Combining SIFT and Individual Entropy Correlation Coefficient for Image Registration

Gan Liu[1], Shengyong Chen[1], Xiaolong Zhou[1], Xiaoyan Wang[1], Qiu Guan[1], and Hui Yu[2]

[1]College of Computer Science and Technology
Zhejiang University of Technology, Hangzhou, Zhejiang Province, China
gan.lau@outlook.com, sy@ieee.org
{zxl,xiaoyanwang,gq}@zjut.edu.cn
[2] School of Computing, University of Portsmouth, Portsmouth, England
hui.yu@port.ac.uk

Abstract. Image registration is an important topic in many fields including industrial image analysis systems, medical and remote sensing. To improve the registration accuracy, an image registration method that combines scale invariant feature transform and individual entropy correlation coefficient (SIFT-IECC) is proposed in this paper. First, scale invariant feature transform algorithm is applied to extract feature points to construct a transformation model. Then, a rough registration image is obtained according to the transformation model. The individual entropy correlation coefficient is used as the similarity measure to refine the rough registration image. Finally, the experimental results show the superior performance of the proposed SIFT-IECC registration method by comparing with the state-of-the-art methods.

Keywords: Image registration, Scale invariant feature transform, Individual entropy correlation coefficient.

1   Introduction

Image registration is the process of spatially aligning two or more images of the same scene acquired with, for example, different sensors or the same sensors at different times [1, 2, 3]. The registration geometrically aligns two images called the reference and floating images, respectively. Image registration has important applications in many fields including remote sensing [4], medical [5], and industrial image analysis systems [6]. In the field of computer vision, image registration is a critical component of image processing, such as image mosaicking [7], image fusion [8], image reconstruction [9] and so forth. Usually, image registration methods are generally categorized into two classes [1, 2]: feature-based method [3] and intensity-based method [6]. In general, the feature-based method is preferably applied when images contain many salient and detectable features, while the intensity-based method is recommended when images contain not enough features or the features are similar. However, in feature-based method, the process of features extracting is sensitive to noise, which

can easily lead the result that features in both reference image and floating image are hard to be detected and/or unstable in time. Therefore, it is hard to make a correspondence between the two feature sets in feature-based method. In intensity-based method, entire images have to be used during the registration steps which cause costly consumption of time and memory. Furthermore, nonlinear illumination changes exert negative effects on the registration result and therefore the intensity-based method does not perform well in stability which may lead to local extremum.

To improve the accuracy and stability, scale invariant feature transform [10] and individual entropy correlation coefficient [11] (SIFT-IECC) are combined to register the reference image and the floating image in this paper. Scale invariant feature transform (SIFT) is an algorithm to detect local features in images. The algorithm was first reported by David G. Lowe in 1999, it became a consummate algorithm till 2004. The SIFT algorithm has been applied in many fields like gesture recognition, object recognition, image stitching, video tracking, 3D modeling, matching moving and so on. In this paper, the SIFT algorithm is applied to extract the feature points that are used to construct a transformation model. A rough registration image is then obtained according to this transformation model. The SIFT-based affine transformation can correct the translation, rotation and scale of the floating image, so we use it first .To get the refined image registration image, the individual entropy correlation coefficient (IECC) is used as the similarity measure.

The flow chart of the proposed method is shown as in Fig. 1. The heavy line represents rough registration process, and thin line represents refining registration process. The corresponding steps of the proposed method are as follows. 1) Extract features of the reference and floating images using SIFT. 2) Match SIFT features of the reference and floating images. 3) Establish the affine transformation model based on the minimum mean square error (MMSE) method of matching feature points. 4) Obtain the rough registered image by transforming the floating image based on the affine transformation model. 5) Initialize the parameters of the IECC-based registration image. 6) Establish an affine transformation model by optimizing initializations or parameters, and then obtain the rough registered image by the affine transformation model. 7) Use IECC as the similarity measure to refine the rough registered image. 8) Output the refined registered image.

2   SIFT-IECC Image Registration

2.1   Feature Extraction Using SIFT

The major stages of the SIFT algorithm are stated in the following.

1. Scale-space extrema detection

The Gaussian kernel has been proved to be the only possible kernel that can produce scale-space [12]. Therefore, the scale-space of an image is defined as a function produced from the convolution of a variable-scale Gaussian kernel with an input image.
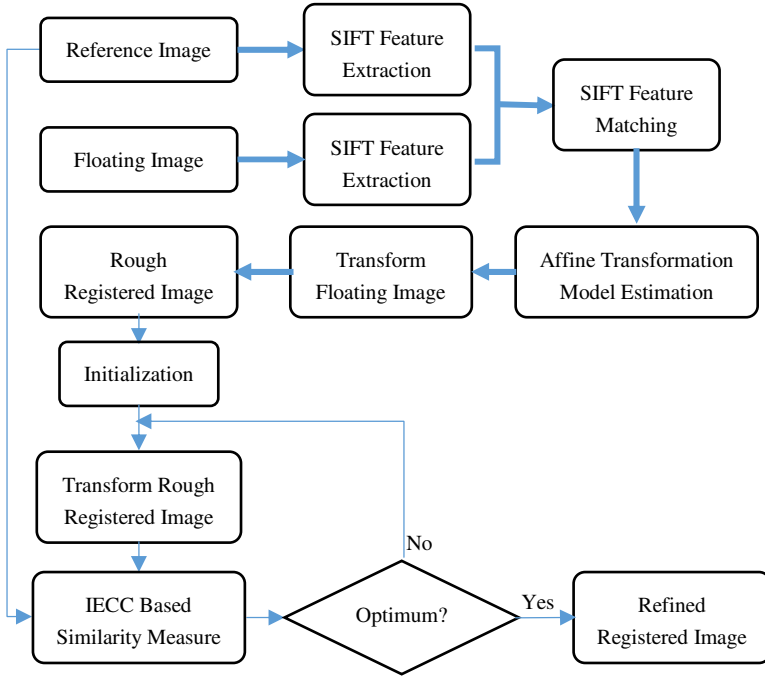
$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \tag{1}$$

Fig. 1. Flow chart of the proposed method

$$G\left(x, y, \sigma\right) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$ (2)

where $L(x,y,\sigma)$ is the scale-space, $G(x,y,\sigma)$ is the Gaussian kernel, $I(x,y)$ is the input image, σ is the scale-space factor, and $*$ is the convolution operation in x and y.

Scale-space extreme is used in the difference-of-Gaussian function convolved with the image for the purpose of detecting stable keypoint locations in scale-space efficiently.

$$D\left(x, y, \sigma\right) = \left(G\left(x, y, k\sigma\right) - G\left(x, y, \sigma\right)\right) * I\left(x, y\right) = L\left(x, y, k\sigma\right) - L\left(x, y, \sigma\right)$$ (3)

where $D(x,y,\sigma)$ is the difference-of-Gaussian function, k is a constant multiplicative factor.

Maxima and minima of the difference-of-Gaussian images are detected by comparing each sample point to its eight neighbors in the current image and nine neighbors in the scale above and below.

2. Keypoint localization

Points those are sensitive to noise or poorly localized along an edge should be rejected to pinpoint the local extreme, enhance the matching stability and improve noise

immunity. $D(x,y,\sigma)$ is approximated by the Taylor expansion   and it can be used to remove the unstable extrema with low contrast by discarding points whose offset is greater than an appropriate threshold. In order to eliminate edge responses, the ratio between the square of Hessian matrix's trace and Hessian matrix's determinant is used. An extremum with a large principal curvature across the edge but a small one in the perpendicular direction will be discarded.

3. Orientation assignment

Using local image properties to assign a consistent orientation to each extremum can guarantee invariance to image rotation. For each Gaussian smoothed image $L(x,y)$, the gradient magnitude $m(x,y)$ and the orientation $\theta(x,y)$ are pre-computed using pixel difference.

$$m(x, y) = \sqrt{\left(L(x+1, y) - L(x-1, y)\right)^2 + \left(L(x, y+1) - L(x, y-1)\right)^2} \qquad (4)$$

$$\theta(x, y) = tan^{-1}\left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)}\right) \qquad (5)$$

4. Keypoint descriptor

Descriptors of the keypoints in both reference image and floating image are needed for matching. A keypoint descriptor is established by first calculating the gradient magnitude and orientation at every point in a region around the keypoint location, which means that the descriptor contains not only the keypoint's information but also pixels' information around the keypoint. These pixels around the keypoint are accumulated into orientation histograms to summarize the contents over 4x4 subregions that can form a descriptor of 8 orientations. 16x16 pixels around a keypoint is used to form 4x4 descriptors of 128 dimensional SIFT feature vector that is used to describe the keypoint, which will achieve invariance to image rotation optimally. To reduce the effects of illumination change, the SIFT feature vector is further normalized.

2.2   Feature Matching and Transformation Model Estimation

1. Feature Matching

Feature matching is used to establish an affine transformation model relying on the correspondences between features in both reference and floating image. An effective method to match feature points is to compare the distance of the nearest neighbor to that of the second nearest neighbor. More specially, with regard to each keypoint in the reference image, we can find the nearest keypoint with the shortest Euclidean distance, d_1, and the second nearest keypoint with second shortest Euclidean distance, d_2, in the floating image. If the ratio, d_1/d_2 is bigger than an appropriate threshold that can be confirmed by testing, the correspondence will be regarded as an incorrect match. K-d tree is used to improve the matching efficiency.

2. Transformation Model Estimation

After feature matching, we can get N matching point-pairs

$$\{(x_{r,i}, y_{r,i}), (x_{f,i}, y_{f,i})\}_{i=1,2,3...,N} \tag{6}$$

where $(x_{r,i}, y_{r,i})$ is the keypoint in the reference image and $(x_{f,i}, y_{f,i})$ is the keypoint in the floating image.

An affine transformation model can be defined as:

$$(x_r, y_r, 1) = (x_f, y_f, 1) \cdot \begin{pmatrix} s \cdot \cos\alpha & s \cdot \sin\alpha & 0 \\ -s \cdot \sin\alpha & s \cdot \cos\alpha & 0 \\ t_x & t_y & 1 \end{pmatrix} \tag{7}$$

where s is the scale factor, α is the rotation angle, t_x is the translation in the x-axis and t_y is the translation in the y-axis.

In this paper, MMSE is used to calculate the parameters of the affine transformation model.

$$s = \frac{\sum_i^N \sum_j^N \sqrt{(x_{r,i} - x_{r,j})^2 + (y_{r,i} - y_{r,j})^2}}{\sum_i^N \sum_j^N \sqrt{(x_{f,i} - x_{f,j})^2 + (y_{f,i} - y_{f,j})^2}} \tag{8}$$

$$\alpha = \frac{1}{N-1} \sum_{i=1}^N \sum_{j=1}^N \left(\tan^{-1} \frac{y_{r,i} - y_{r,j}}{x_{r,i} - x_{r,j}} - \tan^{-1} \frac{y_{f,i} - y_{f,j}}{x_{f,i} - x_{f,j}} \right) \tag{9}$$

$$t_x = \frac{1}{N} \sum_{i=1}^N \left(x_{r,i} - s \cdot \cos\alpha \cdot x_{f,i} + s \cdot \sin\alpha \cdot y_{f,i} \right) \tag{10}$$

$$t_y = \frac{1}{N} \sum_{i=1}^N \left(y_{r,i} - s \cdot \sin\alpha \cdot x_{f,i} - s \cdot \cos\alpha \cdot y_{f,i} \right) \tag{11}$$

Once the parameters confirmed, a rough registered image is obtained by transforming the floating image using affine transformation with bicubic interpolation. Now the rough registered image is regarded as the floating image so the next step is to register the reference image and the rough registered image.

2.3   IECC-Based Similarity Measure

In this paper, the similarity measure is IECC, which is a new similarity measure based on entropy. For two images R and F, we can calculate the marginal probability

distribution, $p(r_i)$ and $p(f_j)$, the joint probability distribution, $p(r_i, f_j)$, of image R and F. The joint probability distribution $p(r_i, f_j)$ can be obtained simply by normalizing the 2D histogram.

$$p(r_i, f_j) = \frac{h(r_i, f_j)}{\sum_{i=1}^{bin} \sum_{j=1}^{bin} h(r_i, f_j)} \tag{12}$$

where r_i is the intensity of image R, f_j is the intensity of image F, $h(r_i, f_j)$ is the 2D histogram calculated from the two images R and F, *bin* is the size of the 2D histogram. The marginal probability $p(r_i)$ and $p(f_j)$ can be obtained by summing $p(r_i, f_j)$ over *f* and *r*, respectively.

$$p(r_i) = \sum_{j=1}^{bin} p(r_i, f_j) \tag{13}$$

$$p(f_j) = \sum_{i=1}^{bin} p(r_i, f_j) \tag{14}$$

IECC can be defined as:

$$\text{IECC}(R, F) = -\sum_{i=1}^{bin} \sum_{j=1}^{bin} \frac{p(r_i, f_j) \cdot \log_2 \left(p(r_i, f_j) / \left(p(r_i) \cdot p(f_j) \right) \right)}{p(r_i) \cdot \log_2 (r_i) + p(f_j) \cdot \log_2 p(f_j)} \tag{15}$$

IECC is used to determine if the reference image and the floating image are registered. The maximum IECC is determined as the best registration between two images.

3   Experiments

In this section, experiments are conducted to evaluate the performance of the proposed image registration method. There are many similarity measures such as mutual information (MI) [5], normalized mutual information (NMI) [5] and entropy correlation coefficient (ECC) [13]   that can be   combined with SIFT to register the image. The proposed SIFT-IECC image registration method is compared with the SIFT [14] method, the SIFT-MI method [15], the SIFT-NMI method and the SIFT-ECC method. The experiments are implemented in Matlab 2012a using a computer with a CPU of Inter Core i5 (2.5GHz) and 4GB memory.
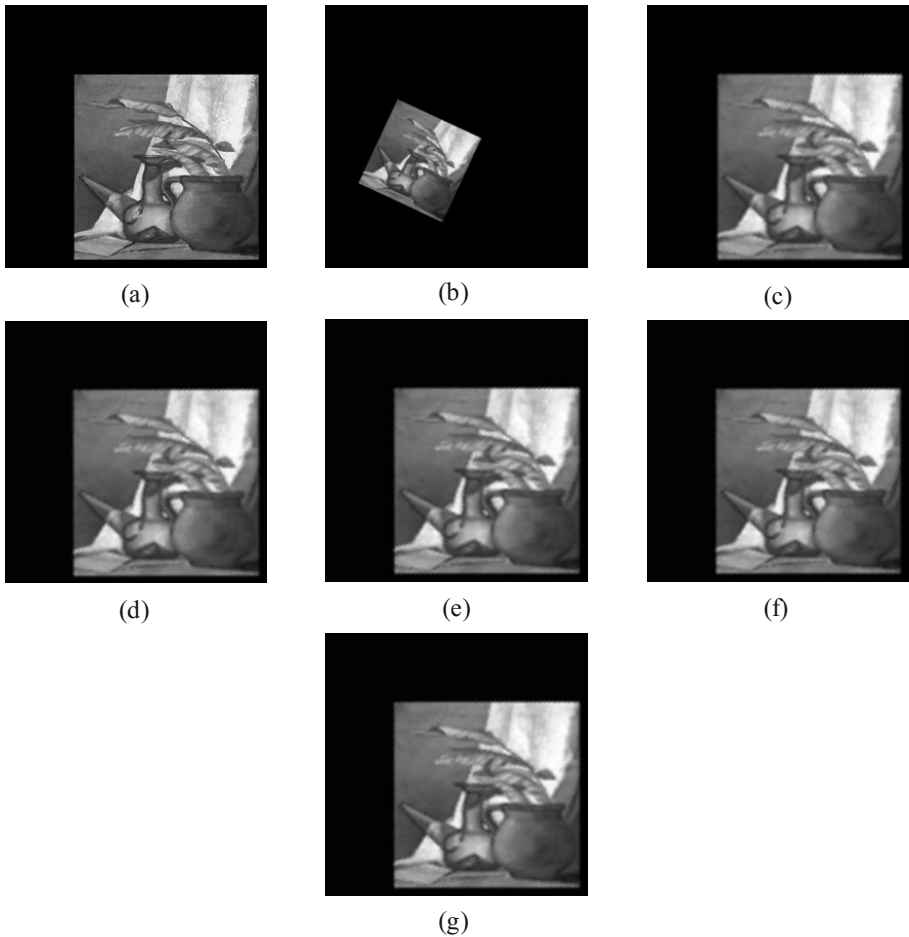
Fig. 2. The image registration results of image 1. (a) Reference   image. (b) Floating image. (c) SIFT registration result. (d) SIFT-MI registration result. (e)   SIFT-NMI registration result. (f) SIFT-ECC registration result. (g) SIFT-IECC registration result.
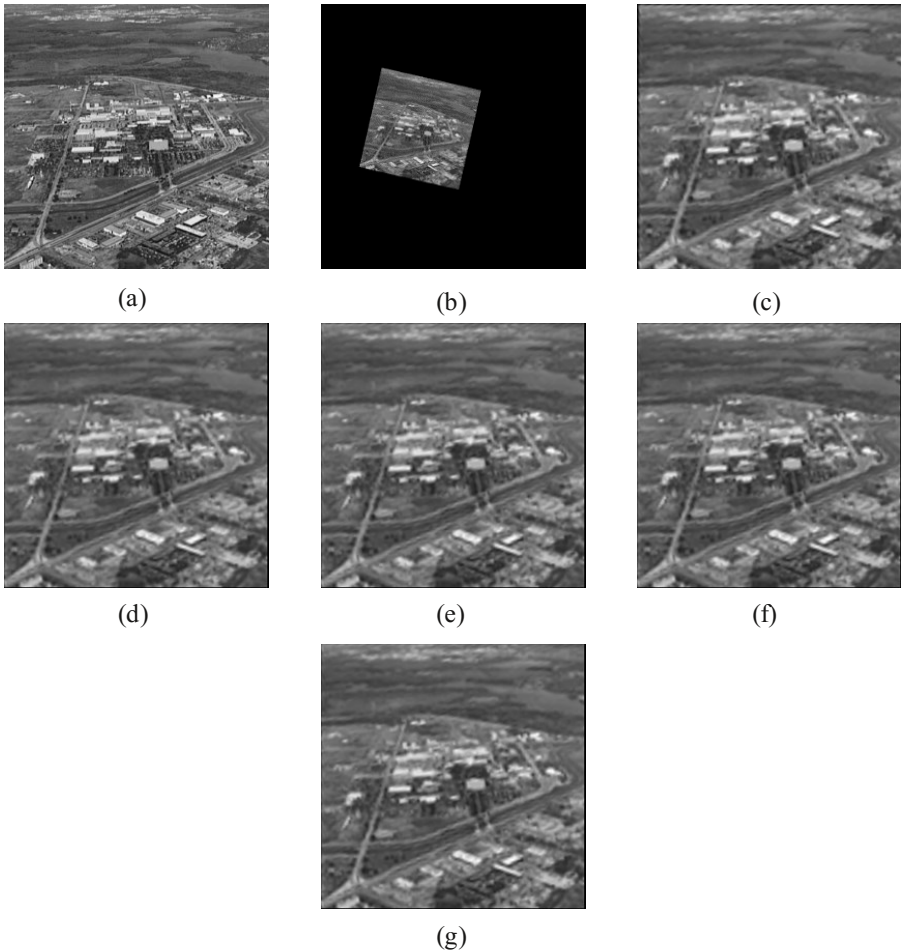
Table 1. The registration results of image 1

Type of Method	t_x (pixel)	t_y (pixel)	α (rad)	s	RMSE (pixel)	Time (second)
SIFT	-137.2329	-39.3152	-0.4356	2.0006	1.6936	85.6430
SIFT-MI	-139.8706	-40.0927	-0.4327	2.0109	0.4917	112.3408
SIFT-NMI	-138.8549	-39.3212	-0.4340	2.0010	0.3272	174.1210
SIFT-ECC	-138.8549	-39.3212	-0.4340	2.0010	0.3272	180.9565
SIFT-IECC	-139.0405	-38.4798	-0.4372	2.0031	0.3192	162.1692

The registration efficiency is assessed by the computation time, and the registration accuracy is assessed by the root mean square error (RMSE) of the detected key points. The smaller RMSE indicates the better registration result. The RMSE between the reference image and the registered image can be defined as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\left(R_{x_i} - Reg_{x_i}\right)^2 + \left(R_{y_i} - Reg_{y_i}\right)^2\right)} \tag{16}$$

where N is the number of key points of the reference image and the registered image, R_{x_i} and R_{y_i} are key points coordinates of the reference image, Reg_{x_i} and Reg_{y_i} are key points coordinates of the registered image.

(a)   (b)   (c)

(d)   (e)   (f)

(g)

Fig. 3. The image registration results of image 2. (a) Reference image. (b) Floating image. (c) SIFT registration result. (d) SIFT-MI registration result. (e) SIFT-NMI registration result. (f) SIFT-ECC registration result. (g) SIFT-IECC registration result.

Table 2. The registration results of image 2

Type of Method	t_x (pixel)	t_y (pixel)	α (rad)	s	RMSE (pixel)	Time (second)
SIFT	-184.1715	-122.3788	-0.2228	2.1743	1.8006	54.2974
SIFT-MI	-185.9644	-122.1512	-0.2228	2.1751	0.3006	223.8533
SIFT-NMI	-185.9236	-122.1541	-0.2229	2.1752	0.2951	183.2656
SIFT-ECC	-185.9236	-122.1541	-0.2229	2.1752	0.2951	182.6300
SIFT-IECC	-185.9054	-122.0525	-0.2229	2.1741	0.2940	183.8825

Two images obtained from www.prenhall.com/gonzalezwoods are used to test the above methods. The reference and floating images are re-sampled to 256x256 for computation convenience. The registration results are shown in Fig. 2, Fig. 3, Tab. 1 and Tab. 2, respectively. In the table, t_x and t_y are the translations in x-axis and y-axis respectively, α is the rotation angle and the s is the scale.

The results demonstrate that the proposed SIFT-IECC image registration method provides a significant improvement in RMSE. Moreover, the processing time of the SIFT-IECC image registration method is comparable to the SIFT-MI method, the SIFT-NMI method and the SIFT-ECC method. It costs more time than the SIFT registration method, because it incorporates the IECC-based similarity measure.

4   Conclusion

In this paper, the SIFT-IECC registration method is proposed. The SIFT algorithm is used to obtain an affine transformation model and get a rough registered image. The reference image and the rough registered image are refined using IECC. The registration results demonstrate that the proposed SIFT-IECC registration method performs better than the SIFT method, the SIFT-MI method, the SIFT-NMI method and the SIFT-ECC method in terms of accuracy. Our future work will focus on improving the efficiency.

References

1. Xing, C., Qiu, P.H.: Intensity-Based Image Registration by Nonparametric Local Smoothing. IEEE Transactions on Pattern Analysis and Machine Intelligence 33(10), 2081–2092 (2011)
2. Oliveira, F.P.M., Tavares, J.: Medical image registration: a review. Computer Methods in Biomechanics and Biomedical Engineering 17(2), 73–93 (2014)

3. Cheng, D., Xie, S.Q., Hammerle, E.: A Robust Local Descriptor Method for Registering Maori Artefacts using Colour Images. In: International Conference on Information and Automation, vols. 1-3. IEEE, New York (2009)
4. Liang, J., Liu, X., Huang, K., Li, X., Wang, D., Wang, X.: Automatic Registration of Multisensor Images Using an Integrated Spatial and Mutual Information (SMI) Metric. IEEE Transactions on Geoscience and Remote Sensing 52(1), 603–615 (2014)
5. Yokoi, T., Soma, T., Shinohara, H., Matsuda, H.: Accuracy and reproducibility of co-registration techniques based on mutual information and normalized mutual information for MRI and SPECT brain images. Annals of Nuclear Medicine 18(8), 659–667 (2004)
6. Xu, H.L., Hua, G.R., Zhuang, J., Wang, S.A.: A Frequency Domain Approach to Fast and Accurate Image Registration. In: International Conference on Information and Automation, vols. 1-3. IEEE, New York (2009)
7. Hurtos, N., Cuf, X., Petillot, Y., Salvi, J., Robotics Society of, J.: Fourier-based Registrations for Two-Dimensional Forward-Looking Sonar Image Mosaicing. In: 25th IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 5298–5305. IEEE (2012)
8. Goshtasby, A.A.: 2-D and 3-D Image Registration: for Medical, Remote Sensing, and Industrial Applications. Wiley (2005)
9. Lee, E.S., Kang, M.G.: Regularized adaptive high-resolution image reconstruction considering inaccurate subpixel registration. IEEE Transactions on Image Processing 12(7), 826–837 (2003)
10. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision 60(2), 91–110 (2004)
11. Itou, T., Shinohara, H., Sakaguchi, K., Hashimoto, T., Yokoi, T., Souma, T.: Multimodal image registration using IECC as the similarity measure. Medical Physics 38(2), 1103–1115 (2011)
12. Lindeberg, T.: Scale-space theory: A basic tool for analyzing structures at different scales. Journal of Applied Statistics 21(1-2), 225–270 (1994)
13. Skerl, D., Likar, B., Fitzpatrick, J.M., Pernus, F.: Comparative evaluation of similarity measures for the rigid registration of multi-modal head images. Physics in Medicine and Biology 52(18), 5587–5601 (2007)
14. Moradi, M., Abolmaesumi, P.: Medical image registration based on distinctive image features from scale-invariant (SIFT) key-points. In: 19th International Congress and Exhibition on Computer Assisted Radiology and Surgery, vol. 1281, pp. 1292–1292. Elsevier (2005)
15. Suri, S., Schwind, P., Reinartz, P., Uhl, J.: Combining mutual information and scale invariant feature transform for fast and robust multisensor SAR image registration. In: Proceedings of the 75 ASPRS Annual Conference (2009)