

Single Image Super Resolution Using Sparse Representation with Image Fusion Principle

Eman A. Sultan¹, Said E. El-Khamy², EL-Sayed El-Rabie¹, Nawal A. El-Fishawy¹,
Fathi E. Abd El-Samie¹

¹Faculty of Electronic Engineering, Menoufia University, Menouf, Egypt

Eman_soltan@yahoo.com, srabaie@hotmail.com, nelfishawy@hotmail.com, fathi_sayed@yahoo.com

²Faculty of Engineering, Alexandria University, Alexandria, Egypt,

elkhamy@ieee.org

Abstract: A suggested single image super resolution algorithm is proposed based on image fusion principle. Magnetic resonance and computed tomography images are interpolated using two algorithms that use sparse-representation modeling with dictionary learning. The MR and CT images are fused either by discrete wavelet or curvelet transforms, then the fused result are interpolated by the same algorithms. Simulation results show that the fused super resolution image provides higher PSNR values than the original CT and MR images by using these interpolation algorithms. Also experimentally we deduce that using the curvelet fusion technique provides better results than using the wavelet, and scaling-up by one of the sparse representation algorithm gives more better results than bicubic and the other sparse representation algorithms in almost all images.

Keywords: about four key words separated by commas.

Introduction

In recent years, sparse representations have got a lot of attention from the signal processing community. This is due to the fact that an important variety of signals such as audio and natural images can be well approximated by a linear combination of a few elements (atoms) of some (often) redundant basis, usually called dictionaries [1]. Sparse modeling aims at learning these non parametric dictionaries from the data itself. Several algorithms have been developed for this task, e.g., the K-SVD [2] and Yang et. al. [3,4]. Super resolution refers to the process of recovering a high resolution image than what is afforded by the physical sensor through post processing, making use of one or more low resolution observations[5].

Single image super resolution SR means only one observed image is available. Single image super resolution (SR) has been studied in recent years due to its practical uses. Single image SR aims at synthesizing the high resolution (HR) version using the information collected from training image data or from the input low resolution (LR) image itself. These SR approaches are typically referred to as example or learning-based SR methods. In example-based SR approaches, for an input LR image a typical search for its similar patches from training LR image data, and the corresponding HR versions are used to synthesize the final SR output [6]. But the learning-based methods aims at modeling the relationship between the images with different resolutions by observing particular image priors, and these models are used to predict the SR output [4]. For example, Yang et al. [3] proposed to learn sparse image representation for SR. In this paper, Yang et al. [3] and Michael et al. [7] learning-based SR approaches will be discussed to be used in the proposed work.

For Yang et al. SR approach, they addressed the problem of generating a super resolution (SR) image from a single low resolution input image. The low resolution image is viewed as down sampled version of a high resolution image, whose patches are assumed to have a sparse representation with respect to an over complete dictionary of prototype signal atoms. That method does not require any learning on the high resolution patches, instead working directly with the low resolution training patches or their features. that approach is motivated by results that linear relationships among high resolution signals can be precisely recovered from their low dimensional projections[8].

For Michal et al. SR approach, they deals with the single image scale up problem using sparse representation modeling. Their goal is to recover an original image from its blurred and down scaled noisy version. Since that problem is highly ill posed, a prior is needed in order to regularize it. Previous work offers various ways to address this problem, ranging from simple linear space invariant interpolation schemes (e.g., bicubic interpolation), to spatially adaptive and non linear filters of various sorts. They build their work on a recently proposed successful algorithm by Yang et. al. [3,4].

The CT (Computing Tomography) image can provide dense structures like bones and implants with less distortion, but it cannot detect physiological changes, while the MR (Magnetic Resonance) image can provide normal and pathological soft tissues information, but it cannot support the bones information. In this case, only one kind of image may not be sufficient to provide accurate clinical requirements for the physicians. Therefore, the fusion of the multimodal medical images is necessary and it has become a promising and very challenging research area in recent years[10]. Image fusion is a very important topic for a wide variety of applications [11, 12]. It has applications in medical imaging, remote sensing and digital photographs.

The fusion of MR and CT images of the same region or organ would result in an integrated image of much more detail. Researchers have made some attempts for the fusion of MR and CT images. Most of these attempts are directed toward the application of the DWT (Discrete Wavelet Transform) for this purpose [12]. Due to the limited capability of the DWT to deal with images with curved structures, the application of the curvelet transform for this purpose gives better results [13]. In this paper all images (CT, MR, wavelet fused, and curvelet fused) will be scaled-up (super-resolution) by the two previous mentioned algorithms (Yang et al. and Michael et al.). Results will show that scaling-up the fused image gives more better results than scaling –up the CT and MR separately.

This paper is organized as follows: Section 2, describes both Michael et al. scale-up algorithm and Yang et al. algorithm, the main difference between the two algorithms will be indicated in this section. Section 3, reviews the wavelet fusion technique with concentration on its limitations and the curvelet fusion technique. In section 4, we present the proposed algorithm by scaling-up the fused MR and CT images, also scaling-up the MR and CT images separately. The experimental results are given in section 5, followed by the conclusions and the more relevant references.

1. Single Image Super Resolution With Sparse Representation

In image processing area, many applications, specially medical imaging, require resolution enhancement while minimizing visual artifacts. The single image super resolution problem can be formulated as follows: for a given input high resolution image $y_h \in \mathbb{R}^{N_h}$, represented as a vector of length N_h pixels and (S, H) are two parameters represent the decimation and blurring operators respectively, we can obtain a low resolution image through the following relation [7].

$$z_l = SHy_h + v \quad (1)$$

The parameter H will apply a known low pass filter to image and S performs decimation by an integer factor s. $z_l \in \mathbb{R}^{N_l}$ is the low resolution (LR) noisy version of original high resolution (HR) y_h image. And v is an additive white Gaussian noise.

For a given low resolution image z_l , the main purpose is to find the super resolute image $\tilde{y} \in \mathbb{R}^{N_h}$, which is close as possible to the original image that $\tilde{y}_h \cong y_h$. In [7] the author uses the sparse-land local model for scale-up problem. This model assumes that each image patch can be represented using a linear combination of few atoms from a dictionary. But for patch building each image patch can be generated by multiplying a dictionary by a sparse (mostly zero) vector of coefficients. This procedure was taken by [3, 4] in their algorithm.

2.1 The Sparse Land Model

The LR image z_l will be scaled-up by bicubic interpolation, where Q represents this interpolation operator. This scaling-up procedure is to avoid complexity caused by the different resolution between z_l and y_h . y_l denotes the scaled-up image by Q operator through the following relation [7]:-

$$y_l = Qz_l = Q(SHy_h + v) = QSHy_h + Qv = L^{all}y_h + \tilde{v} \quad (2)$$

The main required result is to process y_l to get \tilde{y}_h . The patches extracted from y_l image will be used to estimate the corresponding patches from y_h .

2.2 Single Image Scale-Up Algorithm as Proposed in [7]

Michael et al. algorithm will be described in the following two figures (1 and 2) as published in their paper [7], including the training phase and reconstruction phase. We will use this algorithm in our proposed work as it will be presented later in this paper.

After collection the HR images and the corresponding LR images, the next step is pre-processing by using high pass filters just as used in [3,4]. Similar to the approach in [7], the desired pre-processing is applied on the full images not to small-patches, and only then are the patches extracted.

For HR image pre-processing, the low frequency will be removed by computing the difference images $e_h^j = y_h^j - y_l^j$. The authors in [7] applied this subtraction to focus the training on characterizing the relation between the low resolution patches and the edges and texture content within the corresponding high resolution ones. For the pre-processing of the LR images, these LR images are filtered using R high pass filters, in order to extract local features that correspond to their high frequency content.

In training phase several high resolution examples $\{y_h^j\}_j$ are collected. Each one of these images is blurred and down sampled by a factor s leading to $\{z_l^j\}_j$. This blurred and down sampled images are then interpolated back to the original size using Q, resulting $\{y_l^j\}_j$.

1. Training set construction: A set of high resolution training images $\{y_h^j\}_j$ is collected, Low resolution images $\{y_l^j\}_j$ are constructed using scale-down operator L^{all} and pairs of matching patches that form the training database, $P = \{p_h^k, p_l^k\}_k$, are extracted.
2. Each of these patch-pairs undergoes a pre-processing stage that removes the low-frequencies from p_h^k and extracts features from p_l^k . Dimensionality reduction is applied on the features of the low resolution patches p_l^k , making the dictionary training step much faster.
3. A dictionary A_l is trained for the low resolution patches, such that they can be represented sparsely.
4. A corresponding dictionary A_h is constructed for the high-resolution patches, such that it matches the

Figure 1 Training Phase [7]

The training phase illustrated in figure 1 is used to build both the low resolution dictionary and the high resolution one. Following to the training phase it will be the reconstruction phase, which will be applied for any low resolution image to get the super resolution one as shown in figure 2. In our work we will implement two super resolution algorithms Michael et al.[7] to Yang et al. algorithm [3, 4], which are using the same training data base and there are

between them many similarities. There are main important points added by Michael et al. to Yang et al. algorithm, for farther details refer to reference [7].

1. Given a test low-resolution image z_1 to be scaled-up, it is interpolated to y_1 of the destination size, and all that it requires is a spatial non-linear filtering to sharpen it.
2. Pre-processed patches p_i^k are extracted from each location $k \in \Omega$, and then sparse-coded using the trained dictionary A_i .
3. The found representations $\{q^k\}$ are then used to recover the high-resolution patches by multiplying them with A_h .
4. The recovered high-resolution patches $\{p_h^k\}$ are finally merged by averaging in the overlap area to create the resulting image.

Figure 2 Reconstruction Phase [7]

2. Image Fusion

Many techniques for image fusion have been presented and a thorough overview of these methods can be viewed in [14]. In this paper we will concentrate on the wavelet and the curvelet fusion techniques.

3.1 Discrete Wavelet Fusion

In general, the basic idea of image fusion based on wavelet transform is to perform a multi resolution decomposition on each source image; the coefficients of both the low frequency band (approximation part) and high frequency bands (details parts) are then fused with a certain fusion rule. The widely used fusion rule is maximum selection scheme. This simple scheme just selects the largest absolute value of each wavelet coefficient at each location from the input images as the coefficient at the location in the fused image. After that, the fused image is obtained by performing the inverse DWT (IDWT) for the corresponding combined wavelet coefficients.

$$(x, y) = w^{-1}(\mathcal{O}(w(I_1(x, y)), w(I_2(x, y)))) \quad (3)$$

3.2 Curvelet Fusion

One major disadvantage of the DWT in image processing is that it gives a large number of coefficients in all scales corresponding to the edges of the image. So in order to exactly reconstruct the edges in an image many coefficients are required in order to exactly reconstruct the edges in an image. This makes the DWT inefficient for handling long curved edges. Recent approaches such as the ridgelet transform and the curvelet transform are more efficient in handling long linear and curvilinear singularities in an image [15].

In the curvelet transform, the AWT (Adaptive Wavelet Transform) is used instead of the DWT to decompose the image into different subbands, called the detail planes and the approximation plane, and each subband of the detail planes is then partitioned into small tiles. Then, the ridgelet transform is applied on each tile [15]. In this way, the image edges can be represented efficiently by the ridgelet transform, because the image edges will now be almost like small straight lines. IAWT

is used to recover the image after fusion. Thus, the curvelet transform is considered as an effectively extension of the ridgelet transform to detect curved edges.

3. Proposed Image Super Resolution With Image Fusion Principle

In our suggested work we will merge the scaling-up algorithms of [3, 4, 7] with image fusion principle as in [16] (our previous work). We will use the data of [3, 4] for dictionary learning and also it was used by [7]. For Michael et al. [7] we use their training algorithm K-SVD for low resolution dictionary training and also use the pseudo inverse for high resolution dictionary learning. Yang et al. dictionaries will be used as they use.

The following steps will discuss the algorithm details:-

1. The input MR image will be blurred and down-scaled by H and S operators respectively to get z_1 , and then scaled-up again by Q operator to have y_{h1} image as a start for the algorithm in [7]. But in [3, 4] algorithm z_1 is used as the low resolution image, they did not use the Q operator interpolation operator. Both algorithms will provides us the high resolution image \tilde{y}_{h1} which is close as possible to the original y_{h1} .
2. The input CT image will be blurred and down-scaled by H and S operators respectively, and then interpolated again by Q operator to have y_{h2} image as a start point in [7], but in [3, 4] algorithm z_1 is used as the low resolution image. Both algorithms will provides us the high resolution image \tilde{y}_{h2} which is close as possible to the original y_{h2} .
3. Fuse MR image and CT image by wavelet fusion technique to get y_h and scale-up the fused result by both [3, 4] and [7] as followed in step 1 and 2 to get \tilde{y}_h . For the DWT the maximum frequency rule is implemented.
4. Fuse image MR image and CT image by curvelet fusion technique to get, y_h and scale-up the fused result by both algorithms [3, 4] and [7] to get, \tilde{y}_h . For the curvelet fusion, small tiles of dimensions 12×12 , with two pixels of overlapping from each side, are utilized to approximate long curved lines by small straight lines and to avoid the edge effects. A high pass filtering step is used to sharpen the curvelet fused images.

5. Experimental Results

Our experiments are implemented in MATLAB, on Intel Core i3-370M at 2.4 GHZ with 4 GB RAM. We will implement the two algorithm [3, 4] and [7] in our work with the same data set for dictionary learning that was used by [3, 4]. In [7] the original HR image is blurred and scaled -down by bicubic interpolation by $s=3$ (scaling factor), then this image is returned to its original size by bicubic interpolation. Feature extraction from the low resolution image is done by using 4 filters that perform 1-st and 2-nd horizontal and vertical derivatives : $f_1 = [1, -1] = f_2^T$ and $f_3 = [1, -2, 1] = f_4^T$. Around 130,000 training patch-pairs (of size 9×9) are

collected and PCA is applied to reduce feature dimension to $n_1 = 30$. The dictionary training procedure applied 40 iteration of the K-SVD algorithm with $m=1,000$ atoms in the dictionary, and allocating $L=3$ atoms for each representation vector. The reconstruction algorithm is tested on four cases. Each case contains CT1, MR1, Wavelet fused_1, and curvelet fused_1 images. And the second set contains CT2 and MR2, wavelet fused_2, curvelet fused_2 images. The same thing will be repeated for case3 and case4. We will perform scaling-up to all these images by [3, 4] and [7] algorithms. The results will contain PSNR (Peak Signal to Noise Ratio) comparison between bicubic, Yang et al. [3, 4], and Michael et al. [7] as in table 1. This PSNR value will be computed by:-

$$PSNR = 10 \log_{10} \left(\frac{255^2 \cdot N}{\sum_i (y_i - \hat{y}_i)^2} \right) \quad (4)$$

Results in table 1 can be interpreted as follows:-

Scaling-up the fused CT and MR images, either it is fused by wavelet or curvelet fusion technique, provides higher PSNR values than scaling-up the CT or MR image separately. This because, the CT (Computing Tomography) image can provide dense structures like bones and implants with less distortion, but it cannot detect physiological changes, while the MR (Magnetic Resonance) image can provide normal and pathological soft tissues information, but it cannot support the bones information. In this case, only one kind of image may not be sufficient to provide accurate clinical requirements for the physicians. Therefore, the fusion of the multimodal medical images is necessary and it has become a promising. Using curvelet fusion technique provides better results than using the wavelet. Scaling-up by Michael et al. [7] gives more better results than bicubic and Yang et al. algorithms in almost all images. Figure 3 provides the visual comparison between the three different scaling-up algorithms ([3, 4], [7], and bicubic) with both fusion cases wavelet and curvelet.

6. Conclusion

In this paper we use the image fusion principle (either wavelet fusion or curvelet fusion) with single image scale-up sparse-representation algorithms. We scale-up magnetic resonance (MR) and computed tomography (CT) images by two sparse-representation modeling with dictionary learning algorithms [3, 4] and [7]. Then we fuse the MR and CT images either by wavelet or curvelet and scale-up the fused result by the same two algorithms. Yang et al. [3, 4] and Michael et al. [7] algorithms train two low-and high-resolution dictionaries using training images. Experimental results show that scaling-up the fused CT and MR images, either it is fused by wavelet or curvelet fusion technique, provides higher PSNR values than scaling-up the CT or MR image separately. Also experimentally we deduce that using curvelet fusion technique provides better results than using the wavelet, and scaling-up by Michael et al. [7] gives more better results than bicubic and Yang et al. algorithms in almost all images.

References

[1] D. L. Donoho A. M. Bruckstein and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Rev.*, vol. 51, no. 1, pp. 34–81, 2009.

[2] 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.

[3] Yang, J., Wright, J., Huang, T., Ma, Y.: Image super-resolution as sparse representation of raw image patches. In: *IEEE Computer Vision and Pattern Recognition (CVPR)* (June 2008).

[4] Yang, J., Wright, J., Huang, T., Ma, Y.: Image super-resolution via sparse representation *IEEE ICIP*, 2010.

[5] C. V. Jiji, M. V. Joshi, and S. Chaudhuri, "Single-frame image super-resolution using learned wavelet coefficients," *International Journal of Imaging Systems and Technology*, vol.14, no. 3, pp. 105–112, 2004.

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

[7] R. Zeyde, M. Elad, and M. Protter, "On single image scale-up using sparse-representations," *curves and surfaces 2011*, LNCS 6920, pp. 711-730, 2011, Springer-Verlag Berlin Heidelberg 2011.

[8] E. Candes. *Compressive sensing*. Proceeding of International Congress of Mathematicians, 2006.

[9] F. Maes, D. Vandermeulen, and P. Suetens, "Medical image registration using mutual information," *Proceedings of the IEEE*, vol. 91, no. 10, pp. 1699–1721, 2003.

[10] V. Barra and J.-Y. Boire, "A general framework for the fusion of anatomical and functional medical images," *NeuroImage*, vol. 13, no. 3, pp. 410–424, 2001.

[11] Wang, H. J. Sun, and Y. Y. Guan, "The application of wavelet transform to multi-modality medical image fusion," in *IEEE International Conference on Networking, Sensing and Control (IEEE, 2006)*, pp. 270–274.

[12] H. Moustafa and S. Rehan, "Applying image fusion techniques for the detection of hepatic lesions and acute intra-cerebral hemorrhage," in *ITI 4th International Conference on Information and Communications Technology 2006 (IEEE, 2006)*.

[13] F. E. Ali, I. M. El-Dokany, A. A. Saad, and F. E. Abd El-Samie, "Curvelet fusion of MR and CT images," *Prog. Electromagn. Res. C* 3, 215–224 (2008).

[14] Y. Wang and B. Lohmann, "Multisensor image fusion: concept, method and applications," *Tech. Rep.*, Institute of Automatic Technology, University of Bremen, Bremen, Germany, 2000.

[15] J. L. Starck, E. Candes, and D. L. Donoho, "The curvelet transform for image denoising," *IEEE Trans. Image Process.* 11, 670–684 (2002).

[16] E. Sultan, S. Rabie, N. EL-Fishawy, S. E. El-Khamy, F. E. Abd El-Samie, "Image fusion based on wavelet-Based image interpolation with a least squares algorithm", *International Journal of Computer Applications* (0975 – 8887) Volume 55– No.13, October 2012

Table 1 The PSNR Values of Different Scaling-Up Algorithms

Images		Bicubic	Yang et al. [3, 4]	Michael et al. [11]
CT1 image	Case1	24.1	25.8	26.2
MR1 image		23.9	25.5	25.4
Wavelet fused_1 image		20.6	21.6	21.6
Curvelet fused_1 image		24.4	26.4	26.7
CT2 image	Case2	28.8	30.6	30.5
MR2 image		27	28.1	28.5
Wavelet fused_2 image		24.8	25.5	25.6
Curvelet fused_2 image		30.3	31.3	32.3
CT3 image	Case3	23.7	25.4	25.7
MR3 image		24.3	25.4	25.5
Wavelet fused_3 image		21.0	22.1	22.2
Curvelet fused_3 image		25.5	27.0	28.1
CT4 image	Case4	32.2	35.5	36.8
MR4 image		34.2	35.2	36.8
Wavelet fused_4 image		28.2	29.8	30.4
Curvelet fused_4 image		30.8	32.4	33.6

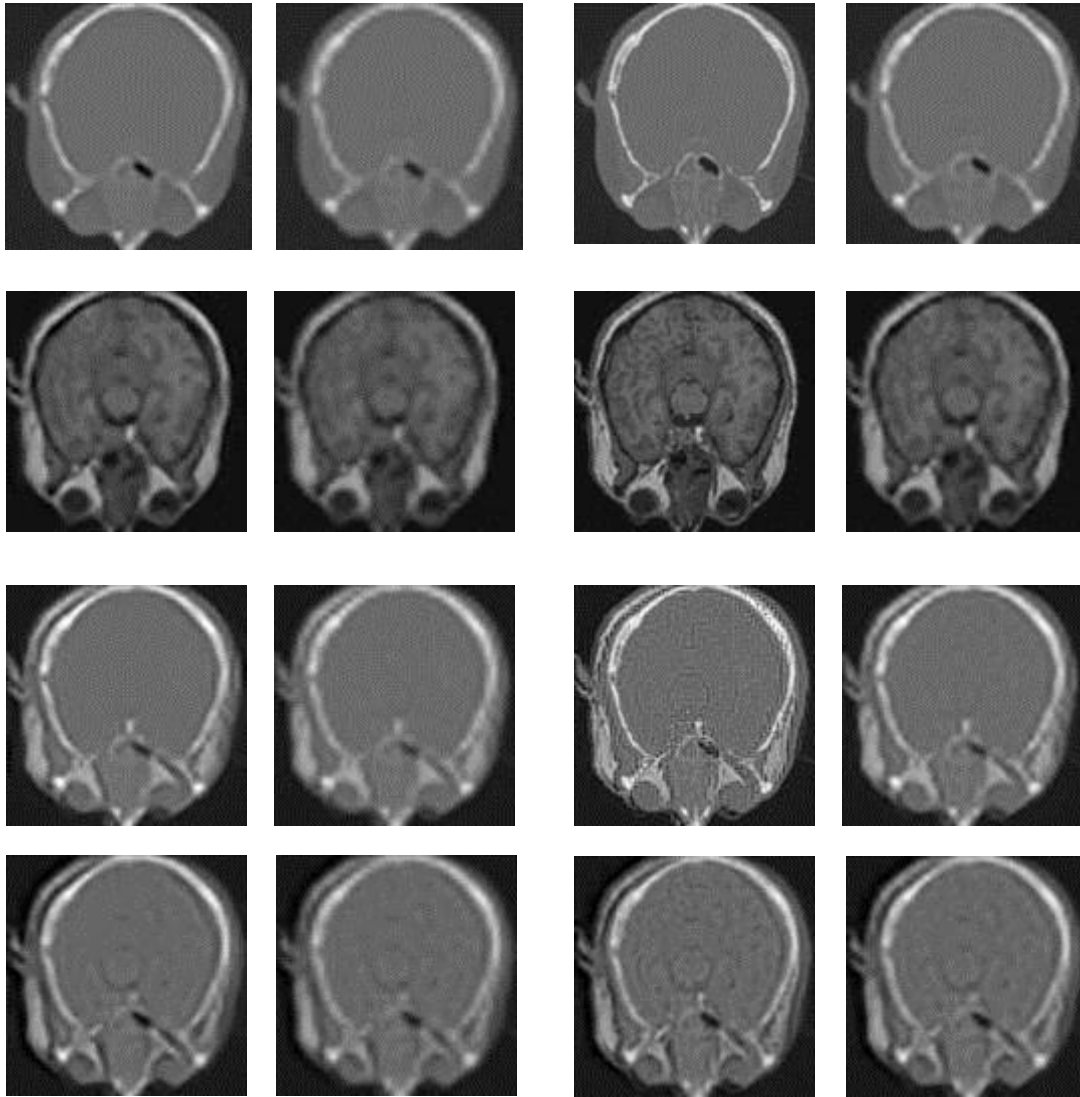


Fig. 3 Visual Comparison : (from top to bottom CT2, MR2, wavelet fused_2, curvelet fused_2 left to right : the original images (including the fused ones), bicubic interpolation, Yang et. al. [3, 4], and Michael et. al. Original Image Bilinear Yang et. All Michael et. all