

Super-resolution for Medical Image *via* Sparse Representation and Adaptive M-estimator

Q Xie^{1,2}, N Sang¹

ABSTRACT

Objective: The goal of super-resolution is to generate high-resolution images from low-resolution input images.

Methods: In this paper, a combined method based on sparse signal representation and adaptive M-estimator is proposed for single-image super-resolution. With the sparse signal representation, the correlation between the sparse representation of high-resolution patches and that of low-resolution patches for the identical image is learned as a set of joint dictionaries and a set of high-resolution patches is obtained for high- and low-resolution patches. Then the dictionaries and high-resolution patches are used to produce the high-resolution image for a low-resolution single image.

Results: At the post-processing phase, the adaptive M-estimator, combining the advantages of traditional L_1 and L_2 norms, is used to give further processing for the resultant high-resolution image, to reduce the artefact by learning and reconstitution, and improve the performance.

Conclusion: Three experimental results show the performance improvement of the proposed algorithm over other methods.

Keywords: Adaptive M-estimator, image dictionary, image super-resolution, learning-based method, sparse representation

Súper resolución de la imagen médica mediante escasa representación y estimador-M adaptativo

Q Xie^{1,2}, N Sang¹

RESUMEN

Objetivo: El objetivo de la súper resolución es generar imágenes de alta resolución a partir de imágenes de input de baja resolución.

Métodos: Este trabajo propone un método combinado basado en la representación escasa de las señales y el estimador-M adaptativo a fin de obtener una imagen de súper resolución. Con la representación de señal escasa, la correlación entre la escasa representación de parches de alta resolución y la de parches de baja resolución para la imagen idéntica, es aprendida como un conjunto de diccionarios asociados, y se obtiene un conjunto de parches de alta resolución para parches de alta y baja resolución. Entonces los diccionarios y parches de alta resolución se utilizan para producir la imagen de alta resolución para sustituir una imagen de baja resolución.

Resultados: En la fase de post-procesamiento, el estimador-M adaptativo, que combina las ventajas de las normas tradicionales L_1 y L_2 , se utiliza para dar mayor procesamiento a la imagen de alta resolución resultante, para reducir los artefactos mediante aprendizaje y reconstitución, y mejorar el rendimiento.

Conclusión: Tres resultados experimentales muestran la mejora del rendimiento del algoritmo propuesto por encima de otros métodos.

Palabras claves: estimador-M adaptativo, diccionario de imágenes, súper resolución de imagen, método basado en el aprendizaje, escasa representación

West Indian Med J 2016; 65 (2): 271

From:¹Institute for Pattern Recognition and Artificial Intelligence, Huazhong University of Science and Technology, Wuhan 430000, China and ²College of Biomedical Engineering, South Central University of Nationalities, Wuhan 430000, China.

Correspondence: Q Xie, Institute for Pattern Recognition and Artificial Intelligence, Huazhong University of Science and Technology, 1037 Luoyu Road, Hongshan District, Wuhan 430074, China, Fax: +86 27 67843220; e-mail: qinlanxie@126.com

INTRODUCTION

The goal of single-image super-resolution (SR) is often to enlarge a single image while preserving the features of the original image, such as sharp edge and texture. Many approaches have been proposed and studied for this task (1–3).

The conventional ways to enlarge a single-image is based on the scaling-up schemes such as interpolation. Simple interpolation methods often result in an over smoothing of the image (4), with some artifacts of ring at the edges. For overcoming these disadvantages, prior knowledge can be used to obtain more satisfying results (5, 6). While these methods are superior in preserving image edges, they are not enough to simulate the visual complexity of the true image. Another category of single-image SR approach is based on learning techniques (1, 7), which attempt to improve the performance of SR by exploiting the prior between low-resolution and high-resolution image patches. The literature (1) proposed an example-based learning strategy. However, the methods typically require enormous databases of millions of high-resolution and low-resolution patch pairs and are therefore computationally intensive. Kim and Kwon (7) propose the strategy of local linear embedding from manifold learning. The algorithm maps the local geometry of the low-resolution patch space to the high-resolution one, generating high-resolution patch as a linear combination of neighbours. Using this strategy, more patch patterns can be represented using a smaller training database. However, using a fixed number K neighbours for reconstruction often result in blurring effects, due to over-or under-fitting.

A method based on the learning method and sparse representation of signal was recently proposed to implement the single-image SR (8–10). The statistical properties of image and sparse representation of signal imply that the linearity relation between high-resolution can be recovered based on their low-resolution projection. Dissimilar to the example-based method (1), the algorithm exploits a certain sparse representation of image patch pairs sampled from the high- and low-resolution image, instead of directly using these image patch pairs, where this sparse representation is obtained by learning-based method from these image patch pairs. Having no direct use for the abundant patch-pairs, the algorithm avoids complicated computation.

This method is relatively simple, and yet it produces a substantial improvement over the conventional interpolation scheme. However, the method based on learning and sparse representation of signal has some space for improvement of the consequent output quality due to a prior of image itself having not been fully utilized.

Various further improvements have been considered to enhance the quality of the scaled-up image. One type of improvement is to migrate certain statistical methods used for multi-frame image super-resolution to improve the quality of the scaled-up image, such as back-projection algorithm, Markov random field (9, 11). The outcome of the method can effectively depress the artifacts due to the over-learning and

mismatching, but it still uses the low-resolution image observation only, instead of the information of high-resolution image.

Using the method of image reconstruction is another option. While the post-processing procedure can be considered as reconstructing the scaled-up image obtained with the learning method and sparse-representation to produce the original high-resolution image, here, the M-estimator is considered as a reasonable post-processing tool.

M-estimator is widely used to solve ill-posed problems (12–16). For SR, the M-estimator based reconstruction method tries to recover the original HR image by minimizing an objective function consisting of a fidelity term and a regularization term. The M-estimator often minimizes two norms for fidelity term and regularization term, respectively (13–16). Farsiu *et al* proposed to use the L_1 error norm (12) and showed that L_1 norm is more robust to large errors that may occur due to intensity outliers. Although L_1 norm-based estimator increases the robustness against outliers, it causes the minimizing function to be unconvex. Therefore, the steepest descent method and conjugate gradient algorithm cannot guarantee the convergence of its solution, and cannot guarantee accurately solving the problems. So, there is need to find an estimator combining both advantages of the L_1 and L_2 norms in the class of convex functions.

In this paper, we propose a robust norm, which can be adaptively close to L_1 and L_2 norms, for post-processing the obtained HR image by sparse representation method for LR single image. Using the proposed norm can preserve large gradients corresponding to edges, while smooth, small gradients usually are the effects of noise. The formed objective function is twice continuously differentiable and strictly convex, and hence the gradient-based optimization technique can find the unique optimal super-resolution image.

MATERIALS AND METHODS

Modeling the problem

A low-resolution image, Z_l , is assumed to be created from the high-resolution one, x_h , by $Z_l = SH x_h + v$, [1] where H and S are the blur and decimation operators, respectively. The operator S performs decimation by an integer factor. The vector v is an additive zero-mean white Gaussian noise, with standard-deviation σ . Let x_p , having the same size as x_h , be the scaled-up image of Z_l .

Let $p_k^h = R_k y_k \in \mathbb{R}^n$ be a high-resolution image patch of size $\sqrt{n} \times \sqrt{n}$, extracted by R_k the operator from the image x_h in location k . We assume that p_k^h can be represented sparsely by $q_k \in \mathbb{R}^m$ over the dictionary $A_h \in \mathbb{R}^{n \times m}$ ie $p_k^h = A_h q$, where $\|q\|_0 \ll n$.

Consider the corresponding low-resolution patch $p_k^l = R_k y_l$, extracted from y_l in the same location k , which size is $\sqrt{n} \times \sqrt{n}$. We assume that $p_k^l = L p_k^h + \tilde{v}_k$, where L is a local operator and \tilde{v}_k is the additive noise in this patch.

Since $p_k^h = A_h q$, we have
 $Lp = LA_h q$. [2]

Exploiting the relation between the low- and the high-resolution patches, $p_k^l = Lp_k^h + \tilde{v}_k$, we thus get

$$LA_h q = Lp_k^h = p_k^l - \tilde{v}_k \quad [3]$$

implying that j

$x_l^j = Lx_h^j + \varepsilon^j$ [4], where ε is related to the noise power σ of \tilde{v} .

The derivations above imply that the low-resolution patch should be represented by the same sparse vector q over the effective dictionary $A_l = LA_h$, with an error ε_l . This further implies that for a given low-resolution patch p_k^l , we should find its sparse representation vector, q , and then we can recover p_k^h by simply multiplying this representation by the dictionary A_h .

Super-resolution

According to the above model, the single-image super-resolution algorithm consists of a dictionary-training phase and super-resolution reconstructing phase.

Training of image dictionary: Before training of image dictionary, there needs to be construction of patch pairs. As examples, we collected some high-resolution image $\{x_h^j\}_j$. Each of these images is blurred and down-scaled by a factors, which forms the corresponding low-resolution image $\{z_l^j\}_j$. These low-resolution images are then scaled-up back to the original size, resulting in the image set $\{x_l^j\}_j$. The procedure can be denoted as: $x_l^j = Lx_h^j + \varepsilon^j$.

Then the different images are computed as

$$e_h^j = x_h^j - x_l^j \quad [5]$$

After some pre-processing steps such as removing their low-frequency components from e_h^j and filtering by high-pass filters, we can extract the local patches, and form the dataset $P = \{p_k^h, p_k^l\}_k$. The dataset is further processed to reduce the dimension.

The first step of the dictionary-training phase is to construct the low-resolution dictionary $A_l \in \mathbb{R}^{n_l \times m}$ from the low-resolution patches $\{p_k^l\}_k$ and obtain the sparse representation vector q_k corresponding to the training patch p_k^l .

Then for constructing the high-resolution dictionary, which is used to recover the patch p_k^h by approximating it as being $p_k^h \approx A_h q_k$, we should seek a dictionary A_h such that this approximation is as exact as it can be. Thus, we define this dictionary to be the solver of the problem

$$\begin{aligned} A_h &= \arg \min_{A_h} \sum_k \|p_k^h - A_h q_k\|_2^2 \\ &= \arg \min_{A_h} \|P_h - A_h Q\|_F^2 \end{aligned} \quad [6]$$

where the matrix P_h contains $\{p_k^h\}_k$ as its columns, and Q contains $\{q_k\}_k$ as its column. The solution of the problem is given by

$$A_h = P_h Q^+ = P_h Q^T (QQ^T)^{-1}. \quad [7]$$

Considering the high-resolution patches overlap, the eventual high-resolution images would be constructed by position-

ing these patches and averaging over their overlaps.

Super-resolution of image: For a given low-resolution image z_l , which has been generated from a high-resolution image x_h by the same blur and scale-down operations as used in the train-ing, the following steps are used for performing the super-resolution:

- Scale this image z_l up by a factor of s using interpolation method, resulting in z_l^s .
- Filter the image z_l^s using the high-pass filters that were used in the dictionary-training phase, and obtain a set of images only with high frequencies. Extract patches from these images, each of size $\sqrt{n} \times \sqrt{n}$. Each patch that corresponds to the same location is to be concatenated to form a patch vector $\{p_k^l\}_k$. All patches form the set $\{p_k^l\}_k$.
- Process the found patches $\{p_k^l\}_k$ by the same method as the training phase to reduce their dimensionality, resulting with the set $\{p_k^l\}_k$, each patch of length n_l . Apply the OMP algorithm on $\{p_k^l\}_k$, allocating L atoms to their representation, and find the sparse representation vectors $\{q_k\}_k$.
- Multiply the representation vectors $\{q_k\}_k$ by the high-resolution dictionary A_h , and obtain the approximated high-resolution patches, $\{A_h q_k\}_k = \{p_k^h\}_k$.
- Construct the final super-resolved image by putting $\{p_k^h\}_k$ to their proper location, averaging in overlap regions, and adding to the final image x_l .

The adaptive M-estimator for post-processing

M-estimation (17, 18) is based on the minimization of a cost function which measures the residual between the captured LR images and the estimated HR image, and can be written as the following minimization problem

$$\begin{aligned} \hat{x} &= \arg \min_{x_l} \rho(SHx_l - z_l) \\ &= \arg \min_{x_l} \sum_{k=1}^M \rho(e_k) \end{aligned} \quad [8]$$

where $e_k = p_k^l - LA_h q$, and $\rho(x)$ called an M-estimator is an even symmetric positive function that has a unique minimizer at $x = 0$ (19). Assuming the zero mean Gaussian noise model in Eq [1], the ML estimation of HR image can be achieved when

$$\hat{x}_l = \arg \min_{x_l} \frac{1}{2} e_k^2 \quad [9]$$

However, because of the L_2 error norm in the cost function in Eq [4], the solution exhibits a poor performance in the presence of intensity outliers or large registration errors (18). Farsiu *et al* made progress by proposing the use of the L_1 error norm as a robust alternative to the L_2 error norm (12) and the reconstructed HR image is solved by

$$\begin{aligned} \hat{x} &= \arg \min_{x_l} \|SHx_l - z_l\|_1 \\ &= \arg \min_{x_l} \sum_{k=1}^M |e_k| \end{aligned} \quad [10]$$

The L_1 error norm is not sensitive to outliers by assigning the same weights (+1 or -1) to all errors (small and large). However, the L_1 norm produces an estimator with higher variance than L_2 norm. As a result, when the residual errors are approximately Gaussian, L_1 norm does not perform as well as L_2 norm. Moreover, $\rho(x) = |x|$ is not differentiable at zero, the gradient descent algorithm used by Zhang and Lam (16) will introduce numerical instability in the iteration process (20).

To increase the robustness to large errors while still being effective to small approximately Gaussian errors, Huber *et al* used Huber function in fidelity term for measuring the difference between the estimated HR image and the captured LR images. Huber estimator can combine the behaviour of L_2 norm when the errors are small, while maintaining the L_1 norm's reduced sensitivity to larger errors (17, 18). To increase robustness further, some redescending M-estimators have been explored recently in multiframe SR reconstruction. An alternate M-estimator using error norm satisfies the above requirements as follows (19),

$$p(x, a) = a \sqrt[3]{a^2 + x^2 - a^2} \quad [11]$$

where a is a positive number. This scale parameter is used to specify the error value at which the influence switches from behaving linearly to behaving like a constant.

With the parameter a fixed, $\rho(x, a) \approx \frac{1}{2}x^2$ when x is small enough and $\rho(x, a) \approx a|x| - a^2$ when x is relatively large. Using the norm in Eq [6], the fidelity term of our robust SR estimation is

$$F = \sum_{k=1}^M \rho(e_k, a) = \sum_{k=1}^M (a\sqrt{a^2 + e_k^2} - a^2)$$

where a is the threshold parameter. And the solution can be obtained by the following minimization problem

$$\begin{aligned} \hat{X} &= \arg \min_x \sum_{k=1}^M (a\sqrt{a^2 + e_k^2} - a^2) \\ &= \arg \min_x \sum_{k=1}^M a\sqrt{a^2 + e_k^2} \quad [12] \end{aligned}$$

Generally, the SR image reconstruction is often an ill-posed inverse problem because of insufficient additional information of HR image (2). Therefore, regularization technique is necessarily applied in SR to well pose this problem. The solution for the regularized SR methods is given by

$$\hat{X} = \arg \min_x \{F(X) + \lambda R(X)\} \quad [13]$$

where $F(X)$, called the fidelity term, measures the closeness of an estimated HR image to the captured LR images. The term $R(X)$, called the regularization term, is utilized to regularize the problem and to achieve a stable solution to the problem. The scalar λ is the regularization parameter to balance the weight between the fidelity term and the regularization term. As Eq [12], the functional to be minimized is defined by

$$C(X) = \sum_{k=1}^M a_1 \sqrt{a_1^2 + e_k^2} + \lambda \sum_{k=1}^M a_2 \sqrt{a_2^2 + \|\nabla X\|^2} \quad [14]$$

where $\|\nabla X\|$ is the first differential of HR image X , and a_1 ,

a_2 are the threshold parameter.

Conjugate gradient (CG) optimization can be used for solving the problem guarantees the CG optimization technique converge to the global minimizer. The procedure for CG optimization of our model (13) is described as follows. The current HR estimate is updated by

$$\hat{X}_{n+1} = \hat{X}_n + \alpha_n P_n, \quad n = 1, 2, \dots$$

where P_n is the conjugate-gradient vector at the n -th iteration with initializing $P_0 = -\nabla C(\hat{X}_0)$. The vector sequence P_n

$$P_{n+1} = -\nabla C(\hat{X}_{n+1}) + \beta_{n+1} P_n, \quad [15]$$

Here

$$\beta_{n+1} = \frac{\nabla C(\hat{X}_{n+1})^t \nabla C(\hat{X}_{n+1})}{\nabla C(\hat{X}_n)^t \nabla C(\hat{X}_n)} \quad [16]$$

RESULTS

We present three experiments that demonstrate the above described super-resolution algorithm.

The experiments are applied to scale-up an image on both directions by a factor $s=3$. Each high-resolution image for training is firstly blurred by separable filter both horizontally and vertically, and then it is down-sampled by a factor $s=3$. Therefore, the scaled-down image z_l is one-ninth of the original image size. The image x_l is created by bicubic interpolation of z_l , with the original size.

Extraction of features from the low-resolution images is done using four filters that perform first and second horizontal and vertical derivatives. These filters are applied such that only sampled pixels are used in the filtering computation. Thus, the patch size used is $n = 9$.

Figure 1 shows the first experimental result, where Fig. 1a is the test image used for the super-resolution experiment, which is extracted from the original one shown in Fig. 1b at both directions by a factor $s=3$; Fig. 1c shows the corresponding consequential image with the proposed method. For comparison, the result only with sparse representation and training method is also shown in Fig. 1d.

Figure 2 shows the second experimental result, of which Fig. 2a is the experimental image used for super-resolution, extracted from the original image (Fig. 2b) at both directions by a factor $s=2$, while Fig. 2c is the corresponding result with the proposed method. For comparison, Fig. 2d is the result

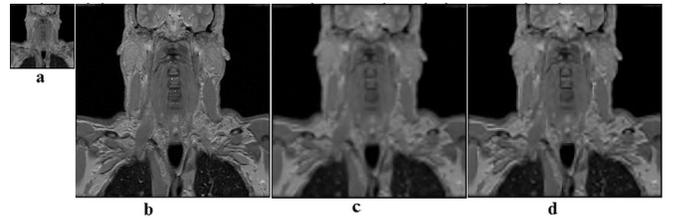


Fig. 1: The original image and consequential high-resolution images of the first experiment. (a) test image, (b) original high-resolution image, (c) high-resolution image produced with proposed method, (d) high-resolution image only with sparse representation and training method.

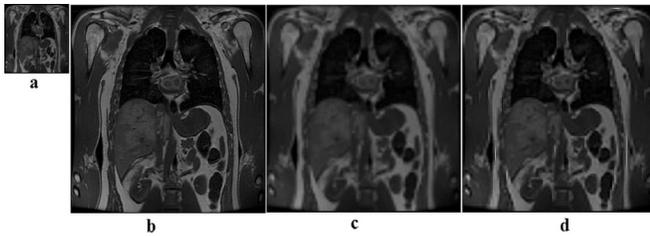


Fig. 2: The test image and corresponding high-resolution images of the second experiment. (a) test image, (b) original high-resolution image, (c) high-resolution image produced with proposed method, (d) high-resolution image only with sparse representation and training method.

Figure 3 shows the third experimental result; Fig. 3a is the experiment image used for super-resolution, extracted from the original image (Fig. 3b) at both directions by a factor $s=3$, while Fig. 3c shows the corresponding consequential image with the proposed method. For comparison, the results only with sparse representation and training method are also shown in Fig. 3d.

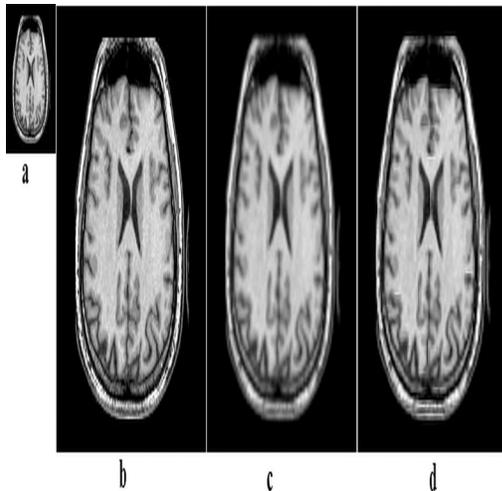


Fig. 3: The original image and consequential high-resolution images of the third experiment. (a) test image, (b) original high-resolution image, (c) high-resolution image produced with proposed method, (d) high-resolution image only with sparse representation and training method.

DISCUSSION

We can observe from Figs. 1–3 that the high-resolution image obtained with the proposed method has some advantage on visual effect by post-processing the scaled-up images with sparse representation and training method, while making the images appear blurry. The results are consistent with the characteristics of two kinds of algorithms. How to restrain the artifacts and reduce the effect of blurring is the key issues of the proposed method. Some further improvements that can be considered are:

- Appropriate parameter a_1, a_2 should be chosen respectively for tradeoff the effective region of L_1 and L_2 .
- Appropriate parameter a_1, a_2 should be chosen for bal-

ancing the weight between the fidelity term and the regularization term.

- A region based post-processing method can be applied to process different regions of the scaled-up image.

CONCLUSION

The paper used a method which combines sparse signal representation and image dictionary learning with adaptive M-estimator to implement single-image super-resolution. The algorithm can be carried out to preserve the profile, sharp edges and details of the original image while enlarging the image. The consequential image with proposed algorithm is prior to the sparse-representation and training method only. However, further work with optimizing the choice of parameters needs to be performed.

ACKNOWLEDGMENT

The work was supported by Hubei Province Natural Science Foundation (Grant no. 2010CDZ057), and the project was supported by the Special Fund for Basic Scientific Research of Central Colleges, South-Central University for Nationalities (Grant no. CZZ11005).

REFERENCES

1. Freeman WT, Jones TR, Pasztor EC. Example-based super-resolution. *IEEE Comput Graph Appl* 2002; **22**: 56–65.
2. Park SC, Park MK, Kang MG. Super-resolution image reconstruction: a technical overview. *IEEE Sig Process Mag* 2003; **20**: 21–36.
3. Katsaggelos A, Molina R, Mateos J. Super resolution of images and video. Morgan and Claypool Publishers; 2007: 12–20.
4. Gonzalez RC, Woods RE. Digital image processing. Upper Saddle River, NJ: Prentice-Hall, Inc; 2002: 568–84.
5. Dai S, Han M, Xu W, Wu Y, Gong Y. Soft edge smoothness prior for alpha channel super resolution. Presented at the IEEE Conference on Computer Vision and Pattern Recognition; Minneapolis; June 17–22, 2007.
6. Sun J, Xu Z, Shum HY. Image super-resolution using gradient profile prior. Presented at the IEEE Conference on Computer Vision and Pattern Recognition; Anchorage; June 23–28, 2008.
7. Kim KI, Kwon Y. Example-based learning for single-image super-resolution. In: Rigoll G, ed. Pattern recognition. Lecture Notes in Computer Science. Proceeding of the 30th DAGM Symposium; Munich, Germany; June 10–13, 2008; 456–65.
8. Elad M. Sparse and redundant representations: from theory to applications in signal and image processing. New York: Springer-Verlag; 2010; 341–52.
9. Yang J, Wright J, Huang TS, Ma Y. Image super-resolution via sparse representation. *IEEE Trans Image Process* 2010; **19**: 2861–73.
10. Yang J, Wright J, Huang T, Ma Y. Image super-resolution as sparse representation of raw image patches. Presented at the IEEE Conference on Computer Vision and Pattern Recognition; Anchorage; June 23–28, 2008.
11. Xie Q, Sang N. Image super-resolution via sparse-representation and iterative back-projection method. *J Comp Inf Sys* 2012; **8**: 6129–37.
12. Farsiu S, Robinson MD, Elad M, Milanfar P. Fast and robust multiframe super resolution. *IEEE Trans Image Process* 2004; **13**: 1327–44.
13. Ng MK, Shen H, Lam EY, Zhang L. A total variation regularization based super-resolution reconstruction algorithm for digital video, *EURASIP J Adv Signal Process* 2007; **2007**: 074585.
14. El-Yamany NA, Papamichalis PE. Robust color image superresolution: an adaptive M-estimation framework. *EURASIP J Image Video Process* 2008; **2008**: 763254.
15. Shankar PM, Neifeld MA. Sparsity constrained regularization for multiframe image restoration. *J Opt Soc Am A Opt Image Sci Vis* 2008; **25**: 1199–214.

16. Zhang X, Lam EY. Superresolution reconstruction using nonlinear gradient-based regularization. *Multidimens Syst Signal Process* 2009; **20**: 375–84.
17. Huber PJ, Ronchetti EM. *Robust statistics*. 2nd ed. Hoboken: John Wiley and Sons; 2009: 167–80.
18. Aster RC, Thurber CH. *Parameter estimation and inverse problems*, 2nd ed. Waltham: Academic Press; 2012.
19. Rabie T. Robust estimation approach for blind denoising, *IEEE Trans Image Process* 2005; **14**: 1755–65.
20. Fouad MM, Dansereau RM, Whitehead AD. Image registration under illumination variations using region-based confidence weighted M-estimators. *IEEE Trans Image Process* 2012; **21**: 1046–60.