Improved IBP for Super-resolving Remote Sensing Images

Feng Li, Donald Fraser, Xiuping Jia

School of Information Technology and Electrical Engineering, ADFA, University of New South Wales, ACT, Australia
E-mail: x.jia@adfa.edu.au

Abstract
The research on super-resolution (SR) image recovery has been carried out in the last two decades. With the fast development of computer technology, more and more efficient algorithms have been put forward in recent years. The Iteration Back Projection (IBP) method is one of the popular methods with SR. In this paper, a modified IBP is proposed for remote sensing image processing. This improved IBP can efficiently deal with local affine transformations within images for SR. Experiments and results are presented using both a synthetic set of images generated from a single Landsat ETM+ channel and a set of Advanced Land Observing Satellite (ALOS) imagery.

Keywords
Super-resolution, image registration, Landsat, ALOS, iteration back projection

I. INTRODUCTION

Super-resolution (SR) recovery has become an important research area for remote sensing images ever since T.S. Huang first published his frequency method in 1984, because of the conflict between the exposure time and the light intensity reaching the high density Charge-Coupled Devices (CCD) of satellite sensors. If the exposure time is too short, then the charge pattern formed by light in the CCD will be swamped by the electronic noise. On the contrary, if the exposure time is too long, the movement along orbits will produce a blur, which is aggravated by the imperfection of optics. A solution to solve this conflict is to obtain SR images from several low-resolution (LR) satellite images based on some post-processing super-resolution techniques. These low-resolution satellite images can be captured as a sequence, recorded at the same time with different sensors or even taken at different time with different sensors.

SR methods have been successfully used in many applications including medical imaging, remote sensing, high-definition television, surveillance systems, video frame freezing and printing. They can be divided into the following main algorithms: the non-uniform interpolation approach, the frequency domain approach, the stochastic reconstruction approach, the Projection Onto Convex Sets (POCS) approach and Iteration Back Projection (IBP) approach. Non-uniform interpolation is the intuitive method for doing SR, and possibly the only choice for doing SR if there is only one LR image available. However, it is hard to guarantee global optimality for all LR images. Frequency domain approach is the first method used for SR from 1984(Huang T, et al., 1984). The major advantages of the frequency domain approach are that it is theoretically simple and easily implemented in a parallel model. However, the shortcoming is that the observation model is only for global translational motion(Park S, et al., 2003), which is not suitable for satellite images (for detail, see the last paragraph of Sect. 2 about the property of satellite images). Maximum A-Posteriori (MAP) is one of the popular stochastic methods. The main advantages of MAP are the direct inclusion of a-priori constraints for the ill-posed problem and it can be used even in the wavelet domain(Zhao S, et al., 2003). According to the method of POCS, incorporating a priori knowledge into the solution can be interpreted as a restriction to getting the global optimality in the convex set. Simplicity is the major advantage of POCS(Borman S, et al., 1998); however, non-uniqueness of solution, slow convergence and high computational cost are the disadvantages. IBP is similar to the back projection used in tomography. The advantage is that it is understood easily and intuitively.

In this paper, a modified IBP algorithm is presented which is efficient in image registration and image spatial resolution improvement.

II. ITERATION BACK PROJECTION

Let us review the IBP algorithm briefly. One critical step for the IBP method is to construct the model for imaging process as:

$$g_k(y) = DH^{psf} \times f(x) + n_k \quad (1)$$

where g_k are the k^{th} observed LR images, y denotes the pixel of LR images influenced by the area of x of the SR image f, H^{psf} is the point spread function of the blur kernel, D means the decimating operator and n_k is an additive noise term. Then, a true SR image is assumed and several LR images are calculated based on the imaging model. If the calculated LR images are the observed LR images, the assumed SR image is regarded as the true SR image. So, the error images between the calculated LR images and the observed LR images are back-projected to the assumed SR image. As the process repeats, the energy of the error becomes smaller and smaller until, finally, a SR image evolves. The scheme of IBP to estimate the SR image is expressed by

1082-4006/06/1202-106$5.00

$$f^{(n+1)}(x) = f^{(n)}(x) + \sum_{y}(g_k(y) - g_k^{(n)}(y)) \times H^{BP} \quad (2)$$

where $f^{(n)}$ is the estimated SR image after n iterations, and $g_k^{(n)}$ are calculated LR images from the imaging model of $f^{(n)}$ after n iterations, H^{BP} is the back projection kernel (a H^{BP} good $H^{BP} = H^{psf}$) choice of is. In determining the model of imaging process, since we cannot know the imaging process exactly, it can be estimated from the degradation of the original small dots or sharp edges.

The other critical step of the IBP method is image registration. As far as remote sensing images are concerned, the image registration method used in [5] limits its efficiency, although Irani and Peleg modified the original registration method in order to consider a more general motion model(Irani M, et al., 1993). In [6], the number of parameters in the motion model was increased from 3 to 8 to describe the 2-D transformation. However, it is still not unable to accurately describe the complex distortions contained in satellite images. Based on Ardy Goshtasby's projective transformation analysis(Goshtasby A, 2005), affine is an acceptable transfer function when registering images taken from a distant platform of a flat scene. The capture of the remote sensing images is reasonably under this condition, as the distance between the sensors on satellites and the scene is often longer than 600 km. Since the distortion caused by cloud, atmospheric turbulence, and other noise is often different in different areas in a LR remote sensing image, methods which consider local warps (like the elastic registration method(Periaswamy S, et al., 2003; Kostelec P, et al., 2003)) and deal with individual local distortions that original method in IBP cannot handle are needed. We reconstruct super-resolution images based on the following improved elastic image registration algorithm in lieu of other methods in this category.

III. ELASTIC IMAGE REGISTRATION

In 2003, Periaswamy and Farid proposed an elastic registration method [8] in circumstances where the intensity varies, which models the transform function between images as local affine. The highlights of this method are that it takes account of intensity variations with a smoothness constraint and is able to deal with local affine warps. It uses an 8-dimensional vector to present the difference between small blocks in source and target images to avoid any wrong influence from global affine and brightness/contrast variations. Four of the eight dimensions represent the affine transformation; two of them are used for translation, and the last two dimensions for the contrast and brightness, respectively. A pyramid method is adopted to deal with the large translation within images. Also, a smooth constraint is introduced to link different local affine transforms and intensity variations. While the result is good, the efficiency is low because the forwards compositional algorithm [10] is adopted to calculate the affine transformation parameters, which means the Hessian matrix has to be calculated at every iteration (for detail, see Baker S, et al., 2004). In this paper, we combine the inverse compositional algorithm with the forward compositional algorithm to improve the efficiency of the elastic registration method.

Here, we use a 7-dimensional vector to describe the warp and the brightness variation between the source and target images. We denote $T(x, y)$ and $S(x, y)$ as the target and source images, respectively. So the motion between them can be modeled locally by an affine transform:

$$T(x, y) = S(m_1 x + m_2 y + m_5, m_3 x + m_4 y + m_6) + m_7 \quad (3)$$

It can be denoted as $T(x, y)=S(W(x, y; M)) + m_7$, where

$$W(x, y; M) = \begin{bmatrix} m_1 & m_2 & m_5 \\ m_3 & m_4 & m_6 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (4)$$

is used to express the affine warp to the coordinate frame and

M is the affine parameters matrix $\begin{bmatrix} m_1 & m_2 & m_5 \\ m_3 & m_4 & m_6 \\ 0 & 0 & 1 \end{bmatrix}$.

As far as brightness and contrast variations are concerned, we model them with one parameter, because it is reasonable to assume both the changes can be approximately described in terms of local intensity variation. In this way, a neat and perfect model embedded with both the inverse compositional algorithm and the forward compositional algorithm is constructed. Experimental results indicate that our assumption is acceptable, as long as we implement a smoothness constraint.

A current M is assumed to be known. Based on Baker and Matthews' inverse compositional algorithm theory(Baker S, et al., 2004), the quadratic error function is defined as:

$$E(\overrightarrow{\Delta m}) = [T(W(x, y; \Delta M)) - S(W(x, y; M)) - \Delta m_7]^2 \quad (5)$$

where

$$\overrightarrow{\Delta m} = [\Delta m_1, \Delta m_2, \Delta m_3, \Delta m_4, \Delta m_5, \Delta m_6, \Delta m_7]^T$$

and

$$\Delta M = \begin{bmatrix} \Delta m_1 & \Delta m_2 & \Delta m_5 \\ \Delta m_3 & \Delta m_4 & \Delta m_6 \\ 0 & 0 & 1 \end{bmatrix}$$

The goal is to minimize ΔM and update the affine parameter matrix M at the end of every iteration. If $T(W(x, y;\Delta M))$ is expanded using a first-order truncated Taylor series, then

$$\begin{aligned} E(\overrightarrow{\Delta m}) \approx [&T(x, y) + (\Delta m_1 x + \Delta m_2 y + \Delta m_5 - x)T_x + \\ &(\Delta m_3 x + \Delta m_4 y + \Delta m_6 - y)T_y - \\ &S(W(x, y; M)) - \Delta m_7]^2 \end{aligned} \quad (6)$$

where T_x and T_y are the spatial derivatives of $T(x, y)$. Therefore the error function can be simplified as following:

$$E(\overrightarrow{\Delta m}) = \sum_{x, y \in \Omega} [k - \vec{c}^T \overrightarrow{\Delta m}]^2 \quad (7)$$

where k and vector $\vec{c}$ are given by

$$k = S(W(x, y; M)) + xT_x + yT_y - T(x, y) \quad (8)$$

and

$$\vec{c} = (xT_x, yT_x, xT_y, yT_y, T_x, T_y, -1)^T \quad (9)$$

Then the error function in Equation (7) can be minimized by differentiating with respect to the unknowns $\overrightarrow{\Delta m}$ as

$$\frac{dE(\overrightarrow{\Delta m})}{d\overrightarrow{\Delta m}} = \sum_{x,y\in\Omega} -2\vec{c}[k - \vec{c}^T\overrightarrow{\Delta m}] \quad (10)$$

Set, $H = \sum_{x,y\in\Omega} \vec{c}\vec{c}^T$ then

$$\overrightarrow{\Delta m} = H^{-1}[\sum_{x,y\in\Omega} \vec{c}k] \quad (11)$$

Note that the $\overrightarrow{\Delta m}$ obtained from Equation (11) is the warp for $T(x, y)$, so the affine transform matrix M for $S(x, y)$ is updated as following:

$$M \leftarrow M \circ \Delta M^{-1} \quad (12)$$

where $\circ$ is a warping operator. The initial M is a 3×3 identity matrix. m_7 is initially set to be 0 and that is updated as $m_7 = m_7 + \Delta m_7$ at every iteration. After several iterations, the source image is warped by M to match the target image.

Compared with the original elastic registration method, H(the Hessian matrix of T) is a constant matrix, because there is nothing in this matrix that depends on M, so it can be pre-computed. The additional step is to calculate Equation (12) at each iteration. By comparison with the calculation of H, the calculation cost of the additional step is much lower. Let N be the number of pixels in the target image, then the saved cost in each iteration is $O(7^2N-7^2)$. Obviously, the proposed method can also be used to improve the efficiency of the original elastic registration algorithm in processing the global affine warp.

In dealing with the local affine problem, however, it is possible that the derivatives of small blocks are found to be zeros (if the intensity is the same within a small block), then H of the small block is not invertible. So it will be impossible to find the movement parameters within the local areas in images. To avoid this, we combine the inverse compositional algorithm for finding the global affine transform with the forward compositional algorithm for finding the local affine transform. Also, images are divided into blocks, such as 5 by 5, to use the forward compositional method in finding the movement parameters within the local area. To avoid expressing the movement parameters with intensity variation only, we adopt the same smooth constraint method used in [8]. Experimental results show that almost the same registration results are obtained as the original elastic algorithm with less time (for detail, see Table 1).

Table 1. Comparison between methods for medical images

Medical images (256×256pixels)	Time	MSE(between the registered image and the target image)
Original registration method	332.781s	0.08208
Improved registration method	212.409s	0.08556

IV. SUPER RESOLUTION WITH IMPROVED IBP

The following procedure is developed to get SR images with the aid of the modified IBP.

Step 1: Following the essence of IBP, we assume that there exists a super resolution image. One of the observed LR images is selected and enlarged by the bi-cubic interpolation method to a finer grid as the initial guess of the SR image.
Step 2: Other observed LR images will be enlarged too for the preparation of image registration. Let the initial SR image be the target image in image registration, then all other enlarged LR images will be registered to the initial SR on the fine grid. This procedure is considerable time consuming; however, it runs once only. In the iterations of back projection, image registration is not required.
Step 3: Following that, the same number of enlarged LR images is estimated based on the assumed SR image and the model of imaging process, then the error images will be calculated by comparison with the enlarged observed LR images. These error images are used to do back projection or modify the assumed SR image to make the error images smaller at the next iteration. As shown in Equation (2), the error image values become smaller and smaller after each iteration.
Step 4: If within the maximum number of iterations, do back projection to the assumed SR and go back to step 3.

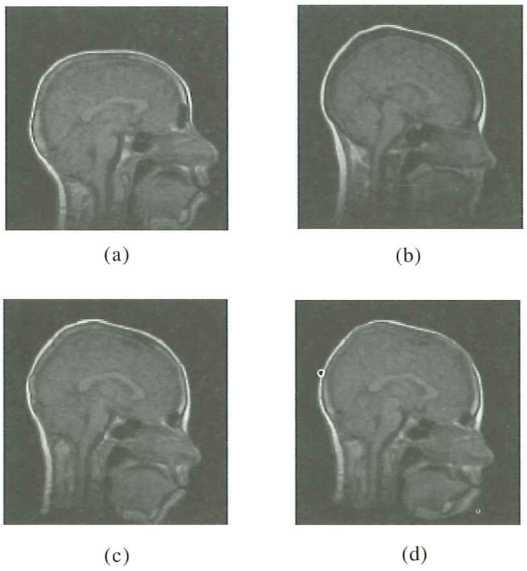
In the original IBP method, one blur kernel is selected for all LR images; however, we believe that the images of different sensors are blurred by kernels of different sizes. Therefore, we treat each LR image separately. Furthermore, we adopt different coefficients for the error images when we do the back projections. The reason is that each LR image contributes different high frequency information in reconstructing the super resolution image.

V. RESULTS

This section is divided into two parts; one is to demonstrate the improved efficiency of image registration method, the other is to present the results of the SR for remote sensing images.

A. Results of improved elastic image registration

For better comparison with the original elastic image registration, we tested the data used in [8] as shown in Fig.1. This medical image has 256×256 pixels with 8 bits per pixel. Around 5 minutes are needed for the original registration method, while the improved registration method needs about



(a) (b)

(c) (d)

Fig.1. Results of the improved elastic registration method used on medical images (from Periaswamy and Farid (Periaswamy S, et al., 2003))
(a) The source image, (b) the target image;
(c) the registered image by the original elastic method;
(d) the registered image by the improved elastic registration method

3 minutes only to provide a similar registration quality as shown in Table 1 and Fig.1. Both algorithms are executed under Matlab for 20 iterations. The computer is a 3.19-GHz Windows machine with 1-GB RAM.

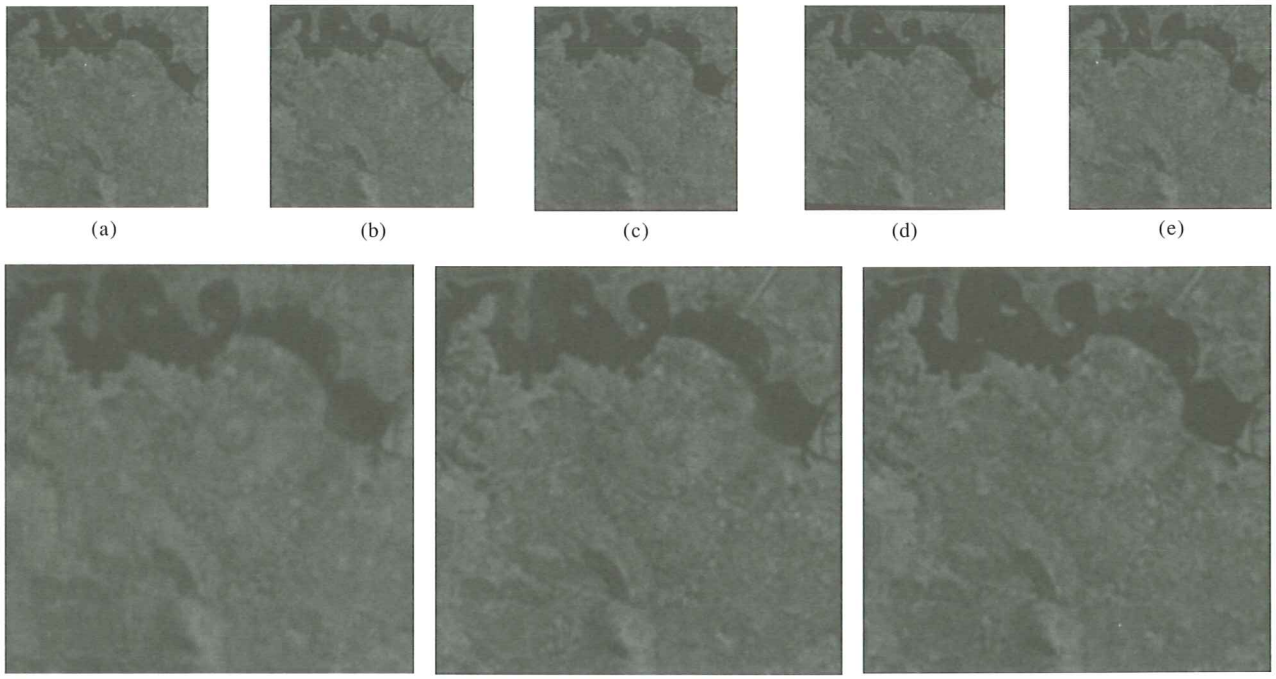
B. Results of SR

Experimental results are provided based on both the synthetic remote sensing images created from a single Landsat ETM+ channel 5 image and a set of ALOS images.

The testing sequence in Fig.2 is created by blurring the original Landsat ETM+ channel 5 image (256×256 pixels of Canberra region, Australia) with a 3×3 Gaussian blur kernel, applying different global affine transformations with different local affine transformations and sub-sampling by a factor of 2. It has 9 low-resolution images and every frame has 128×128 pixels. Fig.2(a) to Fig.2(e) are the 1st, 3rd, 5th, 7th and 9th frame of the LR image sequence. Fig.2(f) is the enlarged mean LR image. Fig.2(g) and Fig.2(h) are the SR images obtained by original IBP and our improved IBP, respectively. Both super-resolved images were obtained after 30 iterations.

Since each LR image has a different local affine transformation, there are shadows at the upper-left corner in Fig. 2(f). Although original IBP performs well compared with the bi-cubic interpolation method, we still can find some shadows at the upper-left corner as shown in Fig. 2(g). However, we can see that the improved IBP reconstructs a better super resolved image (as shown in Fig. 2(h)) without any shadows compared with the original IBP.

Fig.3 shows the experimental result of ALOS images which were captured on Feb. 14, 2006 (JST) and delivered by JAXA (Japan



(a) (b) (c) (d) (e)

(f) (g) (h)

Fig.2. Results of SR used on the synthetically generated LR sequence from a Landsat ETM+ channel 5 image (within Canberra, Australia)
(a) frame 1; (b) frame 3; (c) frame 5; (d) frame 7; (e) frame 9; (f) the mean LR image enlarged by bi-cubic interpolation method; (g) original IBP method after 30 iterations; (h) our improved IBP method after 30 iterations

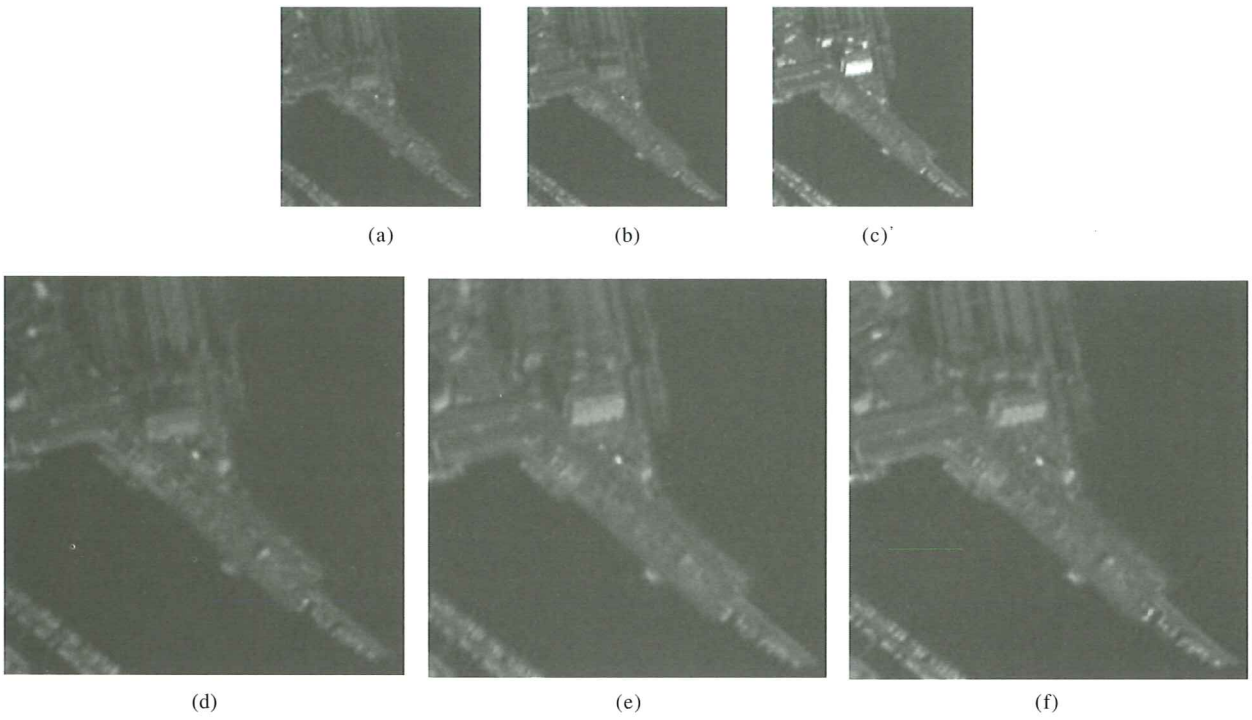


Fig.3. Results of SR used on the ALOS images
(a) The panchromatic forward image; (b) the panchromatic nadir image; (c) the panchromatic backward image; (d) the enlarged forward image by the bi-cubic interpolation method; (e) the original IBP after 30 iterations; (f) the improved IBP after 30 iterations

Aerospace Exploration Agency). ALOS is the Japanese satellite launched in January 2006 and that carries three sensors: Panchromatic Remote Sensing Instrument of Stereo Mapping (PRISM), Advanced Visible and Near Infrared Radiometer type-2 (AVNIR-2) and Phased Array type L-band Synthetic Aperture Radar (PALSAR). The repeat cycle of ALOS is 46 days. The PRISM is a panchromatic radiometer with 2.5-meter spatial resolution and has three independent optical systems for viewing nadir, forward and backward so as to produce a stereoscopic image along the satellite's track. While PRISM is mainly used to construct a 3-D scene, here we use these three panchromatic low-resolution (LR) images captured by nadir, backward and forward sensors to reconstruct one SR image. Fig.3(a) to (c) are the cropped (128×128) panchromatic forward, nadir and backward images of PRISM (within Shimizu port in Shizuoka, Japan), respectively. Fig.3(d) is the forward image enlarged by the bi-cubic interpolation method. Let us compare the SR images obtained after 30 iterations by both the original IBP method in Fig.3(e) and the improved IBP method in Fig.3(f). It can be easily seen that there is more high frequency detail with the SR image obtained by the improved method. In particular, the boats in the left corner in Fig.3(f) can be seen to be much clearer and sharper.

VI. CONCLUSIONS

After analysing the characteristics of remote sensing images, an efficient IBP method is proposed in this paper to achieve image spatial resolution improvement. The contributions of this paper include: (1) we improved the efficiency of the elastic registration method and (2) we embedded the improved registration method in a modified IBP method to reconstruct better SR images.

While we improved the efficiency of the elastic registration method, it still takes considerable time. For instance, about 3 minutes are needed to do image registration for grey images of pixels with 8 bits per pixel. Most of the time is taken in the processing of the local affine warps in the images. Future work will be involved in improving the efficiency further for the elastic image registration in dealing with the local affine warps.

REFERENCES

[1] Huang T., Tsai R., 1984, "Multi-frame image restoration and registration," *Advances in Computer Vision and Image Processing,* 1: 317–339.
[2] Park S., Park M., Kang M., 2003, "Super-resolution image reconstruction: a technical overview," *Signal Processing Magazine, IEEE,* 20(3): 21–36.
[3] Zhao S., Han H., Peng S., 2003, "Wavelet-domain HMT-based image super-resolution," *International Conference on Image Processing,* 2: 953–956.
[4] Borman S., Stevenson R., 1998, "Super-resolution from image

sequences-a review," *Proceedings of the 1998 Midwest Symposium on Circuits and Systems* 5.

[5] Irani M., Peleg S., 1991, "Improving resolution by image registration," *CVGIP: Graph. Models Image Process,* 53(3): 231–39.

[6] Irani M., Peleg S., 1993, "Motion analysis for image enhancement: Resolution, occlusion, and transparency," *Journal of Visual Communication and Image Representation,* 4(4): 324–35.

[7] Goshtasby A., 2005, *2-D and 3-D Image Registration for Medical, Remote Sensing, and Industrial Applications*, Wiley-Interscience.

[8] Periaswamy S., Farid H., 2003, "Elastic registration in the presence of intensity variations," *IEEE Transactions on Medical Imaging,* 22: 865–74.

[9] Kostelec P., Periaswamy S., 2003, "Image Registration for MRI," *Modern Signal Processing,* 46: 161–84.

[10] Baker S., Matthews I., 2004, "Lucas-kanade 20 years on: A unifying framework," *Int. J. Computer. Vision,* 56(3): 221–55.