 DOI: 10.3217/978-3-85125-663-5-44

The Coarse-to-Fine Contour-based Multimodal Image Registration

Maliheh Assadpour Tehrani[1] and Robert Sablatnig[1]

***Abstract*— Image registration brings two images into alignment despite any initial misalignment. Several approaches to image registration make extensive use of local image information extracted in interest points, known as local image descriptors. State-of-the-art methods perform a statistical analysis of the gradient information around the interest points. However, one of the challenges in image registration by using these local image descriptors arises for multimodal images taken from different imaging devices and/or modalities. In many applications such as medical image registration, the relation between the gray values of multimodal images is complex and a functional dependency is generally missing. This paper focuses on registering Mass spectrometry images to microscopic images based on contour features. To achieve more accurate multimodal image registration performance, we proposed a coarse-to-fine image registration framework. The pre-registration process is performed by using contour-based corners and curvature similarity between corners. Image blocking and DEPAC descriptors are used in the fine registration process. A local adaptive matching is performed for the final registration step.**

I. INTRODUCTION

The term image registration describe the procedure of aligning two images that have been acquired by different imaging conditions. The overall aim of image registration approaches is to determine a transformation between the source and target images. There have been multiple approaches during the last decades in order to find such a transformatation[19][17]. According to[19] the process of image registration involves the following steps: Feature detection and matching, transform model estimation and image transformation. Two or more images are called multimodal if different sensors or different imaging devices obtain them. They can provide different information of the same scene but the multimodal registration task is difficult since the images obtained from different modalities can have extreme intensity mapping dissimilarity.

A. Mass Spectrometry Imaging

Mass Spectrometry Imaging (MSI) technologies are powerful tools to investigate the molecular information from biological tissue samples and visualize their complex spatial distributions [4]. MSI allows untargeted analysis of hundreds of molecular species directly from a tissue sample, providing direct spatial correlation between their abundances and histological features. Beside mass-to-charge (m/z) values and their respective intensities, MSI is also recording the spatial position where signals are desorbed and ionized from surface [14].

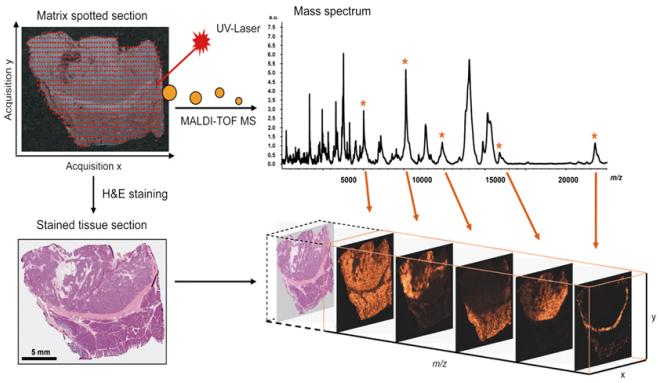


Fig. 1. An example of MSI and Histology Imaging.

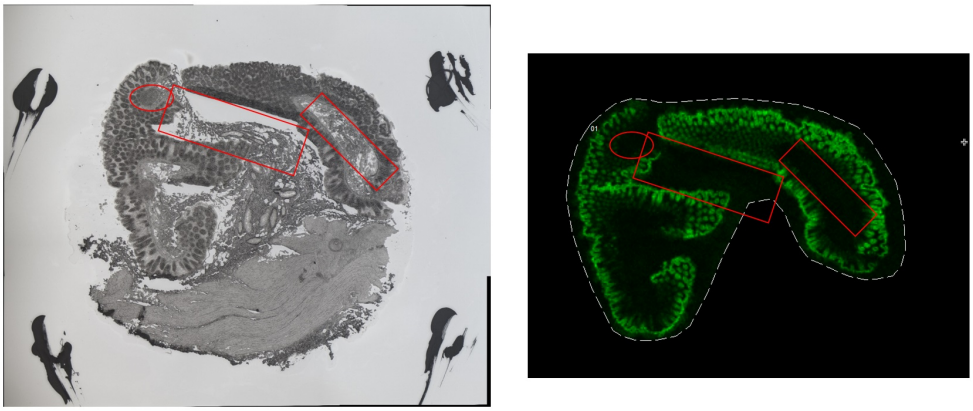
A schematic representation of a MSI is shown in Fig. 1. The sample (tissue section) is covered with matrix in order to allow desorption and ionization of analytes. The resulting ions are transferred to the mass spectrometer, and a mass spectrum is acquired. Then, the next position by a defined distance is analysing. MSI of a selected analyte peak are generated after the measurement by extracting the signal intensity within a certain m/z window. The intensities are plotted as gray scale values for each pixel in a grid representing the corresponding positions on the sample.

B. Challenging multimodal images

MSI can generate biomolecular profiles that describe the spatial distribution of specific biomolecules including metabolites, lipids, peptids and proteins. However, it can not provide histoanatomical and molecular depth information [4]. Therefore, the combination of information from different modalities has proven to be a powerful approach for obtaining molecular signatures from specific cells/tissues of interest. For example, MSI is combined with imaging modalities that have high spatial resolution (Optical Microscopy Imaging-OMI) or tissue structural information (Histology Images). An example of multimodal images is given in Fig.2, an OMI and a MSI. Obviously, there are very large content differences between these two images. It can be seen that each of the regions in the microscopy images contains a large amount of information. In contrast, there is much less information in each of corresponding regions in the MSI. Moreover, Medical microscopic images contain many objects which are visually very similar. These issues are affecting the accuracy of registration results.

*This work was not supported by any organization

[1]Maliheh Assadpour Tehrani and Robert Sablatnig are with Institute of Visual Computing and Human-Centered Technology, Faculty of Informatics, TU Wien, Austria {tehrani,sab}@cvl.tuwien.ac.at



(a)Optical Microscopy (b)Mass Spectrometry

Fig. 2. An example OMI(a) and MSI(b). The selected regions by red lines show very large content difference between images.

C. Image Registration

The term image registration describes the procedure of aligning two images that have been acquired by different imaging conditions. The overall aim of image registration approaches is to determine a transformation between the source and target images. There have been multiple approaches during the last decades in order to find such a transformatation [17][19]. According to [19] the process of image registration involves the following steps: Feature detection and matching, transform model estimation and image transformation.

The extraction of effective features is a crucial step for the application of image registration. Over the last decades, numerous hand-engineered features such as SIFT[12] and Gabor filters[11] are used in the image registration procedures. However, they suffer from limitations for the multimodal image registration because the statistics on the local intensity distribution are insufficient to describe the complex relationship between modalities with different underlying imaging physics[3].

Two or more images are called multimodal if they are obtained by different sensors or different imaging devices. They can provide different information of the same scene but the multimodal registration task is difficult since the images obtained from different modalities can have extreme intensity mapping dissimilarity. Variation in intensities has two possible consequences as 'Gradient Reversal' and 'Region Reversal' which are described in the following part.

The gradient direction of corresponding parts in the multimodal images are changing by exactly 180°[15]. This is called 'Gradient Reversal' which is one of the main reasons that causes SIFT to fail with multimodal images. It also may draw on rotation normalization of regions fault. In a consequence of gradient reversal, the direction of the dominant orientation will reverse which may cause two similar regions remain totally out of phase. We call this property as 'Region Reversal'.

Among local features methods, multimodal variants of SIFT are particularly popular in multimodal image registration, including SIFT-GM (GM: Gradient Mirroring)[10], Symmetric SIFT [7], IS-SIFT (IS:Improved Symmetric) [9], GO-IS-SIFT [16], PIIFD (Partial Intensity Invariant Feature Descriptor) [6]. These variants of SIFT only consider 'Gradient Reversal' and 'Region Reversal' problem of multimodal images, however the real situation may be more complex, such as registering the two images shown in Fig. 2.

In [13], a multimodal image registration method based on a contour-based corner technique (COREG) which is independent of intensity and gradient changes is proposed to overcome these challenging images. COREG algorithm has shown satisfactory registration performance when registering images without large scale difference. Although, when the scale difference increase, the COREG performance is decreasing.

D. Contribution of the Paper

The problem that we try to overcome in this work is caused by large content and scale differences between MSI and OMI. First, the multimodal image registration technique based on a countour-based corner technique that is independent of intensity and gradient changes is investigated [13]. Then, we proposed a method by taking the advantage of COREG algorithm and adding Image blocking and local adaptive matching to the registration procedure for multimodal medical image registration.

The rest of this paper is organized as follow. In Section II, we present our proposed method. Coarse registration is described in II-A, followed by fine registration steps consist of image blocking in II-B, DEPAC descriptor extraction in II-C and local adaptive matching in II-D. The dataset and results are presented in Section III and this paper is concluded in Section IV.

II. THE PROPOSED TECHNIQUE

In order to register these challenging multimodal images, a fully automatic image registration approach is proposed in this paper. Our proposed method is designed based on

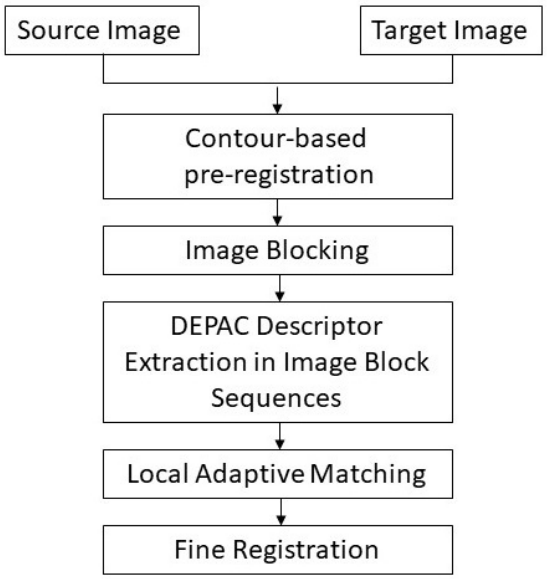


Fig. 3. The proposed algorithm flowchart.

registration process in COREG using the following processes as shown in Fig.3.

In this work, the image registration process is implemented in two main steps. First, the pre-registration process is initialized by detecting corners using Fast-CPDA and then the matching of corners are obtained according to the curvature similarities of all the corners. For each pair of corner triplets, the transformation is computed. By using this transformation, the target edge image is transformed to the reference edge image and the Number of Overlapped Pixels (NOP) is estimated. The transformation related to the maximum NOP is selected for coarse registration.

Second, the algorithm performs uniform image blocking of reference image and transformed target image obtained from pre-registration step. In this step, the DEPAC descriptors for all the corners are computed in each image blocks of reference and transformed target image. The second round matching process is carried out by local adaptive matching strategy.

A. Pre-registration

In this paper, the fast contour-based multi-scale corner detector based on the chord-to-point distance accumulation (Fast-CPDA) technique by [8] has proposed to detect the corners. This technique has proved its robustness over many other single- and multi-scale detectors [8][2][1].The CPDA detector [8] first extracts planar curves from the edge image detected by the Canny edge detector [5].

After corner detection, the curvature similarity of all detected corners in the source and the target image are determined as following.

We have two sets of corners in the reference and target images as

$$C_r = \left\{C_r^1, C_r^2, \ldots C_r^{N_r}\right\}, \tag{1}$$

and

$$C_t = \left\{C_t^1, C_t^2, \ldots C_t^{N_t}\right\}, \tag{2}$$

where N_r and N_t denote the number of corners in the reference and target images respectively. The curvatures of these corners are:

$$K_r = \left\{K_r^1, K_r^2, \ldots K_r^{N_r}\right\}, \tag{3}$$

and

$$K_t = \left\{K_t^1, K_t^2, \ldots K_t^{N_t}\right\}. \tag{4}$$

The curvature similarity of two corners is determined as

$$s^{i,j} = \frac{\left|K_r^i - K_t^j\right|}{K_r^i}, \tag{5}$$

where $1 \leq i \leq N_r$ and $1 \leq j \leq N_t$.

Then all the target corners are ranked based on their curvature similarities to the each source corner. With these matches, the corner triplets are generated. After matching corner triplet pairs, the transformation is computed for each matching pair. This transformation is used to transform the target edge image to the reference edge image and the Number of Overlapped Pixels (NOP) is estimated. The transformation obtained from the pair of corner triplets with the maximum NOP is selected as a coarse transformation.

B. Image Blocking

The large content and scale difference between MSI and OMI are affecting the accuracy of the registration. COREG algorithm has shown satisfactory registration performance when registering images without large scale difference. Although, when the scale difference increase, the COREG performance is decreasing. Moreover, the space of geometric transformations becomes larger which means costing more time to compare corner triplets. Therefore, we employed the image blocking method which dividing each source and transformed target image (warp image) into the image blocks. After pre-registration step is performed, it is assumed that the source image size is $M \times N$ and the warp image size is $P \times Q$. We are dividing them into the $m \times n$ blocks that means the size of source and warp image blocks are $(M/m) \times (N/n)$ and $(P/m) \times (Q/n)$ accordingly. All these blocks are labeled and the similar blocks in the two images will have the same number. Hence, this approach decomposes a matching problem of a whole image into numerous matching problems of image blocks. For each block, the DEPAC method is used to build a DEPAC descriptor for each corner. The DEPAC descriptor is presented in Section II-C.

C. DEPAC Descriptor Extraction

G.Lv. et.al [13] presented the novel corner descriptor which is based on the curvature of a corner in order to capture important edge information in the neighborhood of a corner. The proposed corner descriptor is called Distribution of Edge Pixels Along Contour (DEPAC). The corners and their contours are presented as C_r^i, C_t^j, $\Gamma\left(C_r^i\right)$ and $\Gamma\left(C_t^j\right)$ in

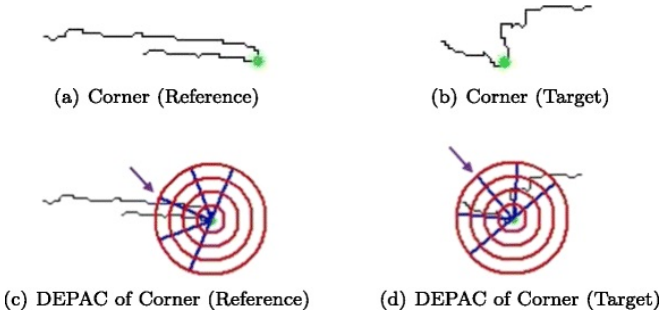


Fig. 4. DEPAC corner descriptor [13].

Fig.4 a and b. The DEPAC corner descriptor is built using C_r^i and $\Gamma\left(C_r^i\right)$. The main steps are as follows.

i. Each corner is used as a center to compute concentric circles are plotted, as shown in Fig. 4. Where R is denoted the radius of the internal circle. The radius of a concentric circle is incremented by R, from inside to outside.

ii. The average orientations of two tangants is considered as the main orientation of the corner, O_m. Arrows in Fig. 4 c,d, show the main orientation.

iii. At each side of main orientation, the orientation bins are defined. Therefore, the four quantized orientations are estimated as $O_1 = O_m - 90°$, $O_2 = O_m - 45°$,$O_3 = O_m$ and $O_4 = O_m + 45°$ in an anticlockwise direction. Finally, we have 16 sub-regions which are defined in the neighborhood of the corner and each sub-region is denoted as (c,o), where $1 \leq c \leq 4$ and $1 \leq o \leq 4$.

iv. The number of edge pixels is increased if an edge pixel, P_e, along the contour falls into this sub-region, i.e.

$$(c-1) \times R < d\left(P_e, C_r^i\right) \leq c \times R, \tag{6}$$

and

$$O_o \leq \overrightarrow{C_r^i P_e} < O_{o+1}, \tag{7}$$

where $d\left(P_e, C_r^i\right)$ is the Euclidean distance between P_e and C_r^i.

v. Final step is the normalizing the number of edge pixels in each sub-region, NEPc,o, into [0,1] by

$$NEP_{c,o} = \frac{NEP_{c,o}}{max\{NEP_{c,o}\}}. \tag{8}$$

D. Local Matching Scheme

After DEPAC descriptor extraction is performed in each image sub-block, we are assuming that $D_{A_i} = \{f_j\}_i$ and $D_{B_i} = \{g_j\}_i$ are a set of DEPAC descriptors where A_i and B_i are image sub-blocks with number i in the source and the target image respectively. Number of extracted DEPAC descriptors in image sub-blocks is shown by j. The Euclidean distance is calculated between each DEPAC descriptor in image sub-block A_i and each DEPAC descriptor in image sub-block B_i.

By ranking the DEPAC descriptor distances, the nearest and the next nearest distance are selected to calculate the ratio distance as following:

$$R_j = d_{j1}/d_{j2}. \tag{9}$$

Where R_j is the ratio distance. The nearest and next nearest distance are denoted by d_{j1} and d_{j2}, respectively.

For each image sub-block, the ratio distance calculation is performed. All R_j are ranked from small to big. The matching corners to the smallest R are selected and merged for fine-registration step. Similar to the final step in coarse-registration, the matching of corner triplets is carried out based on selected corners. These corner triplets are used to estimate transformation which correspond to the higher NOP. Fig. 5 shows the fine-registration process consist of image blocking, DEPAC descriptor extraction and local adaptive matching.

III. EXPERIMENTS

In this section we present comparative study of presented methods for our specific multimodal images to investigate how they perform on registering MSI to Optical microscopy image. Then, we will evaluate our proposed method.

A. Test Data

Our dataset consists of 10 reference microscopy images from tissues and for each reference image, 4 Mass spectrometry images are available, leading to a total of 40 image pairs. The ground truth for all of our test images are processed using FlexAnalysis (version 3.0, Bruker Daltonics). By this software, the MS images are manually registered to the optical images.

B. Evaluation Metric

In order to quantitatively comparing the performance of the presented registration methods, we used the Average Registration Error(ARE) [18]. After aligning the reference and target images with the estimated transformation, ARE is used to measure the overlap error as defined by

$$ARE = \frac{1}{H \times W} \sum_{x=1}^{W} \sum_{y=1}^{H} \|T_e(x,y) - T_g(x,y)\|, \tag{10}$$

where H and W are the height and width of the reference image, T_g is the ground-truth transformation and T_e is the estimated transformation. The smaller the ARE value is, the better the registration performance will be.

C. Performance Comparisons

Fig. 6 compares our presented method with COREG in terms of ARE when registering image pairs of MSI and OMI. From the experimental results, it can be summarized that our proposed algorithm shows clear improvement over COREG. The scale difference between two images in image pairs 24 to 40 (Image Pair ID 5-8) is approximately 1X:10X. ARE values obtained by COREG registration process for these image pairs are obviously bigger compared to pairs 1 to 24. On the other hand, ARE values achieved by our method remain relatively stable in comparison with COREG when increasing scale differences.

In this paper we have proposed coarse-to-fine registration method to overcome the problem of big scale difference

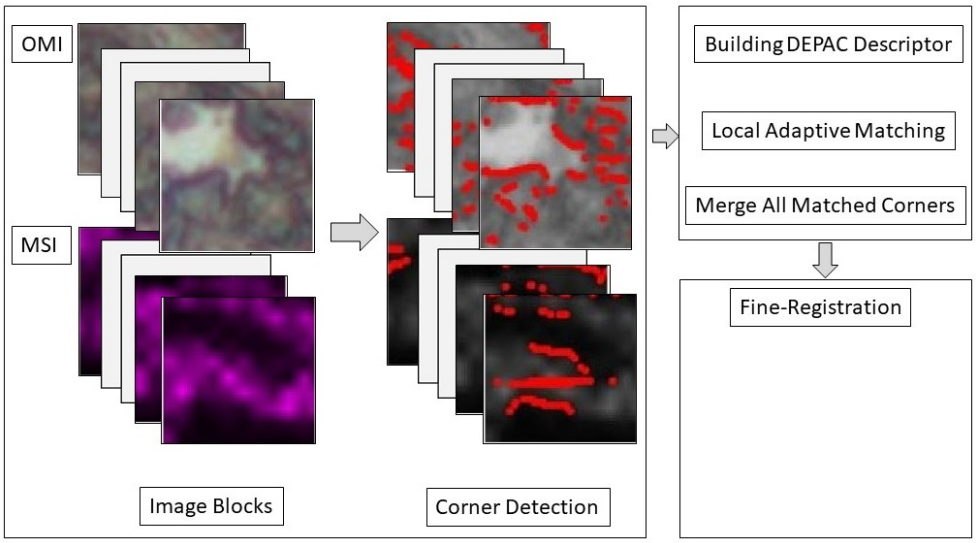


Fig. 5. Fine-registration Process.

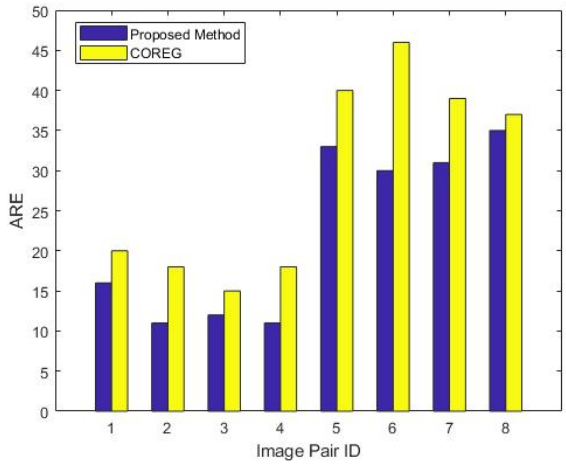


Fig. 6. ARE comparisons between our proposed method and COREG. In this figure image pairs 1-3 have scale difference of 1X:1X and image pairs 4-5 have scale difference of 1X:5X. Image pairs 6-8 have a large scale difference of 10X.

when registering MSI and OMI. The impact of scale difference in terms of ARE is presented in Table I. This proposed method has shown better results with average ARE of 80.45 in comparison of COREG with average ARE of 115.65.

TABLE I
IMPACT OF SCALE DIFFERENCE IN TERMS OF ARE.

Scale Difference	Proposed Method	COREG
$1X:1X$	13	43
$1X:5X$	27.5	38
$1X:10X$	72.67	97.33

IV. CONCLUSION

We have presented a new multimodal image registration method based on contours and DEPAC descriptor. In order to address the large content and large scale difference in MSI and OMI, we have proposed coarse-to-fine registration framework. Our coarse registration is contour-based procedure. The proposed approach utilizes various techniques including the contour matching algorithm, curvature similarity, DEPAC descriptors, image blocking and local adaptive matching as a robust matching mechanism. The algorithm can complete fully automatic registration without any manual intervention. At the same time, all the matching corners extracted in the matching process are distributed uniformly, which can consequently improve the accuracy of registration.

ACKNOWLEDGMENT

Data used in preparation of this paper were obtained from the Institute of Chemical Technologies and Analytics, Vienna University of Technology, Vienna, Austria.

REFERENCES

[1] M. Awrangjeb and G. Lu, "Robust image corner detection based on the chord-to-point distance accumulation technique," *IEEE Transactions on Multimedia*, vol. 10, no. 6, pp. 1059–1072, oct 2008. [Online]. Available: http://ieeexplore.ieee.org/document/4657455/

[2] M. Awrangjeb, G. Lu, C. S. Fraser, and M. Ravanbakhsh, "A fast corner detector based on the chord-to-point distance accumulation technique," in *DICTA 2009 - Digital Image Computing: Techniques and Applications*. IEEE, 2009, pp. 519–525. [Online]. Available: http://ieeexplore.ieee.org/document/5384897/

[3] W. Bingjian, L. Quan, L. Yapeng, L. Fan, B. Liping, L. Gang, and L. Rui, "Image registration method for multimodal images," *Appl. Opt.*, vol. 50, no. 13, pp. 1861–1867, may 2011. [Online]. Available: http://ao.osa.org/abstract.cfm?URI=ao-50-13-1861

[4] A. R. Buchberger, K. DeLaney, J. Johnson, and L. Li, "Mass Spectrometry Imaging: A Review of Emerging Advancements and Future Insights," *Analytical Chemistry*, vol. 90, no. 1, pp. 240–265, jan 2018. [Online]. Available: http://pubs.acs.org/doi/10.1021/acs.analchem.7b04733

[5] J. Canny, "A Computational Approach to Edge Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679–698, nov 1986. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4767851

[6] J. Chen, J. Tian, N. Lee, J. Zheng, R. T. Smith, and A. F. Laine, "A partial intensity invariant feature descriptor for multimodal retinal image registration," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 7, pp. 1707–1718, July 2010.

[7] J. Chen and J. Tian, "Real-time multi-modal rigid registration based on a novel symmetric-SIFT descriptor," *Progress in Natural Science*, vol. 19, no. 5, pp. 643–651, may 2009. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1002007109000240

[8] J. H. Han and T. Poston, "Chord-to-point distance accumulation and planar curvature: A new approach to discrete curvature," *Pattern Recognition Letters*, no. 10, pp. 1133–1144, aug.

[9] M. T. Hossain, G. Lv, S. W. Teng, G. Lu, and M. Lackmann, "Improved symmetric-sift for multi-modal image registration," in *2011 International Conference on Digital Image Computing: Techniques and Applications*, Dec 2011, pp. 197–202.

[10] A. Kelman, M. Sofka, and C. V. Stewart, "Keypoint descriptors for matching across multiple image modalities and non-linear intensity variations," in *2007 IEEE Conference on Computer Vision and Pattern Recognition*, June 2007, pp. 1–7.

[11] R. Kolkila and T. P, "Image Registration Based on Fast Fourier Transform Using Gabor Filter," *International Journal of Computer Science and Electronics Engineering (IJCSEE).*, vol. 2, pp. 31–37., 2014.

[12] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, nov 2004. [Online]. Available: http://link.springer.com/10.1023/B:VISI.0000029664.99615.94

[13] G. Lv, S. W. Teng, and G. Lu, "COREG: a corner based registration technique for multimodal images," *Multimedia Tools and Applications*, vol. 77, no. 10, pp. 12 607–12 634, 2018.

[14] L. A. McDonnell and R. M. Heeren, "Imaging mass spectrometry," *Mass Spectrometry Reviews*, vol. 26, no. 4, pp. 606–643, jul 2007. [Online]. Available: http://doi.wiley.com/10.1002/mas.20124

[15] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Image Registration by Maximization of Combined Mutual Information and Gradient Information," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2000*, S. L. Delp, A. M. DiGoia, and B. Jaramaz, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2000, pp. 452–461.

[16] G. L. Shyh Wei Teng, Md. Tanvir Hossain, "Multimodal image registration technique based on improved local feature descriptors," *Journal of Electronic Imaging*, vol. 24, no. 1, pp. 1 – 17 – 17, 2015. [Online]. Available: https://doi.org/10.1117/1.JEI.24.1.013013

[17] A. Sotiras, C. Davatzikos, and N. Paragios, "Deformable medical image registration: A survey," *IEEE Transactions on Medical Imaging*, vol. 32, no. 7, pp. 1153–1190, jul 2013. [Online]. Available: http://ieeexplore.ieee.org/document/6522524/

[18] M. Xia and B. Liu, "Image registration by "super-curves"," *IEEE Transactions on Image Processing*, vol. 13, no. 5, pp. 720–732, may 2004. [Online]. Available: http://ieeexplore.ieee.org/document/1288197/

[19] B. Zitová and J. Flusser, "Image registration methods: a survey," *Image and Vision Computing*, vol. 21, no. 11, pp. 977–1000, oct 2003. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0262885603001379