

ROBUST AND EFFICIENT 3D REGISTRATION VIA DEPTH MAP-BASED FEATURE POINT MATCHING IN IMAGE-GUIDED NEUROSURGERY

Jie Yang¹, Shaoting Zhang², Xiahai Zhuang³, Long Jiang¹, Lixu Gu¹

¹School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China

²Department of Computer Science, University of North Carolina at Charlotte, NC, USA

³Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai, China

ABSTRACT

In image-guided neurosurgery, preoperatively acquired diagnostic images (e.g., brain MRI) should be accurately registered to the physical space that is specific to the patient's intraoperative neuroanatomy. A popular framework of registration requires manual defining corresponding positions of fiducial markers on the patient head and the preoperative brain MRI. The procedure is time-consuming and subjective to intra- and inter-observer variations. Therefore, markerless-based registration becomes increasingly popular. In this paper, we propose an automated markerless registration framework. Instead of using physical markers, we automatically detect feature points in face depth maps. The preoperative facial depth map is extracted from MRI, while the intraoperative map is reconstructed with structured light projection, using phase shifting interferometry. Then, we automatically detect and match the feature points on these two depth maps, using a robust method based on the extended SIFT algorithm. The transform matrix between the two coordinate systems can be computed accordingly. Our experiments on real data result in reasonable registration efficiency, while synthetic testing reveals promising accuracy. Average online processing time is no more than 1s totally in a MATLAB implementation.

Index Terms— image-guided neurosurgery, registration, depth map, SIFT, robust, efficient

1. INTRODUCTION

Image-guided neurosurgery becomes increasingly popular with the emphasis on minimal-invasive surgery procedures, and it is getting more precisely controlled and well-organized with the advance of volumetric CT and MRI.

In image-guided neurosurgery, preoperative planning (e.g. biopsy needle trajectories based on preoperative volumetric brain MRI) and intraoperative data should be registered. It can be interpreted as an alignment of preoperatively obtained diagnostic image series (e.g., volumetric brain MRI image) to a coordinate system that is specific to the intraoperative patient's neuroanatomy [1].

A popular framework for the registration involves the identification of common landmarks which are localized on the patient and can be identified from the preoperative images.

For example, homogenous landmarks (e.g. the nasal tip, and the inner or outer canthi of the eyes) are selected manually by a tracked pointer on patient face and preoperative diagnostic image series. An alternative way uses probe to sample points on the surface of the patient, and then determines the best match of this point-cloud to an extracted surface from the 3-D patient image [2]. Unfortunately, such manual selection is time-consuming and would be subjective to intra- and inter-observer variations. Therefore, markerless-based registration becomes increasingly popular. However, the achieved accuracy is often reported to be lower. In 2007, R. Konietzschke et al. [3] proposed a markerless and contact free registration method, using DLR handheld 3D-Modeller (3DMo) for acquiring patient intraoperative surface. Their registration accuracy was better than 3mm, yet it is expensive to build.

In this paper, we propose a very efficient and accurate markerless framework for registration. Instead of using physical markers, we use face depth maps to represent face surfaces's 3D location. The preoperative face depth map is extracted from brain MRI, while the intraoperative map is reconstructed with structured light projection, using phase shifting interferometry [4]. Since depth map is in 2D format yet contains 3D information (i.e., depth), it can remarkably speed up the traditional 3D operations without sacrificing the accuracy. To register two surfaces, a set of feature points are automatically detected on each depth map, based on the extended SIFT [5]. These two sets of points are matched based on their locations and attached feature vectors. Once these points are matched, the global transform matrix between the two coordinate systems can be computed straightforwardly using procrustes analysis [6].

Our proposed framework is automatic, robust and fast enough to satisfy the real-time demand owing to 1) the homogenous points detection based on extended SIFT method and 2) dimension reduction from 3D data to 2D depth map. Our experiments on real data result in reasonable registration efficiency, while synthetic testing reveals promising accuracy. Average self-time for online processes, which are extended SIFT point detection, point matching and procrustes analysis in our framework, is totally no more than 1s in a MATLAB implementation.

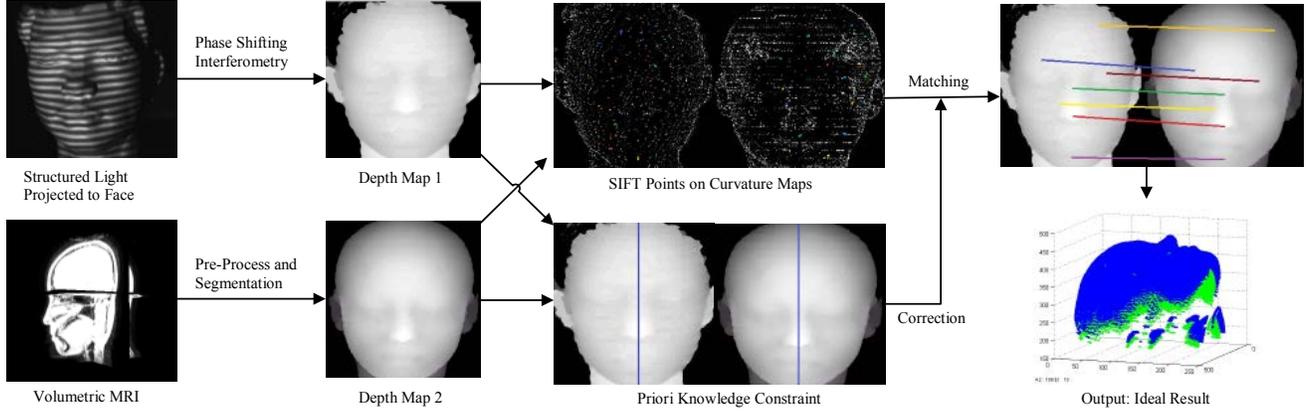


Fig. 1. Framework overview diagram. Depth map 1 is reconstructed with structured light projection, using phase shifting interferometry, and depth map 2 is obtained by volumetric brain MRI segmentation. Coordinate of the whole head can be traced accordingly. Curvature maps are generated from two depth maps, on which SIFT points are detected. The point matching is constrained by priori knowledge. After matching these points, alignment is accomplished using procrustes analysis.

2. METHODOLOGY

Framework Overview: Our proposed framework is based on point matching between preoperative and intraoperative depth maps. The intraoperative depth map is reconstructed with structured light projection, using phase shifting interferometry [4], which computes 3D information of surface by detecting the distortion when structured light are projected to facial surface. The preoperative depth map is segmented from the volumetric MRI. Points in the MRI are clustered with Quick shift [7] for better edge detection, and histogram-based threshold method is employed for computing threshold in each slice. For consistency, we set both face surfaces to be frontal, which can be achieved as follows: the intraoperative projector is placed straightly towards patient's frontal face, and the preoperative brain MRI is aligned to a standardized coordinate system, where the head is upright so as to get frontal face's depth map [8]. The value $D(x, y)$ on depth image maps to the distance between point (x, y) on frontal face and the back plane.

SIFT method is widely used for extracting features from images, these features were proved to be effective in medical image alignment by Matthew Toews et al. [9] in 2012, yet their method was applied on 3D data and required pre-aligned image learning, which is time and space consuming. Our method, on the other hand, is based on 2D maps for better computational efficiency. Indeed, information on depth map is limited. The lack of pattern on depth map prevents a good performance of SIFT method. Therefore, we calculate Gaussian curvature from depth map, reshaping it into a curvature map, where the value of each pixel $K(x, y)$ maps to its Gaussian curvature on the surface. The variety of features and patterns on curvature map meets the requirement of SIFT method, using which two sets of interest points are detected, with specific descriptor attached. These points on two maps are matched as per their locations and descriptors. To improve the robustness, Random Sample Consensus

(RANSAC) [10] is employed to prune outliers during point matching. We also use priori knowledge to constrain the pair matching, which is detected on depth map and can further exclude outliers generated from symmetric characteristic of faces. With the matching results, global transform aligning two coordinate systems is computed straightforwardly. In the following, we introduce several major modules in our framework.

SIFT Point Detection on Curvature Map: SIFT method has been widely used to extract distinctive and invariant features from images that can be used to perform reliable matching between different views of an object. It is able to provide robust matching across a substantial range of affine distortion, changes in 3D viewpoint, presence of image noise, and variations of illuminations, which are common dissimilarities our two depth maps may have. However, simply applying SIFT method on two depth maps does not work well due to limited information, i.e., depth map of face tends to be oversimplified without enough patterns for precise location. Therefore, we compute the curvature from depth map, reshaping it into a curvature map. Curvature is specific to anatomy, tolerant to a range of rotating and meets the demand of pattern variety. The SIFT method is then applied on curvature maps.

In SIFT method, potential interest points are identified by scanning the image over locations and scales. This is achieved efficiently by constructing a Gaussian pyramid and searching for local peaks (termed keypoints) in a series of difference-of-Gaussian (DoG) images.

$$local \ argmax_{x,y,\sigma} |f(x, y, k\sigma) - f(x, y, \sigma)| \quad (1)$$

Where $f(x, y, \sigma)$ is the convolution of the curvature map $K(x, y)$ with a Gaussian kernel of variance σ^2 , k is a multiplicative scale sampling rate, and the expression ' $local \ argmax \ {F(X)}$ ' denotes a set of values of the

argument X that locally maximize $F(X)$. A descriptor vector for each keypoint is built.

Matching with Extended Constraints: For point $P_i \in S_1$, its nearest neighbor P_j in S_2 can be defined as the keypoint with minimum Euclidean distance for the invariant descriptor vector. To discard points that do not have any good match to the database, we should compare the distance of the closest neighbor to that of the second-closest neighbor. In our case, we reject all matches in which the distance ratio is greater than 0.8, i.e., if P_j reliably matches P_i , then $|D_i - D_j| / |D_i - D_k| \leq 0.8$ ($P_k \in S_2, k \neq j$), where D_i denotes the descriptor of point P_i .

To improve the robustness, Random Sample Consensus (RANSAC) [10] is employed to prune outliers, however, there are still outliers generated from symmetric characteristic of face, which cannot be excluded by previous process. We employ priori knowledge for correction, which can be achieved using integral projection [11] and sparse shape composition [12]. For example, facial medial axis can be located by detecting the highest vertical line on depth map, points from different side of the line should not be paired.

Alignment of 3D Data Using Procrustes Analysis: Reshaping the depth maps into 3D surfaces, where the keypoint sets are now $S1\{p_1, p_2 \dots p_n\}$ and $S2\{q_1, q_2 \dots q_n\}$. With one-to-one correspondence matching, the transformation (translation matrix T and rotation matrix R) between the $S1$ and $S2$ is computed easily using procrustes analysis:

$$\operatorname{argmin}_{T,R} \sum_{i=1}^n \|(Rp_i + t) - q_i\|^2 \quad (2)$$

T and R can be employed for rigid alignment between the two surfaces.

3. EXPERIMENTAL SETUP AND RESULT

Experimental Setting: Our proposed framework is validated on 10 volunteers, generating 10 groups of real preoperative volumetric brain MRI series, which were obtained from GE Hdxt 3.0T, using fast SPGR pulse sequence in gradient echo family, with voxel size $0.43 \times 0.43 \times 1\text{mm}$. Their intraoperative surface depth map reconstruction system consists of a low-end CMOS Color camera and a SONY XPL-CX80 LCD projector. Another 20 groups of synthetic data are also used for testing, within which the preoperative MRI series are scanned using sequences above and the intraoperative data are obtained by transferring preoperative data to another coordinate system and changing scales or adding noises. All experiments were performed on a 2.5 GHz PC with 2 cores and 4G RAM, in MATLAB implementation.

Evaluation of the Accuracy and Robustness with Synthetic Data: In our framework, the preoperative depth map is of good quality, but the reconstructed intraoperative depth map would have scale change from the preoperative

map and is interfered by stripe noise. To test the robustness of our algorithm in above occasions, we first introduce tests on synthetic data.

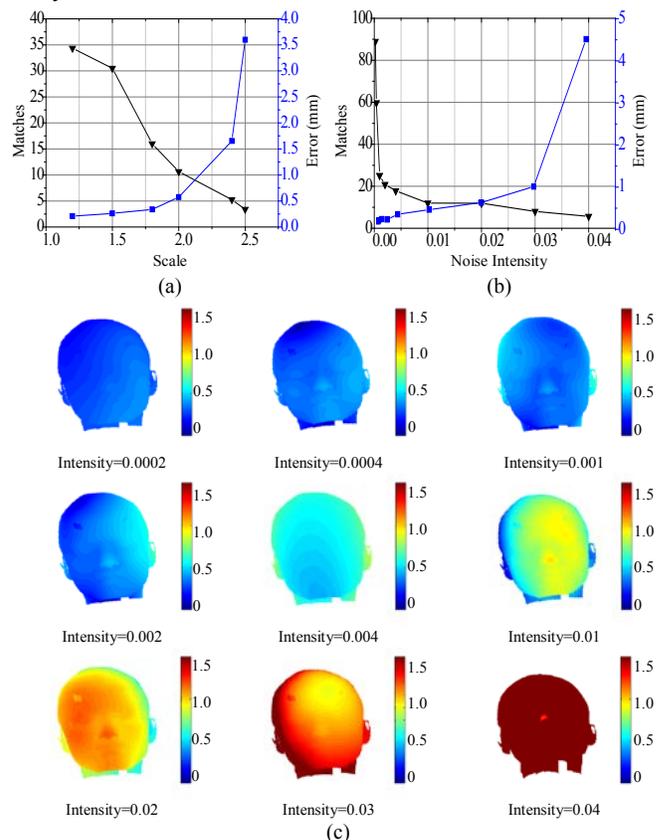


Fig.2. (a) Average number of matched SIFT points and registration error (average point to point distance among the whole surface compared with manually selecting landmarks) when image scale changes in intraoperative map; (b) Average number of matched SIFT points and registration error when stripe noise added while scale ratio is 1.0. The value of noise intensity is defined as $\frac{\max(\text{Noise})}{\max(\text{Depth})}$; (c) Visual comparisons between manually selecting points and using our auto algorithm when noise added. The color maps to registration errors.

In Fig.2. (a), when the scale change is no larger than 2, averagely more than 10 groups of point pairs would be found, ensuring high registration accuracy with the error no more than 0.6 mm. The robustness is good enough because the scale change in our framework is within that range.

In Fig.2. (b), when noise intensity is below 0.03, the registration error is slight (no more than 1mm) and grows slowly, whereas the error would suddenly climb up to 4mm when noise intensity reaches 0.04. Therefore, the extended SIFT method would have good performance when noise intensity is under certain threshold (the threshold is 0.04 in our tests). Fig.2. (c) shows the visualization of such errors.

Evaluation of the Accuracy and Robustness with Real Data: Among the 10 groups of real data, the average registration error is 3.613 mm, and the variance is 0.126.

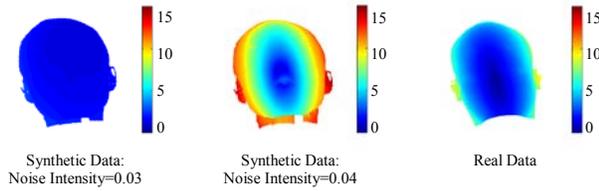


Fig. 3. Visualization of registration error when testing 1) synthetic data with noise intensity of 0.03, 2) synthetic data with noise intensity of 0.04 and 3) real data. The color maps to registration errors.

Fig.3 visualizes the numerical relationship of above three tests. The result of real data is numerically an intermediate state between noise intensity of 0.03 and 0.04 in synthetic data tests. These results indicate that, the noise intensity in real data is above yet close to the intensity threshold. Since patient's head is fixed during neurosurgery, factors such as facial movement can be excluded. Thus the dominant factor affecting accuracy in real data is noise intensity. Therefore, even a slight reduction of noise in real data might lead to changes that are far more positive in accuracy. Had the current reconstruction system improved, the registration would perform far better as well.

Evaluation of Computational Efficiency: The running time of major online processes, i.e., SIFT point detection, point matching and procrustes analysis, are averagely 0.464639s, 0.290894s and 0.079896s, summing up to no more than 1s, being suitable for real-time applications.

4. CONCLUSION AND DISCUSSION

In this paper, we proposed an efficient markerless framework of registration for image-guided neurosurgery. This is achieved by extended SIFT point matching in preoperative and intraoperative depth maps. The preoperative depth map is segmented from MRI and the intraoperative depth map is reconstructed with structured light projection. Compared with previous markerless frameworks, such as what R. Konietschke et al. [3] proposed in 2007, our framework is thoroughly automatic without any intraoperative interaction, reducing setup time and cost at the same time.

In our experiment, the distance between surfaces registered by manually selecting points and our auto method averages to 3.613mm in real data tests, with a variance of 0.126. However, in synthetic data test, with scale changing no more than 2 times and noise intensity less than 0.03, the distance can be restricted within 1mm. Our registration algorithm is proved to maintain reasonable robustness, yet the stripe noise in reconstructed intraoperative depth map notably reduced the registration accuracy. According to the reconstruction process, the main errors in our current reconstructing system result from camera calibration and

coordinate transformation. Had these system errors fixed, the registration accuracy would be further improved, enabling us to realize a markerless framework with greater accuracy yet far lower cost than previous work has achieved. Therefore, our future work would focus on improving the reconstructing system for better quality of intraoperative depth map.

5. ACKNOWLEDGEMENT

This research is partially supported by the Chinese NSFC research fund (61190120,61190124,61271318 and 81301283) and biomedical engineering fund of Shanghai Jiao Tong University (YG2012ZD06).

6. REFERENCES

- [1] M. I. Miga, T. K. Sinha, D. M. Cash, R. L. Galloway, and R. J. Weil, "Cortical surface registration for image-guided neurosurgery using laser-range scanning," *IEEE Trans Med Imaging*, vol. 22, pp. 973-985, 2003.
- [2] T. M. Peters, "Image-guided surgery and therapy: Current status and future directions," in *Medical Imaging*, pp. 1-12, 2001.
- [3] R. Konietschke, A. Busam, T. Bodenmüller, T. Ortmaier, M. Suppa, J. Wiechnik, et al., "Accuracy identification of markerless registration with the dlr handheld 3d-modeller in medical applications," *Proceedings of CURAC*, vol. 6, 2007.
- [4] H. Schreiber and J. H. Bruning, "Phase shifting interferometry," *Optical Shop Testing, Third Edition*, pp. 547-666, 2006.
- [5] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *INT J COMPUT VISION*, vol. 60, pp. 91-110, 2004.
- [6] J. C. Gower and G. B. Dijksterhuis, "Procrustes problems" *Oxford: Oxford University Press*, vol. 3, 2004.
- [7] A. Vedaldi and S. Soatto, "Quick shift and kernel methods for mode seeking," in *ECCV*, pp. 705-718, 2008.
- [8] C. Studholme, D. L. Hill, and D. J. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," *Pattern Recogn*, vol. 32, pp. 71-86, 1999.
- [9] M. Toews and W. M. Wells III, "Efficient and robust model-to-image alignment using 3D scale-invariant features," *Med Image Anal*, 2012.
- [10] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, pp. 381-395, 1981.
- [11] Z.-H. Zhou and X. Geng, "Projection functions for eye detection," *Pattern Recogn*, vol. 37, pp. 1049-1056, 2004.
- [12] S. Zhang, Y. Zhan, M. Dewan, J. Huang, D. N. Metaxas, and X. S. Zhou, "Towards robust and effective shape modeling: Sparse shape composition," *Medical image analysis*, vol. 16, pp. 265-277, 2012.