Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.7766126

COHERENT PIXEL CORRESPONDENCE FOR IMAGE REGISTRATION

B. Sudhir[1], P. Rajesh Kumar[2]
[1]Electronics & Communications Engineering Department, Rajamahendri Institute of Engineering and Technology, AP, India
[2]Electronics & Communications Engineering Department, Andhra University, Vishakapatnam, AP, India

Abstract—Estimating a true set of matching corresponding pixels has always been a challenge for image registration. This paper formulates the alignment problem as an estimation problem using expectation-maximization (EM) algorithm. With a given input of set of correspondence points and their distribution function a posterior probabilities of correspondence can be estimated and maximized until convergence. A simple thresholding can then distinguish between the inliers and outliers which can recover the coherent pixel correspondence between the image pairs. The experimental results have shown efficient results for images with rigid motion of translation and rotation.

Keywords—image registration, expectation, maximum-likelihood, point correspondence, coherence

Introduction

Image registration for commercial images has garnered a lot of importance in image processing techniques lately after its immense applications in many field like medical imaging, satellite imaging, image listing etc. while registration techniques differ based on modality of the images captured, a fully automated image registration have always found its own challenges. Image registration has been found to be more challenging especially when the image pairs used for registration have undergone any rigid or non-rigid motions. This paper describes a method for image registration for image pairs whose set of point correspondence is apriori given and have undergone rigid motion of translation and rotation. The point correspondence set of points are chosen such a way that they include most of the true matches between the image pairs. Hence an inlier set and an outlier set can be defined based on validation of the matching points. A strong discriminator which can estimate these set of points are hence desirable to achieve a reliable image registration.

Many estimating techniques have been used in various image processing techniques in the recent past. With the advance of computer vision and machine learning these estimators have become popular among research community in estimating the unknown parameters from a distribution of data points[1][2]. Maximum-likelihood [3][4][5]and least median of squares [6][7][8]are two techniques that are widely and popularly used in any statistical approach. [9]uses ML technique on mutual information in improving the speed of image registration. [10]has proposed a ML technique for joint image registration and fusion by formulating both these image processing problems as estimation problems. The performance in the fusion step is used to evaluate the accuracy of the image registration thereby optimizing both the processes simultaneously.

A penalized maximum likelihood (PML) function is used in [11] for medical image registration. The transition probabilities of corresponding image pixels in the image pairs under question are used to define this PML and is measured by a joint histogram[12][13]. ML framework has also been used for medical image registration of images that have undergone simultaneous motion as in [14].

Following estimation methods, some works have also explored the possibility of using expectation maximization for robust and accurate results[15][16][17]. [18]incorporates fixed and adaptive bin sizes to estimate the probability distribution of mutual information of retinal images for registration. They also used EM for their final step to increase the accuracy of their process. The geometric transformation parameters are estimated using EM in [19] to align multimodal and multispectral images. The approach has shown efficient image registration results where joint conditional multispectral intensity distribution function has been iteratively updated using EM algorithm.

Establishing a coherent spatial correlation[20][21][22] with the set of points of correspondence is vital in our approach for image registration. This relationship can distinguish well between the inliers and outliers and thereby establishing true match between the pixels of images to be registered. A simple thresholding from empirical data have been used to finally compute the inlier set after EM approach. The experimental results have shown significant efficient image registration results for a data set of different commercial rgb images.

Coherent Pixel Correspondence Of Image Registration Theory

Expectation Maximization

Point to point correspondence of pixels have long been used for image registration. Our approach is to remove points without any correspondence from a given set of point correspondences to establish accurate and reliable image alignment. Towards this regard, first a set of point correspondences S is defined from the image pairs as given below

$$S = \{(\boldsymbol{a}_n, \boldsymbol{b}_n)\}_{n=1}^{N}$$

Where an image point set is defined as $\boldsymbol{a} = (\boldsymbol{a}^x, \boldsymbol{a}^y, 1)^T$ in a homogenous coordinate system. We then try to find a spatial correspondence within the set S between all the image points as that lies within the correspondence which we will call as inliers and those that do not have any correspondence whatsoever as outliers. The objective of the problem formulation then becomes to remove as many outliers as possible and thus include only the correct point set with maximum correspondence as possible. For this a correspondence function $\boldsymbol{C}(\cdot)$forall the image points should be defined. The inlier points and the outlier points can then be identified using this correspondence function such that $\boldsymbol{C}(\boldsymbol{a}, \boldsymbol{b}) = 0$ is satisfied for all the inliers.

With correspondence function defined for the point set correspondence, estimating the correspondence function $\boldsymbol{C}(\cdot)$becomes the first and major task in our approach. An EM framework[23][24] is used to solve this problem by formulating a maximum likelihood for the correspondence function. The assumption that we make here is that the outliers in the image point set is caused by some noise which is Gaussian and have a zero mean and a uniform standard deviation σ. Also the outlier points are assumed to be of uniform distribution $\frac{1}{k}$ where

k is a constant. If we denote the percentage of inlier points as δ then the likelihood can be formulated as

$$\boldsymbol{p(S|\theta)} = \prod_{n=1}^{N} \frac{\boldsymbol{\delta}}{\boldsymbol{2\pi\sigma^2}} \boldsymbol{e}^{-\frac{\|C(a_n,b_n)\|^2}{2\sigma^2}} + \frac{\boldsymbol{1-\delta}}{\boldsymbol{k}}$$

1

Where $\boldsymbol{\theta} = \{\boldsymbol{C}, \sigma^2, \delta\}$ are the parameters to be estimated. Then the true set of parameters $\boldsymbol{\theta}$ maximizes the likelihood function given in $\boldsymbol{p(S|\theta)} = \prod_{n=1}^{N} \frac{\delta}{2\pi\sigma^2} \boldsymbol{e}^{-\frac{\|C(a_n,b_n)\|^2}{2\sigma^2}} + \frac{1-\delta}{k}$
1. This is possible by defining a maximum likelihood for $\boldsymbol{\theta}$as

$$\boldsymbol{\theta^*} = \boldsymbol{arg\, max_\theta p(S|\theta)}$$

2

Alternatively we can write the likelihood in terms of an energy function as

$$E(\boldsymbol{\theta}) = -\sum_{n=1}^{N} ln\, p(\boldsymbol{a}_n, \boldsymbol{b}_n|\boldsymbol{\theta})$$

Of the many parameter estimation models available in the literature we have chosen E-M algorithm for our approach. This is an iterative process where the expectation step uses the current estimate of the parameters to find the expectation of the log-likelihood and the maximization step updated these parameters so as to maximize the expected log-likelihood. A latent variable z_n for each sample in the calculation is assigned based on whether its distribution function is Gaussian or uniform. We have assigned $z_n = 1$ for Gaussian and $z_n = 0$ for uniform distribution of the samples. The E-M problem can now be formulated as

$$\begin{aligned}\boldsymbol{Q(\theta, \theta^{old})} = &-\frac{\boldsymbol{1}}{\boldsymbol{2\sigma^2}} \sum_{n=1}^{N} \boldsymbol{P(z_n = 1|a_n, b_n, \theta^{old})\|C(a_n, b_n)\|^2} \\ &- \boldsymbol{ln\, \sigma^2} \sum_{n=1}^{N} \boldsymbol{P(z_n = 1|a_n, b_n, \theta^{old})} + \boldsymbol{ln\, \delta} \sum_{n=1}^{N} \boldsymbol{P(z_n = 1|a_n, b_n, \theta^{old})} \\ &+ \boldsymbol{ln(1-\delta)} \sum_{n=1}^{N} \boldsymbol{P(z_n = 0|a_n, b_n, \theta^{old})}\end{aligned}$$

3

Equation $\boldsymbol{Q(\theta, \theta^{old})} = -\frac{1}{2\sigma^2}\sum_{n=1}^{N} \boldsymbol{P(z_n = 1|a_n, b_n, \theta^{old})\|C(a_n, b_n)\|^2} - \boldsymbol{ln\, \sigma^2} \sum_{n=1}^{N} \boldsymbol{P(z_n = 1|a_n, b_n, \theta^{old})} + \boldsymbol{ln\, \delta} \sum_{n=1}^{N} \boldsymbol{P(z_n = 1|a_n, b_n, \theta^{old})} + \boldsymbol{ln(1-\delta)} \sum_{n=1}^{N} \boldsymbol{P(z_n = 0|a_n, b_n, \theta^{old})}$
3 can be minimized to a weighted minimal correspondence function as

$$\boldsymbol{Q(C)} = \sum_{n=1}^{N} \boldsymbol{p_n\|C(a_n, b_n)\|^2}$$

4

Expectation Step: by applying Bayes' rule the probability for the n^{th} sample $p_n = P(z_n = 1|\boldsymbol{a}_n, \boldsymbol{b}_n, \boldsymbol{\theta}^{old})$ can be computed as follows

$$\boldsymbol{p_n} = \frac{\boldsymbol{\delta e}^{-\frac{\|\boldsymbol{C}(\boldsymbol{a_n},\boldsymbol{b_n})\|^2}{2\boldsymbol{\sigma}^2}}}{\boldsymbol{\delta e}^{-\frac{\|\boldsymbol{C}(\boldsymbol{a_n},\boldsymbol{b_n})\|^2}{2\boldsymbol{\sigma}^2}} + \frac{\boldsymbol{2\pi\sigma^2(1-\delta)}}{\boldsymbol{k}}}$$

5

Maximization Step: the estimation parameter$\boldsymbol{\theta}$is updated in this step to$\boldsymbol{\theta}^{\boldsymbol{new}}$as follows

$$\boldsymbol{\theta}^{\boldsymbol{new}} = \boldsymbol{arg\,max_{\theta}\,Q}\left(\boldsymbol{\theta}, \boldsymbol{\theta}^{\boldsymbol{old}}\right)$$

6

The maxima is obtained by taking the derivative of the $Q(\cdot)$ with respect to the variance σ^2 and δ and then setting each derivative to zero. The updated variance and percentage of inliers will be then

$$\boldsymbol{\sigma^2} = \frac{\boldsymbol{Q(C)}}{\sum_{\boldsymbol{n=1}}^{\boldsymbol{N}} \boldsymbol{p_n}}$$

7

$$\boldsymbol{\delta} = \sum_{\boldsymbol{n=1}}^{\boldsymbol{N}} \frac{\boldsymbol{p_n}}{\boldsymbol{N}}$$

8

The correspondence function $\mathbf{C}(\cdot)$ can be obtained as the EM steps convergefinally. This is explained in the next section. The set of pixels that are coherent from the correspondence function can then be binned into a set by thresholding. The posterior probabilities are expected to be practically either 0 or 1 and hence the choice of threshold is very easy to make. These set of pixels can be denoted as $\mathcal{P} = \{n|p_n > \tau, n = 1, \dots, N\}$, where $\tau$is the threshold chosen.

Correspondence function

The correspondence function is estimated for image pairs that have undergone only two kinds of rigid motions. Our research has concentrated only on translational and rotational motions of the images. We define a constraint for both these motions. For any general translation motion the correspondence function can be written as

$$\boldsymbol{C}(\boldsymbol{a}, \boldsymbol{b}) = \boldsymbol{a}^{\boldsymbol{T}}\boldsymbol{\mathcal{T}}\boldsymbol{b}$$

9

Where $\boldsymbol{\mathcal{T}}$ is a matrix that defines any translation motion that might have happened to the image pair. Similarly for a rotational motion the correspondence function is

$$\boldsymbol{C}(\boldsymbol{a}, \boldsymbol{b}) = \boldsymbol{b} - \boldsymbol{\mathcal{R}}\boldsymbol{a}$$

10

Where $\boldsymbol{\mathcal{R}}$ is a matrix that defines rotation.From$\boldsymbol{Q}(\boldsymbol{C}) = \sum_{\boldsymbol{n=1}}^{\boldsymbol{N}} \boldsymbol{p_n}\|\boldsymbol{C}(\boldsymbol{a_n}, \boldsymbol{b_n})\|^{\boldsymbol{2}}$
4, 11 can be rewritten as

$$\boldsymbol{Q}(\boldsymbol{\mathcal{T}}) = \sum_{n=1}^{N} \boldsymbol{p_n}(\boldsymbol{a_n}^{T}\boldsymbol{\mathcal{T}}\boldsymbol{b_n})^2 \qquad 11$$

We observe that the translational matrix $\boldsymbol{\mathcal{T}}$ has a rank of 2 in two dimensional plane and can be diagonalised with positive entries. Whenever the posterior probabilities are estimated to be zero, it corresponds to outliers and in such cases the point correspondence set will not be used for estimating the translation motion. This is apparent since we are only interested in the pixels that experience any motion.

The weighted minimal correspondence function for any rotation motion can be similarly computed as follows

$$\boldsymbol{Q}(\boldsymbol{\mathcal{R}}) = \sum_{n=1}^{N} \boldsymbol{p_n}\|\boldsymbol{b_n} - \boldsymbol{\mathcal{R}}\boldsymbol{a_n}\|^2 \qquad \mathbf{14}$$

The same logic for posterior probabilities as explained for translation is applicable here as well. The entire process we have followed for image registration has been summarized in **Error! Reference source not found.**.

Table 1

Algorithm 1 Image Registration Algorithm through coherent pixel correspondence
Input: Set of point correspondence $S = \{(\boldsymbol{a}_n, \boldsymbol{b}_n)\}_{n=1}^{N}$,uniform distribution parameter k; **Output:** Set of inliers;
1: Initialize all the parameters. 2: Compute the posterior probabilities in Expectation step using Eq. $\boldsymbol{p_n} = \frac{\boldsymbol{\delta e}^{-\frac{\|C(a_n,b_n)\|^2}{2\sigma^2}}}{\boldsymbol{\delta e}^{-\frac{\|C(a_n,b_n)\|^2}{2\sigma^2}} + \frac{2\pi\sigma^2(1-\delta)}{k}}$ 5 3: Maximize the correspondence in maximization step until EM converges. 4: Using a threshold determine the set of inliers and outliers

Experimental Results

The proposed algorithm is experimented for registration of a set of color images with a deliberate movement incurred. This is explained in this section and all the results are shown in from Fig 1 to Fig 18. An image is first chosen and then deliberately translated and rotated to incur movement. This movement and rotation incurred image is taken as the second image input for registration. For example, the image in Fig 1 is considered to be the first set of point correspondence. The same image is moved and rotated anti-clockwise 30 degrees. These set of pixels are then subjected to the expectation problem to find the maximum correspondence between them. The final registered results through coherent pixel correspondence is given in Fig 3. The results show that our algorithm could also easily find the correspondence between the hand signature as well shown in the image.

The registration metrics used for all the tests we have done are tabulated in Table 1. The mutual information calculated for the set of images is given in this table. We have tested for various values of parameters explained before. The variations in the parameters used in the tests are tabulated in Table 3. The starting value of gamma for all the tests have been practically chosen as 0.9 and after the iterations we have obtained the gamma value for optimal results of registration. The energy change rate for the corresponding gamma values are also shown in the same table. Interestingly the energy change rate is minimum for image pairs with minimal correspondence points and is found to be more in case of image pairs with more set of feature points or matching corresponding pixels. In Table 4 we have tabulated different computational time taken for each image pairs for corresponding sigma values against the true value of correspondence in each case.

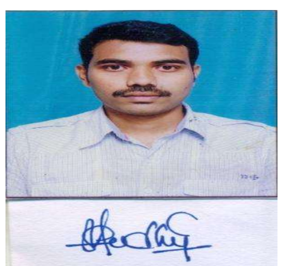

Fig:1 Test1 Input-1

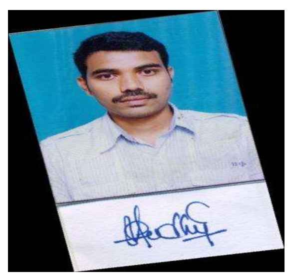

Fig:2 Test1 Input-2

Fig:3 Test 1 Registered Image

Fig:3 Test 1 Registered Image

A similar set of example is shown in Fig 4 to Fig 6, however there is 45 degrees anti-clockwise rotation. Also there are more feature points to be matched than the previous example and has been accurately done for coherent correspondence by our algorithm and visualized using blue lines in Fig 6. Fig 7 to Fig 9 shows another example where the correspondence points are less than the previous example and evidently takes lesser time in computation.

Fig:4 Test2 Input-1

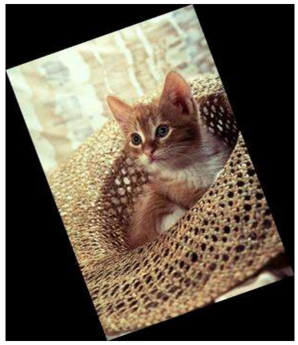

Fig:5 Test2 Input-2

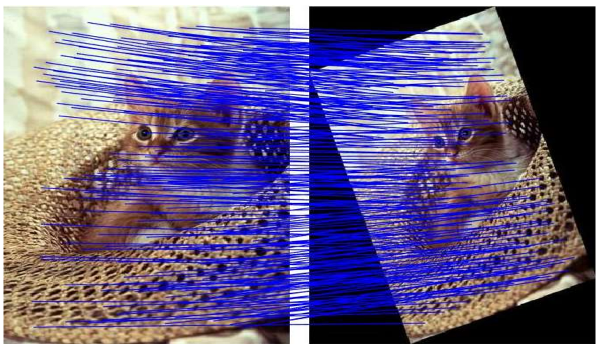

Fig:6 Test 2 Registered Image

Fig:7 Test3 Input-1

Fig:8 Test3 Input-2

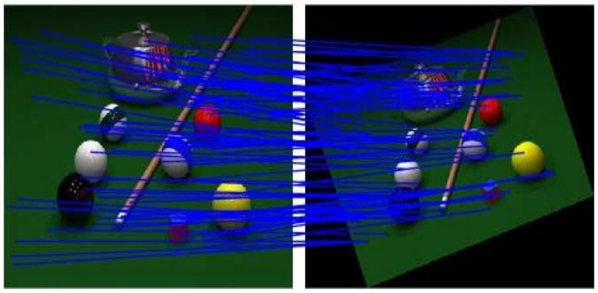

Fig:9 Test 3 Registered Image

 Fig:10 Test 4 Input-1	 Fig:11 Test 4 Input-2

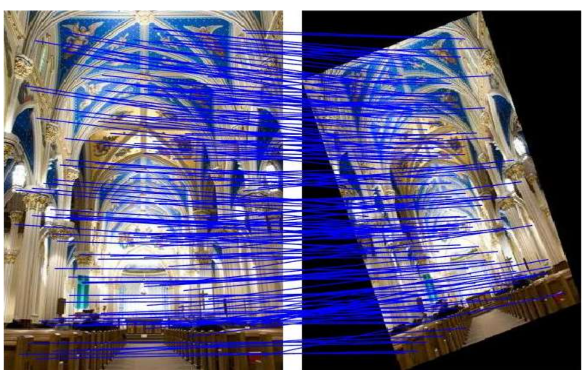

Fig:12 Test 4 Registered Image

Fig:13 Test 5 Input-1

Fig:14 Test 5 Input-2

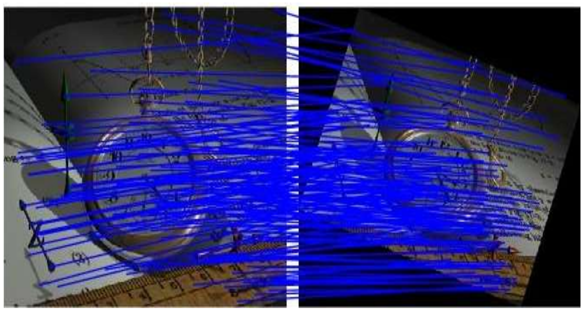

Fig:15 Test 5 Registered Image

The next three tests contain more mutual information and hence more pixel set for correspondence. This is evident in the mutual information tabulation and energy change rate. The corresponding time elapsed is also more for these last three tests. The final test image has taken more time in computation despite having lower correspondence. This is because there is more mutual information between the image pairs. The algorithm has worked very well in terms of final image registration in all the test cases we have put in. the results thus strongly conforms with the proposed method.

Fig:16 Test 6 Input-1

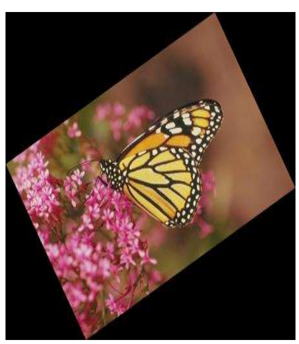

Fig:17 Test 6 Input-2

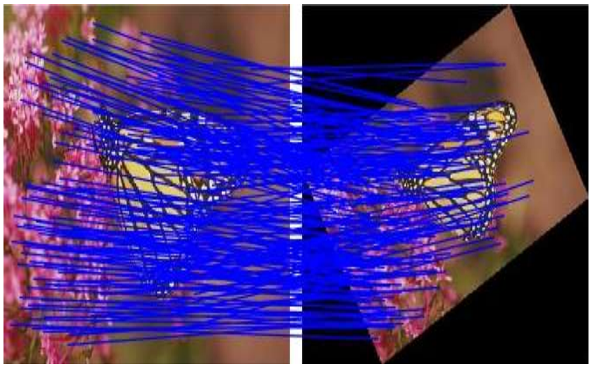

Fig:18 Test 6 Registered Image

Table 2: Registration Metrics

Test Sample	Information Ratio (IR) count	Lower bound on Information Ratio (LIR) count	Mutual Information Ratio (MIR) count	Lower bound on Mutual Information Ratio (LMIR) count
Test-1	143.38	80.62	53.79	148.70
Test-2	167.62	90.94	55.53	175.32
Test-3	104.68	61.35	21.32	115.33
Test-4	167.18	90.71	42.88	176.83
Test-5	143.85	82.1	20.79	158.10
Test-6	146.37	82.85	29.12	160.61

Table 3: Variation in parameters

	Gamma		The Energy Change Rate	
Test Sample	At Starting	At Saturation	At Starting	At Saturation
Test-1	0.9	0.4070	1.0033	0.000009
Test-2	0.9	0.3067	1.0018	0.000001
Test-3	0.9	0.3344	1.0113	0.000006
Test-4	0.9	0.3955	1.0030	0.544872
Test-5	0.9	0.4055	1.0027	0.000003
Test-6	0.9	0.243	1.0040	0.031723

Table 4: Variation in parameters and Elapsed time for algorithm

	Sigma			
Test Sample	At Starting	At Saturation	Elapsed time in sec	Correct correspondence (%)
Test-1	0.334	0.000011	0.665	40.703
Test-2	0.363	0.000014	4.758	30.679
Test-3	0.525	0.000258	0.646	33.447
Test-4	0.381	0.000014	16.876	39.556
Test-5	0.322	0.000011	1.383	40.559
Test-6	0.436	0.000026	44.54	24.375

Conclusion

We have proposed a method for image registration through coherent pixel correspondence in this paper using probability and maximum likelihood. The experimental results have shown excellent results under very less computation time for maximum correspondence. The results also have been achieved for image pairs that has experienced translation and rotational movements for various parameters. The method has been tested on color images and could thus be generalized for any application.

References

[1] L. Zhou, Z. Wang, Y. Fu, J. Chen, R. Xiang, R. Zhong and S. Wang, "Depth image in-loop filter via graph cut," in International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 2016.

[2] T. Zuk and M. Atkins, "A comparison of manual and automatic methods for registering scans of the head," IEEE Transactions on Medical Imaging, vol. 15, no. 5, 1996 .

[3] A. Roche, G. Malandain and N. Ayache, "Unifying maximum likelihood approaches in medical image registration," International Journal of Imaging Systems and Technology, vol. 11, no. 1, pp. 71-80, 2000.

[4] W. Li and H. Leung, "A maximum likelihood approach for image registration using control point and intensity," IEEE Transactions on Image Processing , vol. 13, no. 8, pp. 1115 - 1127, 2004 .

[5] S. Chen, Q. Guo, H. Leung and É. Bossé, "A Maximum Likelihood Approach to Joint Image Registration and Fusion," IEEE Transactions on Image Processing, vol. 20, no. 5, pp. 1363 - 1372, 2011 .

[6] E. Castillo, R. Castillo, B. White, J. Rojo and T. Guerrero, "Least median of squares filtering of locally optimal point matches for compressible flow image registration," Physics in Medicine & Biology, vol. 57, no. 15, 2012.

[7] D. Mintz, P. Meer and A. Rosenfeld, "Analysis of the least median of squares estimator for computer vision applications," in CVPR, 1992.

[8] A. Bab-Hadiashar and D. Suter, "Robust optic flow estimation using least median of squares," in Proceedings of 3rd IEEE International Conference on Image Processing, Switzerland, 1996.

[9] M. I. Patel and V. K. Thakar, "Speed improvement in image registration using maximum likelihood based mutual information," in International Conference on Advanced Computing and Communication Systems, 2015.

[10] S. Chen, Q. Guo, H. Leung and É. Bossé, "A Maximum Likelihood Approach to Joint Image Registration and Fusion," IEEE Transactions on Image Processing, vol. 20, no. 5, pp. 1363 - 1372, 2011.

[11] M.-E. Lee, W.-H. Cho, S.-W. Kim, S.-H. Kim and X. Zhao, "Intensity-Based Registration of Medical Images Using Penalized Maximum Likelihood," in IEEE International Workshop on Semantic Computing and Systems, 2008.

[12] H.-m. Chen and P. Varshney, "Mutual information-based CT-MR brain image registration using generalized partial volume joint histogram estimation," IEEE Transactions on Medical Imaging, vol. 22, no. 9, pp. 1111 - 1119, 2003 .

[13] X. Lu, S. Zhang, H. Su and Y. Chen, "Mutual information-based multimodal image registration using a novel joint histogram estimation," Computerized Medical Imaging and Graphics, vol. 32, no. 3, pp. 202-209, 2008.

[14] G. Ramos-Llordén, A. J. d. Dekker, G. V. Steenkiste, B. Jeurissen, F. Vanhevel, J. V. Audekerke, M. Verhoye and J. Sijbers, "A Unified Maximum Likelihood Framework for Simultaneous Motion and T1 Estimation in Quantitative MR T1 Mapping," IEEE Transactions on Medical Imaging, vol. 36, no. 2, 2017.

[15] X. L. &. Z. Hu, "Rejecting Mismatches by Correspondence Function," International Journal of Computer Vision , p. 1–17, 2010.

[16] J. Ma, J. Zhao, J. Tian, A. L. Yuille and Z. Tu, "Robust Point Matching via Vector Field Consensus," IEEE Transactions on Image Processing , vol. 23, no. 4, pp. 1706 - 1721, 2014.

[17] W. Tao and K. Sun, "Robust Point Sets Matching by Fusing Feature and Spatial Information Using Nonuniform Gaussian Mixture Models," IEEE Transactions on Image Processing, vol. 24, no. 11, 2015 .

[18] P. S. Reel, L. S. Dooley, K. C. P. Wong and A. Börner, "Enhanced retinal image registration accuracy using expectation maximisation and variable bin-sized mutual information," in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2014.

[19] E. Arce-Santana, D. U. Campos-Delgado, A. Mejia-Rodriguez and I. Reducindo, "An innovative multimodal/multispectral image registration method for medical images based on the Expectation-Maximization algorithm," in International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2015.

[20] I. Abdulhalim, "Competence between spatial and temporal coherence in full field optical coherence tomography and interference microscopy," Journal of Optics A: Pure and Applied Optics, vol. 8, no. 11, 2006.

[21] M. Zhang, C. Yang, L. Wei, Z. Zhang, R. Chen and H. Zhou, "Non-rigid Point Set Registration via Coherent Spatial Mapping and Local Structures Preserving," in International Symposium on Parallel and Distributed Computing (ISPDC), 2016.

[22] S. Cho, T. G. Brown and M. A. Alonso, "Simple methods for measuring spatial coherence and their relation to the Wigner function," in Euro-American Workshop on Information Optics, 2012.

[23] N. Ryan, C. Heneghan and P. Chazal, "Registration of digital retinal images using landmark correspondence by expectation maximization," Image and Vision Computing, vol. 22, no. 11, pp. 883-898, 2004.
[24] M. Zhang, N. Singh and P. T. Fletcher, "Bayesian Estimation of Regularization and Atlas Building in Diffeomorphic Image Registration," in International Conference on Information Processing in Medical Imaging, 2013.
[25] K. Zhang, L. Zhang, K.-M. Lam and D. Zhang, "A Level Set Approach to Image Segmentation With Intensity Inhomogeneity," IEEE Transactions on Cybernetics, vol. 46, no. 2, pp. 546 - 557, Feb. 2016.