising, image in-painting etc.

Learning based SR approach [7] predicts the HF details which are explicitly missed in the HR image. The HF details are learned by training LR-HR patch-pairs. Learning based methods can introduce counterfeit HF details, if the training samples are not relevant with the input LR image. In order to overcome this, many self-learning SR approaches which uses only the patch-pairs extracted from the observed LR image have been proposed [8]. Small image patches extracted from an image will be represented as lines, dots, arcs etc. and are called as image primitives. These image primitives will redundantly appear both within and across various scales of the image. Many approaches to find the correspondence between the LR-HR patch-pairs have been proposed. In recent years, learning based SR algorithm is executed by training two dictionaries (LR and HR) within the framework of sparse representation.

In this current work, an efficient learning based SR algorithm to up-scale the LR thermal IR image which learns the in-similarity within the self-example patch-pairs extracted from the test IR thermal is presented. IR thermal images captured by non-coolant based FPA thermal imaging device is down-scaled by a scale factor s and is represented as LR image. It is up-scaled by a simple interpolator such as bilinear interpolator to the desired magnification size to form the pseudo-HR image. Image patches are extracted from same location (i, j), from both LR and pseudo-HR image to form the self-example patch-pairs. An efficient sampling strategy based on sample maximum mean square error (SMMSE) is

http://dx.doi.org/10.21611/qirt.2015.0092

employed to sample and select the self-example patch-pairs. Two effective dictionaries (LR and HR) are constructed from the patch-pairs attained by the proposed strategy and is trained with state-of-the-art K-SVD algorithm. It is used within the sparse representation framework to reconstruct the up-scaled IR thermal Image. Qualitative and quantitative evaluation with medical and industrial IR thermal images shows that the proposed method preserves sharp edges and is free from artifacts.

2. SR methodology based on sparse representation

Under mild conditions [9], theory of compressive sensing (CS) suggests that it is possible to recover HR patches by the sparse representations of its down-sampled LR patches. The sparse representations of all the patches extracted from the observed IR thermal image is found and its sparse co-efficient are used to learn the HF details that are explicitly missed in the LR image by self-example patch-pairs. In this subsection, sparse representations of image patches is addressed.

2.1.1. Self-example patch-pairs

Learning based SR algorithms rely on the training set to extract LR-HR patch-pairs. If the training set is not relevant to the target image, it will introduce counterfeit HF details. In order to avoid this, self-example patch-pairs are extracted from the observed LR image both within and across various scales of the image. Though natural images tend to appear very complex with sharp discontinuities, mathematically it can be observed that small image patches extracted from these image will have very simple structures such as points, lines etc. It is obvious that these simple structures will repeat itself redundantly both within and across various scales of the image. These simple patches are extracted from the same location of observed LR image and its up-scaled version and are called as self-example patch-pairs.

Let $Y_l \in \mathbb{R}^{n\times n}$ be the observed test LR image. It is magnified to the with a magnification factor s using simple interpolator and is denoted as $Y_h \in \mathbb{R}^{kn\times kn}$. Let the in-place self-similar patch-pairs extracted from Y_l and Y_h are p_l and p_h respectively. The self-similar patch-pairs are extracted to form the LR-HR patch-pairs for the learning based SR problem. For an image patch p_h extracted from an origin (i,j) of the pseudo-HR image Y_h, a self-similar patch p_l exist at the origin (i_k, j_k) in the observed LR image Y_l, where $i_k = [i/k + 0.5]$, and $j_k = [j/k + 0.5]$. The correspondence across the patch-pairs is learned effectively to form the LR-HR self-example patch pairs $\{p_l, p_h\}$. These self-example patch-pairs are used as a prior for the SR problem.

2.1.2. Self-example patch-pair sampling strategy

Learning-based SR algorithms are computationally expensive as it involves an exhaustive search over a large dictionary with all possible patches from the input observed LR test images. To reduce the computational complexity, it is essential to decrease the number of patches in the dictionary. It can be achieved by choosing the sample patches from the test image based on a threshold on sample maximum mean square error SMMSE [10]. Only those patches with higher SMSE are selected and the rest are discarded. Instead of searching all the test patch-pairs, it is enough to search merely the sample patches, thereby the computational complexity will be reduced. Recent results in image statistics has revealed that most of the images do have repeated structures. For very small patches, the redundancy is very frequent and henceforth many patches which carry similar details will exist. These redundant patches will not provide HF details and can be discarded. The main task is to eliminate the patches with less variance and to select those with high variance.

From the magnified LR thermal IR image $Y_h \in \mathbb{R}^{kn\times kn}$, K image patches are extracted based on raster scan method and are represented as column vector of size $(k \times 1)$. The sample maximum mean square error SMMSE for the patches are found using

$$SMMSE_{p_i} = \frac{\sum_{j=1}^{n^2}\left(p_{ij} - \sum_{j=1}^{n^2} p_{ij}/k^2\right)^2}{k^2-1} \quad (1)$$

In (1), p_i represents the i^{th} column vector and p_{ij} represnts the j^{th} row element of p_i. Many similar patches exists in the test image which repeatedly appear within it. These repetitive patches are eliminated by efficiently selecting only those patches which has higher variance are selected. This is accomplished by a threshold on SMMSE value. For Patches with smooth details the variance will be less and hence the SMMSE will be lesser. For patches which carry HF details, the SMMSE value tend to be higher. By finding the SMMSE value, the patches can be efficiently sampled and selected. IT reduces the number of atoms in the dictionary thereby the computational complexity is reduced. Only 10% of the patches will have high variance and the rest can be discarded for dictionary initialization.

2.1.3. Dictionary learning from self-example patch-pairs

In this section, an effective dictionary learning from self-example patch-pairs is proposed. It is evident from the theory sparse and redundant representations, it is possible to represent image patches as linear combination of few atoms from an over-complete [6] dictionary. The initial dual dictionaries (LR and HR) are constructed from the self-example patch-

http://dx.doi.org/10.21611/qirt.2015.0092

pairs. The LR and HR dictionaries D_l and D_h are initialized with sample patch-pairs $\{p_l^i\}_{i=1}^{L}$ and $\{p_h^i\}_{i=1}^{K}$ respectively. The dictionaries will have L atoms such that each atom corresponds to L patch-pairs extracted with the sampling strategy explained in section 2.1.2. The dictionaries are unit normalized and are trained with state-of-the-art dictionary learning algorithm K-Singular Value Decomposition (K-SVD) [11]. LR image patches are sparse represented using the trained dictionaries.

A small LR image patch $\wp_l \in \mathbb{R}^{n\times 1}$ taken from the test IR image can be well represented as sparse linear combination of few atoms from the trained dictionary such that,

$$\wp_l^k = D_l\alpha^k \quad \forall k, \;\; s.t \left\|\alpha^k\right\|_0 << L \tag{2}$$

In Eq. (), α represents the sparse co-efficient vector. Let us denote $\wp_l = [\wp_l^1, \wp_l^2 \dots \wp_l^P] \in \mathbb{R}^{n\times P}$ as a matrix of P LR patches arranged as column vector. Such that,

$$\wp_l = D_l\mathrm{A} \tag{3}$$

Where, $\mathrm{A} = [\alpha_1, \alpha_2, \dots \alpha_M] \in \mathbb{R}^{L\times M}$ denotes the sparse co-efficient vectors obtained by the sparse coding stage during the training process. The dictionary is trained such that,

$$D_l{}^* = \underset{D_l,\alpha^k}{\operatorname{argmin}} \|\mathrm{X}_l - D_l\mathrm{A}\|_F^2 \tag{4}$$

The minimization problem stated in Eq. (4) is solved using the K-SVD dictionary learning algorithm. The algorithm searches for the best sparse linear combination of atoms from the dictionary to sparse represent the LR image. Each and every atoms in the dictionary and their corresponding sparse co-efficient vector is updated at every stage such that the error will be minimized.

During the dictionary learning process, all the atoms will be updated one by one and the rest will remain unchanged. The l^{th} atom in the dictionary D_l is denoted as d^l. The sparse co-efficient vector corresponding to d^l is represented as α^l. The error contributed by the l^{th} atom is given by $E_l = \wp_l - \sum_{j\neq l} d^j\,\alpha^j$, Such that

$$\left\|\wp_l - \textstyle\sum_{j=1}^{L} d^j\,\alpha^l\right\|_F^2 = \left\|E_l - d^l\,\alpha^l\right\|_F^2 \tag{5}$$

The error contributed by l^{th} atom is reduced by updating the l^{th} atom and its corresponding sparse vector with the optimization step given as

$$d^l = \frac{E_l(\,\alpha^l)^T}{\left\|E_l(\,\alpha^l)^T\right\|_2} \tag{6}$$

And

$$\alpha^l = (\,d^l)^T E_l \tag{7}$$

The HR dictionary is trained by taking the inverse of the sparse co-efficient vector given by

$$\wp_h^k = D_h\alpha^k \quad \forall k, \;\; s.t \left\|\alpha^k\right\|_0 << L \tag{8}$$

The HR dictionary is approximated with X_h and A by

http://dx.doi.org/10.21611/qirt.2015.0092

$$D_h = \mathrm{X}_h \mathrm{A}^{\dagger} \tag{9}$$

The trained LR and HR dictionaries D_l and D_h are used to sparse represent the image patches extracted from the thermal IR images.

2.1.4. IR thermal image super-resolution based on sparse-representation

Given a down-sampled IR thermal image $Y_l \in \mathbb{R}^{n\times n}$, it is desired to super-resolve it with a magnification factor s, such that the super-resolved HR image is denoted as $X_h \in \mathbb{R}^{sn\times sn}$. It is accomplished by using the sparse-prior model to effectively sparse represent the LR patches with the trained dictionaries.

Let us assume that the observed downs-sampled IR thermal image $Y_l \in \mathbb{R}^{n\times n}$ be the degraded version of the unknown HR thermal image X_h , such that

$$Y_l = SX_h \tag{10}$$

Where $S = BH \in \mathbb{R}^{p\times q}$ is the degradation operator such that B and L represents the blurring and down-sampling operator respectively. The input LR image is up-scaled to the size of the target HR image by simple interpolation operator and is referred as pseudo-HR image given by

$$Y_h = QX_l \in \mathbb{R}^{sn\times sn} \tag{12}$$

The image patch extracted from Y_h is represented as $p_l \in \mathbb{R}^{p\times p}$. The corresponding HR image patch p_h is to be recovered from p_l. This problem is extremely ill-posed.

$$p_l = S.p_h \tag{13}$$

From theory of CS, it is evident to recover HR image patch from its LR observation within the frame-work of sparse prior model. The HR image patch p_h is represented as sparse linear combination of elements from a dictionary D_h

$$p_h = D_h\alpha \tag{14}$$

Where $\alpha \in \mathbb{R}^K$ is the coefficient vector,$\|\alpha\|_0 \leq L$, where L is the maximum allowed number of non-zero elements in the atom and K is the number of atoms in the dictionary. Substituting (14) in (13) we get

$$p_l = S.D_h\alpha = D_l\alpha \tag{15}$$

The HR image patch is recovered using the optimization problem stated in (16)

$$p_h = \underset{\|\alpha\|_0}{\operatorname{argmin}} \|p_h - D_h\alpha\|_F^2 \tag{16}$$

The dictionaries are constructed with the self-example patch-pairs selected based on SMMSE threshold and are trained using K-SVD algorithm. Within the framework of sparse representation, the HR image patches are recovered by solving the optimization problem stated in Eq. (16).

3. Result and Discussions

The proposed algorithm for single image SR based on sparse representation is used to super-resolve the IR thermal images thereby the quality of the existing IR thermal image is enhanced. The proposed method is validated by testing it on IR thermal images [12] obtained from Terravic Facial IR Database and Terravic Weapon IR Database. The images are captured with Raytheon L-3 Thermal-Eye 2000AS. The images collected from the database are up-scaled by a factor of $2\times$ and $4\times$. The LR and HR dictionary is formed from the self-example patch-pairs extracted from the test

http://dx.doi.org/10.21611/qirt.2015.0092

example IR image. The size of the patch is chosen as 3×3. The number of atoms in the dictionary is fixed as 1024. The proposed SR algorithm was applied to various IR images and its results are compared both qualitatively and quantitatively with recent state-of-the-art SR approaches.

3.1 Qualitative Analysis

The quality of SR algorithm is evaluated based on the sharpness of the image reconstructed. The reconstructed image should be free from artifacts. Three IR test images are up-scaled by $4\times$ by state-of-the-art approaches such as Yang's method and bi-cubic interpolation method.

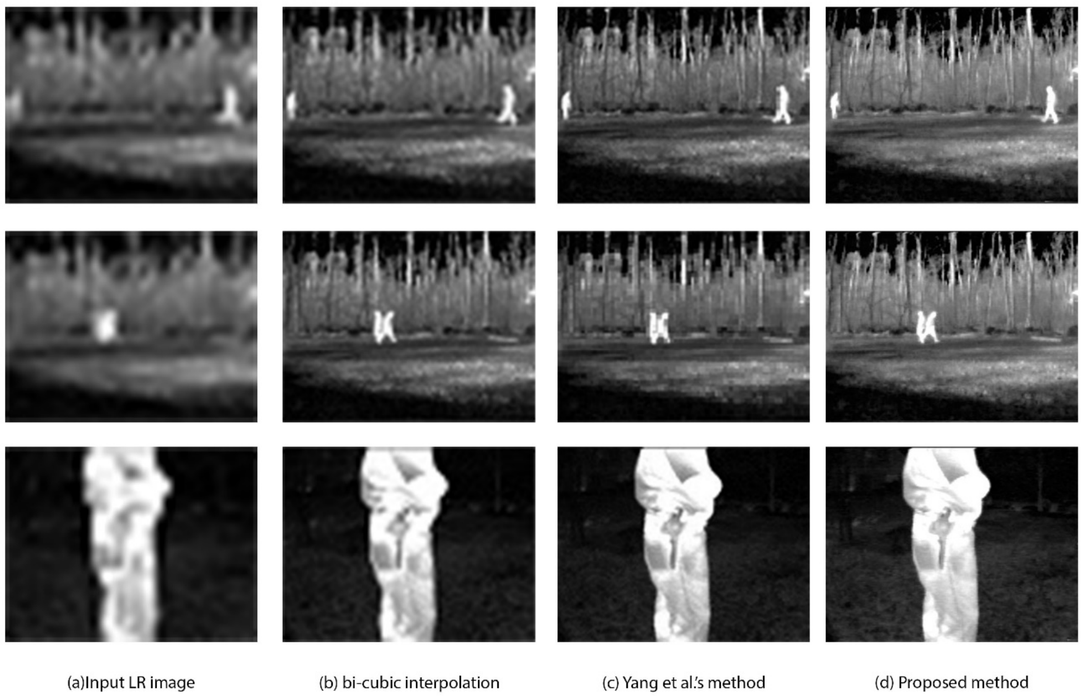

Fig.-1 *(a) Input IR thermal Image (b) Up-scaled by bicubic interpolation (c) Yang et al.'s method (d) Proposed method*

In Figure 1, the input LR image is down-sampled by a factor of 4. The input LR thermal IR image in figure 1-a shows three different input IR images. In the first image, 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 by bicubic interpolation method. The results in Figure 1-b can identify the persons but it does not give more information and are suffering from jaggy artifacts. Figure 1-c shows the image up-scaled by Yang et.al's method and figure 1-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 the second image, two persons are very close and it is very well characterized in the proposed method. Similarly, in the third image a person carrying weapon is shown. The input LR image is very much burred such that the weapon is not visible. The proposed method shows better details about the weapon and it can be easily traced.

3.2 Quantitative Analysis

To quantitatively measure the performance of the proposed method two quantitative measures such as peak signal to noise ratio (PSNR) and root mean square error (RMSE) and Structural similarity index measurement (SSIM) are used. The qualitative metrics for the three input images is summarized in Table-1. The proposed self-similarity based SR algorithm has the best PSNR values, as it combines the advantages of in-place example patches and their corresponding local self-similarity learned using the dictionary-based approach. The PSNR value is obtained by taking the ground truth information provided in the database. The PSNR results for the proposed method is higher than other existing methods. The SSIM value of the proposed method is similar to Yang et al.'s method and the overall performance of the proposed method is better than other state-of-the art approaches.

http://dx.doi.org/10.21611/qirt.2015.0092

Table 1. *RMSE, PSNR, and SSIM comparison for various state-of-the-art SR approaches*

Test Image	Bi-cubic Interpolation			Yang et al.'s method			Proposed method		
	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM
Test Image 1	6.30	32.14	0.982	5.85	32.81	0.992	5.70	33.02	0.992
Test Image 2	13.45	25.56	0.841	11.138	27.20	0.892	11.05	27.25	0.891
Test Image 3	10.20	27.95	0.892	7.48	30.65	0.912	7.33	30.82	0.910

4. Conclusion

In this paper, an efficient SR algorithm for IR thermal based on sparse representation is presented. Self-example patch-pairs are sampled and selected based on a threshold on sample mean square error (SMSE) from the given input LR image. The selected self-example patch-pairs are used to construct two (LR and HR) dictionaries. The dictionaries are trained with state-of-the-art K-SVD algorithm. Within the framework of sparse representation, the trained dictionaries are used to reconstruct the HR image effectively. Both qualitative and quantitative analysis claim that the proposed method can effectively super-resolve existing low resolution IR thermal images without introducing counterfeit artifacts.

REFERENCES

[1] [1] Jen Hong Tan, E. Y. K. Ng, U. Rajendra Acharya and Caroline Chee, Infrared thermography on ocular surface tempaerature: A review.

[2] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Process. Mag.*, vol. 20, no. 3, pp. 21–36, May 2003.

[3] . Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in Proc. of IEEE Int. Con. on Computer Vision, IEEE, Kyoto, Japan (2009).

[4] X. Li and M. T. Orchard, "New edge-directed interpolation," IEEE Trans. Image Process. 10(10), 1521–1527 (2001).

[5] S. Dai, M. Han, W. Xu, Y.Wu, and Y. Gong, "Soft edge smoothness prior for alpha channel super resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Classification*,Minneapolis,MN, USA, Jun. 2007, pp. 1–8.

[6] J. Yang et al., "Image super-resolution via sparse representation," IEEE Trans. Image Process. 19(11), 2861–2873 (2010).

[7] Y. Tang, L. Li, and X. Li, "Learning similarity with multikernel method," IEEE Trans. Syst., Man, Cybern. B, Cybern., vol. 41, no. 1, pp. 131–138, Feb. 2011.

[8] G. Freedman and R. Fattal, "Image and video upscaling from local self-examples," ACM Trans. Graph. 30(2), 1–10 (2011).

[9] N. Fan, "Super-resolution using regularized orthogonal matching pursuit based on compressed sensing theory in the wavelet domain," in Proc. IEEE Conf. Comput.Graph., Imag. Vis.,Aug. 2009, pp. 349–354.

[10] W. C. Ogle, H. E. Witzgall, M. A. Tinston, J. S. Goldstein, and P. A.Zulch, "Independent sample mean squared error for adaptive detectionstatistics," in Proc. IEEE Aerosp. Conf., BigSky,MT, USA,Mar. 2005,pp. 2222–2227.

[11] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," IEEE Trans. Signal Process., vol. 54, no. 11, pp. 4311–4322, Nov. 2006.

[12] 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.