

# Registration of Images With Outliers Using Joint Saliency Map

Binjie Qin, *Member, IEEE*, Zhijun Gu, Xianjun Sun, and Yisong Lv

**Abstract**—Mutual information (MI) is a popular similarity measure for image registration, whereby good registration can be achieved by maximizing the compactness of the clusters in the joint histogram. However, MI is sensitive to the “outlier” objects that appear in one image but not the other, and also suffers from local and biased maxima. We propose a novel joint saliency map (JSM) to highlight the corresponding salient structures in the two images, and emphatically group those salient structures into the smoothed compact clusters in the weighted joint histogram. This strategy could solve both the outlier and the local maxima problems. Experimental results show that the JSM-MI based algorithm is not only accurate but also robust for registration of challenging image pairs with outliers.

**Index Terms**—Image registration, joint saliency map, mutual information, outliers, weighted joint histogram.

## I. INTRODUCTION

IMAGE registration can be considered as finding the optimal transformation  $T$  between the reference image  $I_R$  and the floating image  $I_F$  to maximize a defined similarity measure such as mutual information (MI). Since 1995 [1], [2], MI has been proved to be very effective in image registration. The MI between  $I_R$  and  $I_F$  (with intensity bins  $r$  and  $f$ ) is defined as:

$$MI = H(I_R) + H(I_F) - H(I_R, I_F) \quad (1)$$

where  $H(I) = -\sum_i p(i) \log p(i)$  and  $H(I_R, I_F) = -\sum_{r,f} p(r, f) \log p(r, f)$  are the entropy of the intensities of image  $I$  and the entropy of the joint intensities of two images,  $p(i)$  is the intensity probabilities with  $p(r) = \sum_f p(r, f)$  and  $p(f) = \sum_r p(r, f)$ ,  $p(r, f)$  is the joint intensity probabilities estimated by the joint histogram  $h(r, f)$ .

MI-based registration methods take advantage of the fact that properly registered images usually correspond to compactly-clustered joint histograms [3]. They measure the joint histogram dispersion by computing the entropy of the joint intensity probabilities. When the images become misregistered, the compact

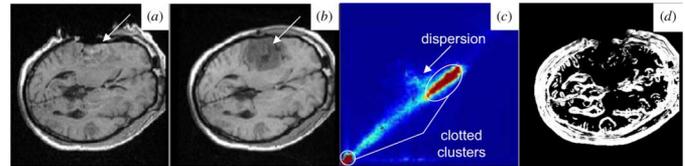


Fig. 1. (a)–(b) Intra-operative and pre-operative MR image with a large tumor resection. (c) Joint histogram dispersion with two clotted clusters (dark red in pseudo color). (d) Joint saliency map for (a) and (b).

clusters become disperse sets of points in the joint histogram and the entropy of the joint intensity probabilities increases. Making no assumptions about the form of the intensity mapping between the two images, MI is sensitive to the unmatched outliers, e.g. the tumor resection in the intra- and pre-operative brain images [see Fig. 1(a)–(b)]. To reject the outliers, some approaches are proposed including consistency test [4], intensity transformation [5], gradient-based asymmetric multifeature MI [6] and graph-based multifeature MI [7]. However, all these methods do not emphasize the corresponding salient structures in the two images to suppress the outliers. Furthermore, MI likely suffers from local and biased maxima [8] which are caused by the ambiguities in defining structure correspondence.

Spatial information, i.e. the dependence of the intensities of neighboring pixels, has been included in MI [9]–[12] to improve registration. Nevertheless, almost all MI-based methods equally treat each overlapping pixel pair as a separate point in the overlap area to calculate the joint histogram. This could raise three issues: 1) when we equally consider the outlier pixel pairs, the noncorresponding structures overlap and the histogram will show certain clusters for the grey values of the outliers. These clusters easily introduce the histogram dispersion [see Fig. 1(c)] with increasing misregistration; 2) while registration can be achieved by maximizing the compactness of the histogram, the undesired clotted clusters [see Fig. 1(c)] related to many noisy pixel pairs in the structureless regions, such as background and white matter in the brain image, increase the MI ambiguities and the local maxima [8] (Fig. 5(c) shows that the normalized MI [1], [20] is in a biased global maximum when the whole background areas in the two endoscopic images are exactly aligned); 3) when we group the intensity pairs as separate points into the histogram, the independence of the neighboring bins could increase the MI ambiguities and the local maxima. To solve this problem, joint histogram smoothing (or blurring) [5], [8] has been used to increase the dependence of the neighboring histogram bins. We address these issues above as follows.

In fact, image registration is to match the corresponding salient structures in both images. To suppress the outliers and the homogeneous pixel pairs, the corresponding pixel pairs in the corresponding salient structures should contribute more to the joint histogram. For example, the corresponding salient

Manuscript received August 08, 2009; revised September 27, 2009. First published October 06, 2009; current version published October 28, 2009. This work was supported in part by the NSFC (60872102), NBRPC (973 Program 2010CB834303), the Science Foundation of Shanghai Municipal Science & Technology Commission (04JC14060), Shanghai Municipal Health Bureau (2008115), and the Small Animal Imaging Project (06-545). The associate editor coordinating the review of this manuscript and approving it for publication was Prof. H. Vicky Zhao.

B. Qin, Z. Gu, and X. Sun are with the Department of Biomedical Engineering, School of Life Sciences & Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: bj Qin@sjtu.edu.cn; gzj0126@gmail.com; sxj\_sun@sjtu.edu.cn).

Y. Lv is with the Department of Mathematics, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: yslv@sjtu.edu.cn).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/LSP.2009.2033728

pixel pairs in the normal brain tissues should be given more weight in the histogram than the homogeneous and the tumor resection pixel pairs. To weight each overlapping pixel pair when computing the joint histogram, we propose a novel joint saliency map (JSM) to assign a joint saliency value between 0 and 1 to the pixel pair. The idea of JSM is demonstrated schematically in Fig. 1(d), where the high joint saliency values are assigned to the corresponding salient pixel pairs rather than the outlier and the homogeneous pixel pairs.

The JSM is determined by correlating each overlapping pixel pair's respective regional saliency vectors (RSVs). The RSV characterizes the regional salient structure around each underlying pixel after a principal axis analysis (PAA) of the pixel's regional saliency distribution. In the JSM-weighted joint histogram (WJH), the contributions of the corresponding salient structures are distributed over neighboring histogram bins. This leads to the smoothing of the compact clusters for the grey values of the corresponding salient structures, which can solve both the outlier and the local maxima problems.

The proposed JSM-MI has been applied to the rigid registration of 2-D images. Experimental results show that, compared to other MI-based registration methods, JSM-MI method achieves better robustness and higher accuracy for the registration of challenging image pairs with outliers. The letter is organized as follows. We first introduce the JSM for WJH in MI. Next, we report some experiment results to identify the registration performance on accuracy and robustness. Finally, the conclusions close this letter.

## II. METHODS

### A. Regional Saliency Vector

We use visual saliency operator to enhance the regional salient structures we are interested in. Many techniques have been developed to define the saliency of image, i.e., using edge gradient, local phase [12], salient regions [13], corner and keypoints [14]. Gradient map has been incorporated into the MI-based registration methods [9]–[11]. However, gradient is a local feature and sensitive to noise. Local phase [12] and salient regions [15] suffer from high computational complexity. Corner and keypoint can not be defined for each image pixel. Inspired by the center-surround mechanism [16], [17] which has defined the intensity-contrast-based visual saliency map, we define a two-step scale and rotation invariant saliency operator based on intensity contrast as follows:

$$S_l(v) = \sum_{u \in N_v} (I_l(v) - I_l(u))^2 \quad (2)$$

where  $N_v$  is the 1-pixel radius circular neighborhood of the pixel position  $v = (x, y)$  at scale  $l$ ,  $S_l(v)$  is the local saliency computed for the intensity  $I_l(v)$  in the Gaussian image pyramid [18] at scale  $l$ ,  $I_l(u)$  is the intensity of the pixel in the  $I_l(v)$ 's neighborhood. The multiscale local saliency map  $S(x, y)$  at the finest scale is reconstructed by summing up all the saliency maps at the coarser scales.

In the second step, a PAA of the saliency distribution in a certain region assigns *regional saliency* to each pixel based on the inertia matrix:

$$\mathbf{M} = \begin{bmatrix} \mu_{20} & \mu_{11} \\ \mu_{11} & \mu_{02} \end{bmatrix} \quad (3)$$

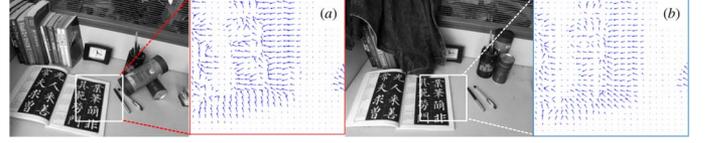


Fig. 2. (a)–(b) RSVs for the subblocks in the reference and the floating images (size:  $400 \times 300$  pixels).

where  $\mu_{jk} = \sum (x - g_x)^j (y - g_y)^k S(x, y)$ ,  $(g_x, g_y) = (m_{10}/m_{00}, m_{01}/m_{00})$  and  $m_{jk} = \sum x^j y^k S(x, y)$  are the central  $(j, k)$ -moment, the centroid and the  $(j, k)$ -moment of the saliency distribution  $S(x, y)$  in the 5.5-pixel radius circular neighborhood around each pixel. This regional saliency distribution describes a 2-D regional salient structure. The two eigenvectors of the matrix  $\mathbf{M}$  represent the orthogonal coordinate system within the regional salient structure, while the corresponding eigenvalues give information about the length of the respective axes. Because the regional information about the orientation of the salient structure is mostly stored along the first eigenvector corresponding to the largest eigenvalue, the first eigenvector referred as the RSV is enough to represent the regional salient structure around a pixel [see Fig. 2(a)–(b)].

### B. Joint Saliency Map

Given two RSVs of each overlapping pixel pair, JSM is ready to describe the matching degree between the two RSVs. The inner product of two RSVs is a measure of their co-linearity and naturally can be used as their similarity measure. The essential idea of JSM is an assumption, which is always valid in practice according to the empirical experience in image registration: for two precisely aligned multimodal (or multitemporal) images, the majority of the corresponding pixel locations are very likely to produce the RSVs with similar orientations [see Fig. 2(a)–(b)]. This is because the two images under registration fundamentally depict the same image structures. As a result, the RSVs of the corresponding pixel locations from two images could present relatively coincident orientations in general. Therefore, the angle  $\theta$  between the two RSVs ( $\mathbf{x}_R, \mathbf{x}_F$ ) is simply calculated, making  $\cos \theta$  the scalar measure of the joint saliency value  $w(v)$ :

$$w(v) = \cos \theta(\mathbf{x}_R, \mathbf{x}_F) = \langle \mathbf{x}_R, \mathbf{x}_F \rangle / \|\mathbf{x}_R\| \cdot \|\mathbf{x}_F\| \quad (4)$$

A JSM value near one suggests that the underlying pixel pair originates from the corresponding salient structures. Contrarily, a JSM value near zero indicates that the underlying pixel pair comes from either the outliers or a homogeneous region. To speed up the registration without reducing accuracy, the pixel with a small saliency value below a threshold value (10% of the maximum saliency value) is assigned a zero JSM value directly. The JSM would primarily respond to the high-gradient edge pixels if a high threshold value is chosen. However, the JSM does not simply emphasize the common image gradients in the two images. Fig. 3(d)–(f) present the image gradient and the JSM profiles of the same line (marked as dashed lines across the tumor areas) at the two registered images [see Fig. 3(a)–(b)]. As shown in the figures, the image gradient features in Fig. 3(d)–(e) are very noisy and do not agree with each other at each overlapping location. The JSM in Fig. 3(f) can accurately preserve the corresponding salient structures in larger capture range with smaller variability than the image gradients.

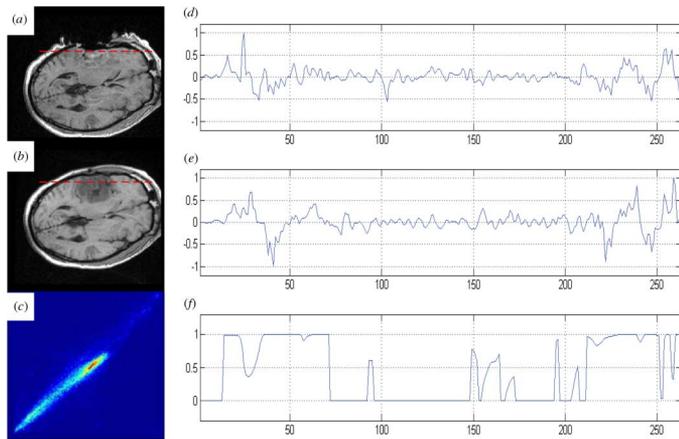


Fig. 3. (a)–(b) The reference and the floating images for the gradient magnitude and the JSM magnitude. (d) Compact JSM-WJH smoothing for (a)–(b). (d)–(e) Gradient value profiles of the lines in (a)–(b), which are marked as dashed lines. (f) JSM value profiles of the lines in (a)–(b).

### C. JSM-Weighted Joint Histogram

The contribution of the interpolated floating intensity  $f(v_f)$  to the joint histogram is weighted by a  $w(v)$  of the JSM (the pixel positions  $(v_r, v_f)$  are overlapped at the position  $v$ ). For 2-D image registration, if using a nearest neighbor or a bilinear interpolation, the value  $w(v)$  should be added to the histogram entry  $h(r, f)$ . In bilinear partial volume distribution (PV) interpolation, the contribution of the  $f(v_f)$  to the histogram, distributed over the intensity values of all nearest neighbors of the reference pixel position  $v_r$  on the grid of  $I_R$ , is weighted using the  $w(v)$ . Similarly, JSM could be easily incorporated into other interpolation schemes and Parzen-based joint histogram.

In the JSM-WJH, the outliers and homogeneous regions have little impact on the histogram distribution. Furthermore, each histogram entry for the corresponding salient structures is the sum of smoothly varying fractions of one, such that the histogram changes smoothly in the neighboring bins related to those structures. As a result, the compact histogram smoothing [see Fig. 3(c)] is introduced by highlighting the grey values of the corresponding salient structures. Computed from the compact and smooth histogram, the MI is then maximized to achieve robust and accurate rigid registration.

### D. Computational Complexity

The JSM should be recalculated with the transformation changing the overlap area at each registration iteration. The RSV orientation for a JSM calculation could be easily reoriented as it is done in the diffusion tensor image registration [19]. Nevertheless, to ensure the numerical stability and the computation speedup, a new JSM at each iteration can be simply updated from the JSM of the previous iteration through the PV interpolation. The JSM could be re-calculated after  $n$  iterations ( $n = 10 \sim 15$ ) to reflect the updated correspondence between the salient structures in the two images.

## III. EXPERIMENTAL RESULTS

We evaluated our JSM-MI-based (JMI) algorithm on 11 challenging image pairs including CT-PET tumor images, MR brain tumor resection images, optical images with background/foreground clutter and etc. We implemented the JMI algorithm using the simplex optimization in a multiresolution scheme [18]. The

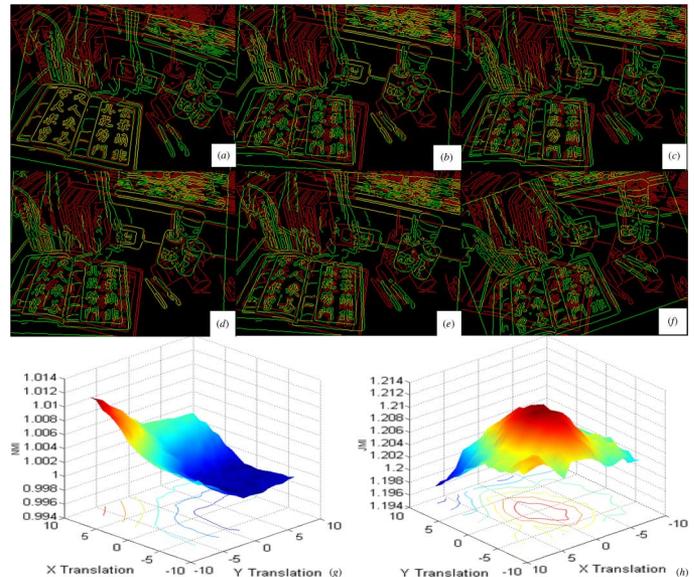


Fig. 4. Registration results for the two images in Fig. 2(a)–(b). (a) JMI. The yellow contour overlap of the book validates the registration accuracy owing to the additive mixing of red and green. (b) NMI. (c) RMI. (d) HMI. (e) GMI. (f) PMI. (g)–(h) NMI and JMI similarity surfaces plotted as a function of  $x$  and  $y$  translation (within a range of  $\pm 10$  pixels around the matching position).

TABLE I  
REGISTRATION RESULTS FOR FIGS. 4 AND 5 (THE TRANSLATIONS  $X$  AND  $Y$  ARE IN PIXELS IN THE  $x$  AND  $y$  DIRECTIONS, THE ROTATION  $\beta$  IS IN DEGREES AROUND THE CENTER OF THE IMAGES)

Cases	Correct( $X, Y, \beta$ )	Computed( $X, Y, \beta$ )
1	$-23.11, 45.59, 11.43^\circ$	$-22.34, 45.30, 11.03^\circ$
2	$37.91, -36.78, 4.43^\circ$	$37.46, -38.18, 4.68^\circ$

algorithm stops if the current step length is smaller than  $10^{-5}$  or if it has reached the limit of 200 evaluation numbers. The challenging image pairs include some complex outliers that the normalized MI-based method and four of MI-based adaptations with incorporating spatial information fail to deal with. Due to space restrictions, we only show some typical experimental results in this letter.

Fig. 4 shows the various registration results for the two images at Fig. 2 with a foreground book and the large changes of background appearance. To facilitate the visual assessment of registration accuracy, the green floating contours and the red reference contours obtained by Canny–Deriche edge detector have been overlaid over each other. The subpixel registration accuracy (see Table I, case 1) of our JMI algorithm can be validated by the book’s yellow contour overlap, which is due to the additive color mixing of the green and the red contour [see Fig. 4(a)].

Using particle swarm optimization (PSO) to deal with the local maxima, the other methods based on normalized MI (NMI) [1], [20], regional MI (RMI) [21], high-dimensional MI (HMI) [22], MI with gradient information (GMI) [10], and phase MI (PMI) [12] show different misregistration results in Fig. 4(b)–(f). The PSO is conducted with 20 particles and allowed to experience 200 evaluation iterations. The algorithm stops if it has reached the limit of evaluation numbers or if the minimum error ( $10^{-5}$ ) conditions is satisfied. The computation time needed for the different algorithms are listed in Table II.

Fig. 4(g)–(h) show that the NMI and JMI similarity surfaces are plotted as a function of  $x$  and  $y$  translation. In this case, the

TABLE II  
COMPUTATION ITERATIONS AND RUNTIME IN SECONDS FOR FIG. 4.  
(MATLAB 6.5, SINGLE CORE INTEL CELERON 2.8 GHZ, RAM 2 GB)

	JMI	NMI	RMI	HMI	GMI	PMI
Iter.	64	41	45	46	50	29
Time	157.4	296.7	297.1	1060.1	329.1	3049.3

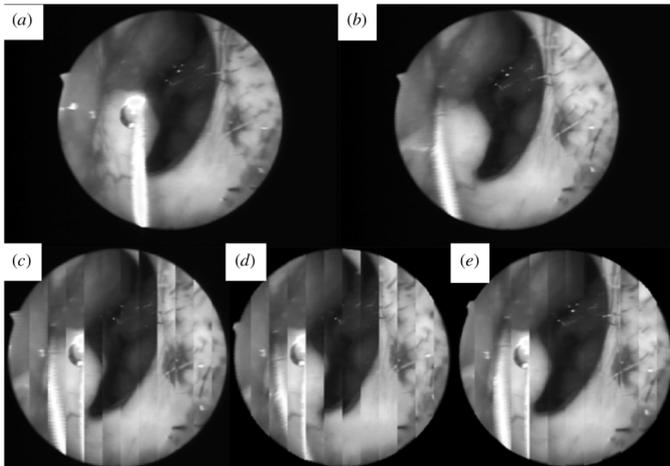


Fig. 5. (a)–(b) Reference and floating endoscopic images (size:  $720 \times 572$  pixels) with a surgical tool and illumination changes. The two images are fused using a mosaic pattern. (c) NMI. (d) PMI. (e) JMI.

JSM removes all local maxima and achieves the global maximum at the registration position, while the NMI suffers from the biased maximum at the mismatching position.

Fig. 5(a)–(b) show the reference and floating endoscopic images ( $720 \times 572$  pixels) including a surgical instrument with different illuminations. Using a mosaic pattern to fuse the two images, Fig. 5(c)–(d) show the NMI-based and PMI-based misregistration results. Fig. 5(e) shows our accurate JMI-based registration result (see Table I. case 2).

#### IV. CONCLUSION

We propose an effective JSM to solve the problems of outliers and local maxima in MI-based image registration. Representing the corresponding salient structures in the two images to be registered, JSM is easily integrated into other intensity-based similarity measures for 3-D nonrigid registration. Independent of this work but subsequent to our preliminary conference papers [23], [24] which this letter elaborates on and extends, Ou *et al.* [25] developed a similar mutual saliency map for outlier rejection in 3-D nonrigid image registration.

Additionally, our method is an intensity-based method and also sensitive to the initial conditions. It is necessary in principle to set the proper initial conditions close to a correct alignment solution, which can be achieved by coarse alignment techniques such as principal axes based method. Nevertheless, all instances of correct registration in this letter are directly performed by our method without any coarse alignment.

#### ACKNOWLEDGMENT

The authors thank S. K. Warfield, M. Irani, R. Barnett and E. Vrscay for allowing the use of image, R. Ali for the phase

recovery source code, S. Tong and all reviewers for their useful comments, and W. Wang for her help to our algorithm.

#### REFERENCES

- [1] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Trans. Med. Imag.*, vol. 16, no. 2, pp. 187–198, 1997.
- [2] P. Viola and W. M. Wells, III, "Alignment by maximization of mutual information," *Int. J. Comput. Vis.*, vol. 24, no. 2, pp. 137–154, 1997.
- [3] J. Orchard, "Multimodal image registration using floating regressors in the joint intensity scatter plot," *Med. Image Anal.*, vol. 12, no. 4, pp. 385–396, 2008.
- [4] M. Auer, P. Regitnig, and G. A. Holzzapfel, "An automatic nonrigid registration for stained histological sections," *IEEE Trans. Image Process.*, vol. 14, no. 4, pp. 475–486, 2005.
- [5] P. Likar and F. Pernus, "A hierarchical approach to elastic registration based on mutual information," *Image Vis. Comput.*, vol. 19, no. 1–2, pp. 33–44, 2001.
- [6] D. Tomazevic, B. Likar, and F. Pernus, "3-D/2-D registration by integrating 2-D information in 3-D," *IEEE Trans. Med. Imag.*, vol. 25, no. 1, pp. 17–27, 2006.
- [7] M. Staring, U. A. van der Heide, S. Klein, M. A. Viergever, and J. P. W. Pluim, "Registration of cervical MRI using multifeature mutual information," *IEEE Trans. Med. Imag.*, vol. 28, no. 9, pp. 1412–1421, 2009.
- [8] J. Tsao, "Interpolation artifacts in multimodality image registration based on maximization of mutual information," *IEEE Trans. Med. Imag.*, vol. 22, no. 7, pp. 854–864, 2003.
- [9] R. Gan, A. C. S. Chung, and S. Liao, "Maximum distance-gradient for robust image registration," *Med. Image Anal.*, vol. 12, no. 4, pp. 452–468, 2008.
- [10] J. P. W. Pluim, J. B. Maintz, and M. A. Viergever, "Image registration by maximization of combined mutual information and gradient information," *IEEE Trans. Med. Imag.*, vol. 19, no. 8, pp. 809–814, 2000.
- [11] R. Bhagalia, J. A. Fessler, and B. Kim, "Accelerated nonrigid intensity-based image registration using importance sampling," *IEEE Trans. Med. Imag.*, vol. 28, no. 8, pp. 1208–1216, 2009.
- [12] M. Mellor and M. Brady, "Phase mutual information as a similarity measure for registration," *Med. Image Anal.*, vol. 9, no. 4, pp. 330–343, 2005.
- [13] T. Kadir and M. Brady, "Saliency, scale and image description," *IJCV*, vol. 45, no. 2, pp. 83–105, 2001.
- [14] D. G. Lowe, "Distinctive image feature from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.
- [15] H. Luan, F. Qi, Z. Xue, L. Chen, and D. Shen, "Multimodality image registration by maximization of quantitative-qualitative measure of mutual information," *Pattern. Recognit.*, vol. 41, no. 1, pp. 285–298, 2008.
- [16] L. Itti and C. Koch, "Computational modeling of visual attention," *Nature Rev. Neurosci.*, vol. 2, no. 3, pp. 194–203, 2001.
- [17] F. Liu and M. Gleicher, "Region enhanced scale-invariant saliency detection," in *IEEE ICME 2006*, 2006, pp. 1477–1480.
- [18] M. Jenkinson and S. Smith, "A global optimization method for robust affine registration of brain images," *Med. Image Anal.*, vol. 5, no. 2, pp. 143–156, 2001.
- [19] D. C. Alexander, C. Pierpaoli, P. J. Basser, and J. C. Gee, "Spatial transformation of diffusion tensor magnetic resonance images," *IEEE Trans. Med. Imag.*, vol. 20, no. 11, pp. 1131–1139, 2001.
- [20] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," *Pattern Recognit.*, vol. 32, no. 1, pp. 71–86, 1999.
- [21] D. Russakoff, C. Tomasi, T. Rohlfing, and C. Maurer, "Image similarity using mutual information of regions," in *ECCV 2004*, 2004, pp. 596–607.
- [22] J. Kybic, "High-dimensional mutual information estimation for image registration," in *IEEE ICIP 2004*, 2004, pp. 1779–1782.
- [23] Z. Gu and B. Qin, "Multi-modal and multi-temporal image registration in the presence of gross outliers using feature voxel-weighted normalized mutual information," in *IEEE NSS-MIC 2006*, 2006, pp. 3209–3212.
- [24] B. Qin and Z. Gu, "Robust adaptive non-rigid image registration based on joint salient point sets in the presence of tumor-like gross outlier," in *Proc. SPIE*, 2007, vol. 6833, pp. 683320-1–683320-11.
- [25] Y. Ou and C. Davatzikos, "DRAMMS: Deformable registration via attribute matching and mutual-saliency weighting," in *IPMI 2009*, 2009, vol. 5636, LNCS, pp. 50–62.