16th European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, August 25-29, 2008, copyright by EURASIP

ELASTIC REGISTRATION OF REMOTE-SENSING IMAGES BASED ON THE NONSUBSAMPLED CONTOURLET TRANSFORM

Chahira Serief [1], Mourad Barkat[2] (IEEE Fellow), Youcef Bentoutou[1]

[1] National Center of Space Techniques (CNTS), 01 Avenue de la Palestine, P.O. Box 13, Arzew 31200, Algeria
serief07@hotmail.com; bentyou@yahoo.com
[2] Department of Electrical Engineering, American University of Sharjah, P.O. Box 26666, Sharjah, UAE
mbarkat@gmail.com

ABSTRACT

Image registration plays a critically important role in remote sensing applications. Due to the large volumes of remote-sensing data available today, automated registration of multitemporal and/or multisensor images is highly desired. In this work, a new automatic approach for elastic image registration of remotely sensed images is proposed. The critical elements for an automated image registration procedure are explored. The elements include control-point (CP) extraction, CP matching, and transformation parameters estimation. In the proposed algorithm, a new CPs extractor has been developed. This technique exploits the nonsubsampled contourlet transform (NSCT) to automatically extract a set of CPs where misalignment between images can be expected to appear. The proposed algorithm has been successfully applied to register multitemporal ALSAT-1 images from urban and agricultural areas. The experimental results demonstrate the robustness, efficiency and accuracy of the proposed algorithm.

1. INTRODUCTION

Image registration plays an important role in the analysis and fusion of information between multiple images. Examples can be found in remote sensing, robotics, and medical imaging, among others. For all of these applications, image registration is the process of finding the transformation which best matches, according to some similarity measure, two or more images that differ in certain aspects but essentially represent the same object. Such images may be acquired from separated positions, by different sensors and/or at different times. A large number of methods have been developed to solve different variants of this problem. Thorough surveys of image registration techniques have been published by Brown in 1992 [1] and Zitova and Flusser in 2003 [2].

Our interest in the problem of registration stems from remote sensing applications. Image registration is frequently used in remote sensing for a wide variety of tasks such as change detection using multiple images acquired at different times, cartography and photogrammetry using imagery with overlapping coverage, and fusion of image data from multiple sensor types. In remote sensing applications, the traditional procedure for registering a pair of images requires the manual selection of ground control points (GCPs) at significant landmarks of the images. These GCPs are then used to estimate a transformation model which is subsequently used to align the two images, by warping one image with respect to the other using any interpolation function. The primary drawback to this approach is that a trained expert is needed to manually select each individual GCP in the remotely sensed images. This is very laborious and time consuming, especially when dealing with the large volumes of remote-sensing data available today. Therefore, an automatic method of aligning such images is highly desired. Automation of this procedure requires the replacement of the manual control-point selection with automatic algorithms for locating corresponding points in both images [1]. Several techniques have been developed to automate the process of image registration.

This paper proposes a new approach to automated image registration for remotely sensed images alignment. The critical elements for an automated image registration procedure are explored. The elements include control-point (CP) extraction, CP matching, and transformation parameters estimation. In the proposed algorithm, a new control-points extractor has been developed. This new approach exploits a nonsubsampled directional multiresolution image representation recently introduced [3], and called nonsubsampled contourlet transform (NSCT), to capture significant image features across spatial and directional resolutions. This extraction method is a further development of the one proposed in [4]. CP matching is one of the most important tasks in automated image registration. Indeed, the similarity measure used to create the cost function is one of the factors that most influences the quality of registration. In the proposed algorithm, Zernike moments-based similarity measure is used to establish the correspondence between the control points detected from the two images. The CP matching process is performed by computing the correlation coefficients of the Zernike moments descriptor vectors of a circular neighborhood centered on each CP in the two images. The reason for this choice is due to the fact that Zernike moments have been shown to be superior in terms of their orthogonality, rotation invariance, low sensitivity to image noise, information redundancy [5], fast computation, and ability to provide faithful image representation [6]. One of the important factors to achieving accurately registered images is the model that describes the spatial mapping between the images to be registered. The proposed registration method is performed by applying the elastic thin-plate spline (TPS) mapping function on sets of CP pairs [7]. Consequently, the designed algorithm has the capability to register images both globally and locally with subpixel accuracy. By introducing local processing on areas of interest based on control pixels, this new algorithm is fast compared to all traditional algorithms where entire images are searched. The proposed algorithm has been successfully applied to register multitemporal ALSAT-1 images from urban and

agricultural areas. The experimental results demonstrate the robustness, efficiency and accuracy of the algorithm.

The rest of this paper is structured as follows. In Section II, we discuss the main characteristics of the NSCT transform and we give the motivations which brought us to the utilization of this transform for CP detection. Section III describes the proposed registration algorithm in detail. In Section IV, we give the experimental results along with a discussion and assess the registration performance. The conclusion is given in Section V.

2. MOTIVATION

In recent years, the fast growth of multiscale geometric analysis has brought out abundant tools for image processing, such as contourlet transform. Indeed, the contourlet transform as introduced by Do and Vetterli [8], is a new efficient image decomposition scheme, which provides sparse representation at both spatial and directional resolutions. The contourlet transform employs Laplacian pyramids to achieve multiresolution decomposition and directional filter banks to yield directional decomposition, such that, the image is represented as a set of directional subbands at multiple scales. The contourlet transform with its extra feature of directionality achieves better results than the discrete wavelet transform and yields new potentials in image processing applications [8]. Due to downsampling and upsampling, the contourlet transform is shift-variant [9]. However, shift-invariance is desirable in image analysis applications such as edge detection, contour characterization, and image enhancement. Recently, Cunha et *al*. [3], [9] proposed the nonsubsampled contourlet transform (NSCT) which is a shift-invariant version of the contourlet transform. The NSCT is built upon iterated nonsubsampled filter banks to obtain a shift-invariant directional multiresolution image representation. Compared to the contourlet transform, the NSCT is a fully shift-invariant, multi-scale, and multi-direction image decomposition that has a better frequency selectivity and regularity, and a fast implementation [3]. The NSCT has proven to be very efficient in image denoising and image enhancement as shown in [3].

Consequently, the primary motivation of this work is to exploit the NSCT to detect salient features for remotely sensed images registration. The choice of using NSCT to detect image features from remotely sensed images, is justified by the fact that NSCT features are more robust to noise as well as local intensity variations and time-of-day conditions than pure gray levels. The other motivation for detecting features on the basis of NSCT is its fast implementation.

3. PROPOSED REGISTRATION ALGORITHM

Given two images, I_1 (defined as a reference image) and I_2 (defined as a sensed image) to match the reference image, the goal of remotely sensed images registration is to rectify the sensed image into the coordinate system of the reference image and to make corresponding coordinate points in the two images fit the same geographical location. The robustness, accuracy, and efficiency of image registration depend on: 1) the performance of the control-points extraction, 2) the performance of the CPs correspondence step, and 3) the class of spatial transformation used for registration. The goal of this paper is to take each of these factors into consideration in the registration process to provide a practical automatic nonrigid image registration algorithm for remotely sensed images.

In this paper, the registration process is carried out in the following three steps.

3.1 Proposed NSCT-based CPs extraction method

The performance of CP detection step is one of the factors that most influences the quality of registration. This requirement brought us to the utilization of the nonsubsampled contourlet transform to perform the control-points extraction step.

Knowing that the NSCT do not only provide multiresolution analysis, but also geometric and directional representation, and knowing that it is shift-invariant such that each pixel of the transform subbands corresponds to that of the original image in the same location, we can therefore gather the geometric information pixel by pixel from the NSCT coefficients. The proposed feature extraction method is a further development of the one proposed in [4]. Indeed, this new method makes use of the scale-interaction model [10] and is described in the following algorithm:

1. Compute the NSCT coefficients of the image for N levels and L directional subbands.
2. Compute the difference between each directional subband at one level and the corresponding one at another level. L difference subbands will be obtained at the end.
3. At each pixel location, compute the maximum magnitude of all obtained difference subbands. These points are called "maxima of the NSCT coefficients".A hard thresholding procedure is then applied on the NSCT maxima image in order to eliminate non significant feature points. A point is recorded if
$$\text{NSCT maxima} > Th, \quad (1)$$
where $Th = c\,(\sigma + \mu)$, c is a parameter whose value is defined by the user, and σ and μ are the standard deviation and mean of the NSCT maxima image.
4. Take each maxima point as the central point of a block neighbourhood of size $w \times w$ and find one local maximum in each neighbourhood, this will eliminate maxima that are very close to each other.
5. The locations of the obtained thresholded NSCT maxima are taken as the extracted feature points.

An example of the CP candidates extracted using this method is shown in Fig.1.

3.2 CP candidates matching

After the CP candidates have been selected from both reference and sensed images, a correspondence mechanism between these two control-points sets must be established. The objective is that each CP in the reference image is paired with its correspondent in the sensed image.

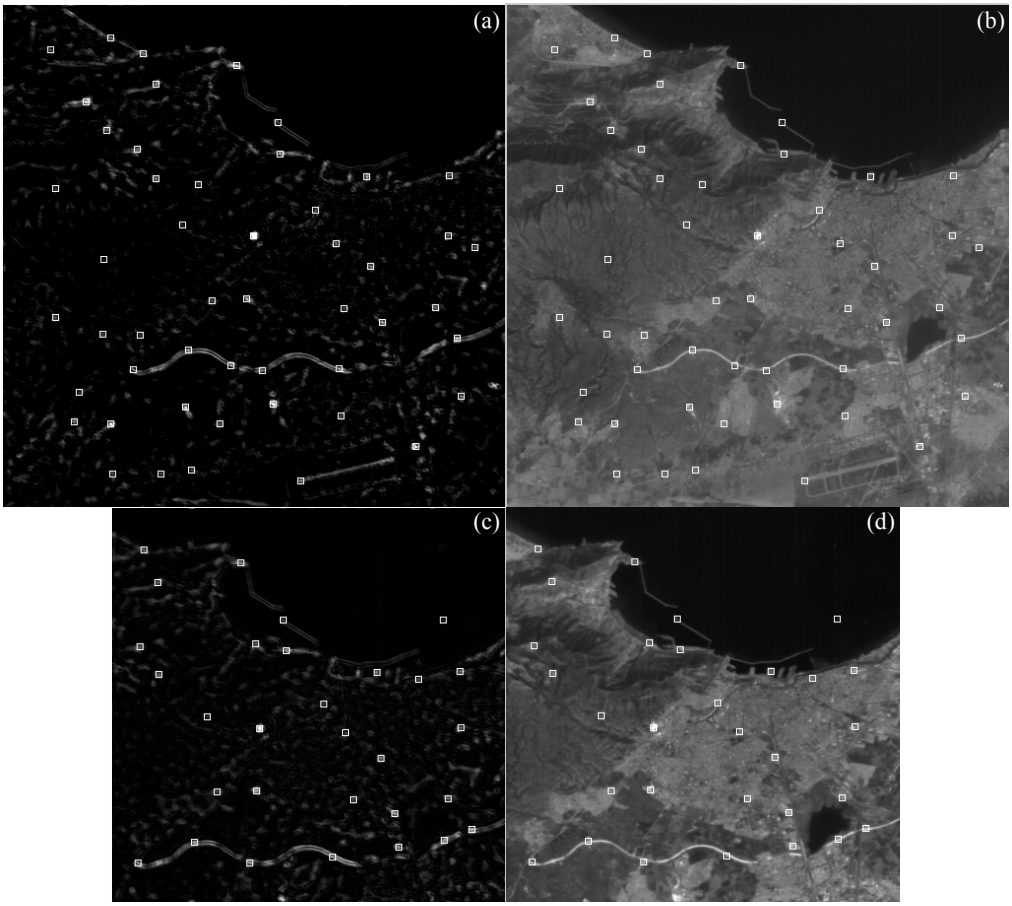

Figure 1 – Extracted CP candidates (as indicated by boxes). NSCT maxima image of the reference image (a) and the reference image (b) marked by the extracted CPs. NSCT maxima image of the sensed image (c) and the sensed image (d) marked by the extracted CPs

In the proposed algorithm, Zernike moments-based similarity measure is used to establish the correspondence between the two images. This correspondence is evaluated using a circular neighborhood of radius R centered on each CP. The reason for selecting the complex Zernike moments as feature descriptors is that they possess a useful rotation invariance property [6]. Rotating the image does not change the magnitudes of its Zernike moments. Hence, they can be used as rotation invariant image features [6]. The defined features on the Zernike moments are only rotation invariant. To obtain scale and translation invariance, the image is first subject to a normalization process using its regular geometrical moments. The rotation invariant Zernike features are then extracted from the scale and translation normalized image [6].

The correspondence between the two sets of CPs is obtained as follows:

- For every point P_i, choose a circular neighborhood of radius R centered at this point and construct a Zernike moments descriptor vector P_z as

$$P_z = \left(|Z_{1,1}|, \ldots, |Z_{p,q}|, \ldots, |Z_{10,10}| \right) \quad (2)$$

  where $|Z_{p,q}|$ is the magnitude of Zernike moments of a non-negative integer order p, $p-|q|$ is even, and $|q| \le p$. While higher order moments carry fine details of the image, they are more sensitive to noise than lower order moments [6]. Therefore, the highest moments order used in the descriptor vector P_z (10 in the algorithm) is chosen to achieve a compromise between noise sensitivity and the information content of the moments.
- The CP matching process is performed by computing the correlation coefficients of the two descriptor vectors. The matched points are those who give the maximum coefficient correlation value in both directions. The correlation coefficient C of two feature vectors V_1 and V_2 is defined as

$$C = \frac{(V_1 - m_1)^T (V_2 - m_2)}{\left\| (V_1 - m_1) \right\| \left\| (V_2 - m_2) \right\|} \quad (3)$$

  where m_1 and m_2 are the means of the two vectors V_1 and V_2 respectively.

3.3 Image warping and resampling

Finally, given the two sets of corresponding CPs, $P = \{p_i\}$ and $Q = \{q_i\}$, the sensed image is transformed based on the final transformation model and then resampled using an interpolation method such as bilinear and bicubic interpolation depending on the level of image quality and the level of computational performance required. Satellite image warping can be performed by applying an affine transform. The affine mapping function is appropriate for data with flat topography. For some data with nonlinear or local geometric distortions (such as hilly

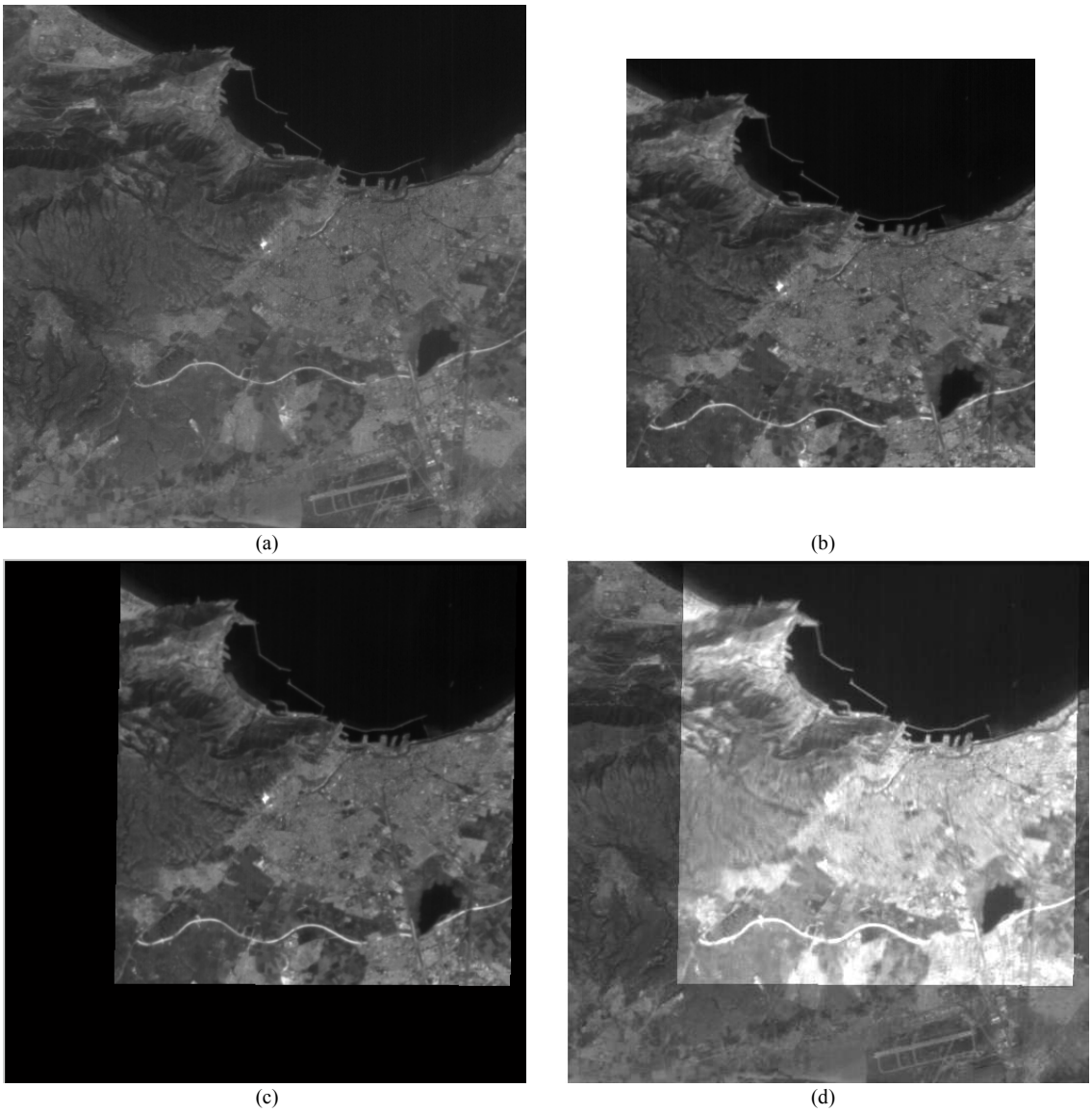

Figure 2 – Registration of ALSAT-1 images of the region of Oran (Algeria), acquired at different dates. (a) The reference image. (b) The sensed image. (c) The registered sensed image obtained after application of the proposed algorithm. (d) The registration result.

areas with terrain changes), more complex transformation functions may be needed to produce better interpolation results. These mapping functions include thin-plate interpolation [7]. In this paper, sensed image warping has been performed by applying the TPS interpolation function on the two sets of corresponding CPs.

One of the most important attributes of TPS is its ability to decompose a space transformation into a global affine transformation and a local non-affine warping component. The TPS are interpolating functions representing, exactly, the distortion at each CP and defining a minimum-curvature surface between CPs. The TPS function is a flexible transformation that allows rotation, translation, scaling, and skewing. It also allows lines to bend according to the TPS model. Therefore, a large number of deformations can be characterized by the TPS model [11]. The TPS interpolation function can be written as

$$\mathbf{h}(\mathbf{x})=\mathbf{A}\mathbf{x}+\mathbf{t}+\sum_{i=1}^{n}\mathbf{W}_i\, K\left(\left\|\mathbf{x}-\mathbf{x}_i\right\|\right) \qquad (4)$$

where **A** and **t** are the affine transformation parameter matrices, $\mathbf{W}_i$ are the weights of the nonlinear radial interpolation function K, and x_i are the CPs. The function $K(\lambda)$ is the solution of the biharmonic equation ($\Delta^2 K=0$) that satisfies the condition of bending energy minimization, namely $K(\lambda)=\lambda^2\log(\lambda^2)$.

4. EXPERIMENTAL RESULTS

To measure the performance of the proposed NSCT-based registration algorithm with respect to registration accuracy and robustness, it was applied to the registration of multitemporal satellite images derived from the first Algerian micro-satellite ALSAT-1 (32-m resolution). In this experiment, a set of two ALSAT-1 images of Band-1 (Green) representing the region of Oran (Algeria) was used. These images are eight-bit grayscale images. Subscenes of around 512x512 pixels and 400x400 pixels from the original images were used. The set consists of a

Table 1. Root Mean Square Errors at the control-points (in pixels)

CP	1	2	3	4	5	6	7	8	9	10	Total CPs RMSE
RMSE	0.0109	0.0843	0.0331	0.0148	0.0382	0.0350	0.0310	0.0125	0.0509	0.1374	0.0582

reference image (512x512 pixels) acquired on 18/08/2003 and a sensed image (400x400 pixels) acquired on 27/05/2003.

The algorithm was implemented in MATLAB. The experiment was conducted according to the following settings: the NSCT decomposition of images, performed using the NSCT toolbox, was carried out with N=5 resolution levels and L=4 directional subbands; the parameter c =1; the block neighbourhood size w = 25; and the Zernike moments-based descriptor neighbourhood's radius R = 11. The particular level-pair used in this experiment is (3, 5).

Given a number of matched CPs from two images, the transformation parameters can be estimated. The TPS transformation function is appropriate for satellite images with nonlinear, local geometric distortions and is needed to produce better interpolation results and to achieve subpixel accuracy. For this type of transformation, the affine transformation parameters and the weights of the nonlinear radial interpolation function are computed in order to resample and transform the sensed image according to the reference image. Since the determination of the affine transformation parameters requires the knowledge of only three CPs, three pairs of CPs with maximum correlation coefficient were used to estimate these parameters. Knowing the TPS mapping function, the sensed image was transformed and resampled using bilinear interpolation. The result of the application of the proposed registration technique to a pair of ALSAT-1 images acquired at different dates are shown in Fig.2. Four images are chosen for this pair: (a) the reference image, (b) the sensed image, (c) the registered sensed image, and (d) the registration result of the two images. It can be seen from Fig.2 (d) that the two images are well superposed. This demonstrates the efficiency and robustness of the proposed NSCT-based registration algorithm.

The traditional measure of registration accuracy is the Root Mean-Square Error (RMSE) between the extracted CPs from the sensed image and the recovered ones, which is given by the following equation.

$$\mathrm{RMSE}=\sqrt{\frac{1}{N}\left(\sum_i \left\|(x,y)_i-(x',y')_i\right\|^2\right)} \qquad (5)$$

where $(x',y')_i = T_{TPS}(x, y)_i$, where T_{TPS} is the computed TPS transformation, $\left\|(x,y)_i-(x',y')_i\right\|$ is the Euclidean distance and N is the number of CPs pairs.

The registration accuracy of the proposed algorithm was estimated using the RMSE of (5) at every CP. Only ten points were selected among the total number of CPs. As shown in Table 1, the results show that a registration accuracy of less than 0.15 pixels has been achieved at each individual CP. A total RMSE of less than 0.1 pixels has been obtained for the ten CPs, which is a relatively high accuracy achieved by the proposed automated algorithm.

5. CONCLUSION

In this paper, we have introduced a new approach based on the nonsubsampled contourlet transform for performing efficient, robust and accurate elastic registration of remotely sensed images. The performance of the proposed NSCT-based registration algorithm was demonstrated by registering two multitemporal ALSAT-1 images. The experimental results show that registration performance and accuracy are relatively high. Indeed, a registration accuracy of less than 0.1 pixels has been achieved for ten CPs.

REFERENCES

[1] L.G. Brown, "A survey of Image registration techniques," *ACM Computing Surveys*, vol. 24, no. 4, pp. 325-3760, Dec. 1992.

[2] B. Zitova and J. Flusser, "Image registration methods: A Survey," *Image Vision Computing*, vol. 21, no. 11, pp. 977-1000, Apr. 2003.

[3] A.L. Cunha, J. Zhou, and M.N. Do, "The nonsubsampled contourlet transform: theory, design, and applications," *IEEE Trans. Image Processing*, vol. 15, no. 10, pp. 3089-3101, Oct. 2006.

[4] C. Serief, M. Barkat, and Y. Bentoutou, "An automatic image registration scheme based on the nonsubsampled contourlet transform," presented at the *IEEE Int. Symposium Signal Processing and its Applications ISSPA* 2007, Feb. 2007.

[5] Cho-Huak Teh and R.T. Chin, "On image analysis by the method of moments," *IEEE Trans. PAMI*, vol. 10, no. 4, pp. 496-513, 1988.

[6] A. Khotanzad and Y.H. Hong, "Invariant image recognition by Zernike moments," *IEEE Trans. PAMI*, vol. 12, no. 5, pp. 489-497, May 1990.

[7] F.L. Boostein, "Principal warps: Thin-plate splines and the decomposition of deformations," *IEEE Trans. PAMI*, vol. 11, no. 6, pp. 567-585, June 1989.

[8] M.N. Do and M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation," *IEEE Trans. Image Processing*,, vol. 14, no. 12, pp. 2091-2106, Dec. 2005.

[9] J. Zhou, A.L. Cunha, and M.N. Do, "The nonsubsampled contourlet transform: construction and application in enhancement," *in Proc. IEEE Int. Conf. Image Processing, ICIP 2005*, vol. 1, pp. I-469-472, Sep. 2005.

[10] B.S. Manjunath, C. Shekhar, and R. Chellappa, "A new approach to image feature detection with applications," *Pattern Recognition*, vol. 31, pp. 627-640, 1996.

[11] Y. Bentoutou and N. Taleb, "An automatic image registration for applications in remote sensing," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 43, no. 9, pp. 2127-2137, Sep. 2005.