

Volumetric Ultrasound Panorama Based on 3D SIFT

Dong Ni¹, Yingge Qu¹, Xuan Yang¹, Yim Pan Chui¹,
Tien-Tsin Wong¹, Simon S.M. Ho², and Pheng Ann Heng¹

¹ Department of Computer Science and Engineering,
The Chinese University of Hong Kong, Hong Kong SAR, China

² Union Hospital, Hong Kong SAR, China

Abstract. The reconstruction of three-dimensional (3D) ultrasound panorama from multiple ultrasound volumes can provide a wide field of view for better clinical diagnosis. Registration of ultrasound volumes has been a key issue for the success of this panoramic process. In this paper, we propose a method to register and stitch ultrasound volumes, which are scanned by dedicated ultrasound probe, based on an improved 3D Scale Invariant Feature Transform (SIFT) algorithm. We propose methods to exclude artifacts from ultrasound images in order to improve the overall performance in 3D feature point extraction and matching. Our method has been validated on both phantom and clinical data sets of human liver. Experimental results show the effectiveness and stability of our approach, and the precision of our method is comparable to that of the position tracker based registration.

1 Introduction

Panoramic imaging is one of the key technologies used in widening the field of view of medical ultrasound images for clinical diagnosis and measurement. 2D ultrasound panorama has been prevalent in routine clinical practices. Over the past few years, 3D ultrasound has been popular, and some research work on 3D ultrasound panorama has been conducted accordingly. Gee et al. [1] proposed an algorithm for the alignment of multiple freehand 3D ultrasound sweeps. An intensity based registration method was performed on a single dividing plane between two sweeps instead of on the entire overlapping region to improve the registration speed. Poon et al. [2] proposed a system to create a mosaic view from a set of ultrasound volumes acquired from a dedicated 3D probe, which provided higher imaging resolution in the elevation direction than a freehand probe. Two block-based registration methods were developed to correct the misalignment based on measurements obtained from position trackers. Wachinger et al. [3] focused on the global registration strategy among multiple ultrasound volumes. Unfortunately, intensity-based algorithms including sum of squared differences, normalized cross-correlation, mutual information, and correlation-ratio may not perform well on ultrasound data due to the low signal to noise ratio and the existence of shadows, speckles and other noises in ultrasound data. On the other

hand, position tracker based registration methods may not always be well-suited to routine clinical use, since the trackers may be affected by metals (for magnetic sensors) or sight occlusions (for optical sensors).

In addition to the noisiness of ultrasound images, there are other problems which make the registration of ultrasound volumes difficult. Since the ultrasound probe has to be arbitrarily oriented during the acquisition of multiple volumes, such variation of rotation imposes a great challenge to traditional registration methods in feature matching. Besides, the intensity of the same structure may differ in multiple ultrasound volumes obtained under different scan angles, and this affects the registration accuracy as well. Recently, SIFT in 2D space has been successfully applied to object recognition [4], point tracking [5], panorama creation [6], etc. Chen et al. [7] utilized SIFT for rapid pre-registration of multi-modality medical images. Moradi et al. [8] applied SIFT and a B-spline based interpolation technique for registering elastically deformed MRI and ultrasound images. Scovanner et al. [9] proposed a 3D SIFT descriptor for video analysis, and recognized actions based on the similarity of their feature representations. SIFT features are invariant to rotation and provide robust feature matching across a substantial range of addition of noise and change in illumination (intensity). This motivates us to apply SIFT based registration to ultrasound data.

In this paper, we adopt a modified 3D SIFT algorithm to create 3D ultrasound panorama. We first pre-process the ultrasound data to remove the artifacts in the ultrasound volume. Then 3D feature points are detected from multiple ultrasound volumes, and described as 3D SIFT features. Finally these features are matched for stitching multiple ultrasound volumes to form panorama.

2 Method

2.1 Ultrasound Preprocessing

An inherent characteristic of ultrasound imaging is the presence of shadows and speckles. Since the contact between a dedicated 3D probe and skin is not as good as that using a 2D probe, some of the emitted echoes are attenuated by air resulting in some large dark regions in the image. These artifacts may greatly affect the detection and matching of feature points. Therefore it is essential to remove them before registration. For speckle reduction, a well-known technique



Fig. 1. (a) the original ultrasound image, (b) the ultrasound mask generated from (a)

is to low-pass filter the image, e.g. by a Gaussian filter [10]. Since a Gaussian filter has been used in the SIFT algorithm, speckles are reduced effectively. The removal of shadows makes use of the fact that shadow is mainly caused by a loss of signal along the direction of the beam, due to strong acoustic reflection. We scan the ultrasound image along the echo direction from the bottom to the transducer surface. The image is labeled as shadows if certain threshold is reached. Large dark regions are segmented and removed using the Hidden Markov Random Field Model and the Expectation-Maximization algorithm [11]. The results of these two steps are combined to generate a ultrasound mask (Figure 1), which will be used in the process of feature detection.

2.2 3D Feature Detection and Descriptor Construction

In this subsection, we discuss the modified 3D SIFT approach.

Detection of 3D Keypoints. To detect keypoints efficiently, a Difference of Gaussian (DoG) scale-space is first built based on the original volume data [4]. Specifically, an input volume $I(x, y, z)$ is first convolved with a 3D Gaussian filter $G(x, y, z, k\sigma)$ of different scale, $k\sigma$, to obtain the scale space $L(x, y, z, k\sigma)$ i.e.,

$$L(x, y, z, k\sigma) = G(x, y, z, k\sigma) * I(x, y, z). \quad (1)$$

Then the difference of Gaussian (DoG) volumes is taken as:

$$D(x, y, z, k_i\sigma) = L(x, y, z, k_i\sigma) - L(x, y, z, k_j\sigma) \quad (2)$$

Figure 2(a) shows the construction process. The left column shows the volume pyramid, and each level is created by convolving the original volume with a

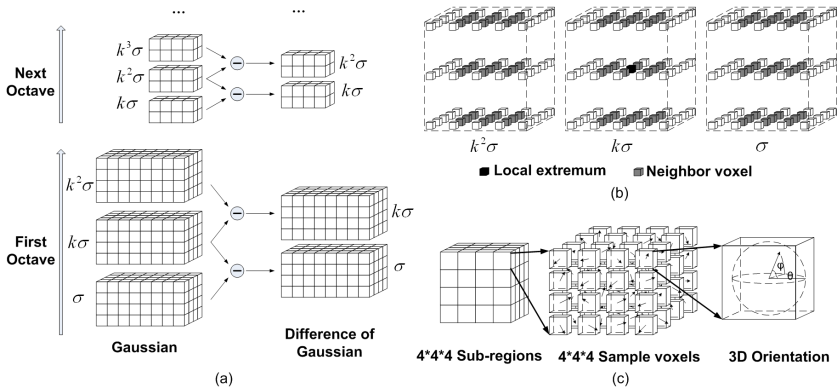


Fig. 2. (a) Volume pyramid: each level is created by convolving the original volume with Gaussian of various scale, and to generate DoG volumes from adjacent scales. (b) An extremum (in black) is the maximum or minimum of its 80 neighbors (in grey). (c) Construction of 3D SIFT descriptor. The 3D gradient orientation is shown (right).

Gaussian of scale $\sigma, k\sigma, k^2\sigma$ and so on. The right column shows the generated DoG from the difference of neighboring filtered volume on multi-scales.

Keypoint candidates are taken as the local maxima/minima of the DoG scale space $D(x, y, z, \sigma)$. To achieve this, each sample point is compared to its 80 neighbors, where 26 neighbors in the current volume, and each 27 neighbors in the scale above and below, as shown in Figure 2(b). A keypoint candidate is selected only if it is larger than all of its neighbors or smaller than all of them. If the position of the keypoint candidate is outside the ultrasound mask generated in the previous step, this candidate is eliminated. The next step is to eliminate the false candidates with low contrast, or those poorly localized along an edge. Edge responses in 3D space can be measured by a 3×3 Hessian matrix $H = (D_{ij})_{3 \times 3}$, which describes the curvatures at the keypoint in each 3D sub-plane, where D_{ij} is the second derivatives computed at the location and scale of the keypoint in the DoG space. A poor candidate usually presents a high principal value across the edge direction but a small value in the perpendicular direction. Given the ordered eigenvalues of H ($|\lambda_1| < |\lambda_2| < |\lambda_3|$) with corresponding eigenvectors (e_1, e_2, e_3), the eigenvectors define an orthogonal coordinate system aligned with the direction of minimal (e_1) and maximal (e_3) curvature. Then we refer a keypoint to be valid if the following constraint is satisfied: $r = \frac{|\lambda_3|}{|\lambda_1|} < T$. Since the ultrasound image is noisy and not as sharp as the normal image, we lessen this threshold restriction so as to obtain more keypoint candidates in 3D space. In our experiments, the threshold T is set as 15.

Construction of 3D SIFT Descriptor. Based on [9], we adopt a modified 3D-SIFT feature descriptor to volumetric ultrasound data. Each keypoint detected in the above section is first assigned with a dominant orientation (θ^*, ϕ^*) according to local image properties so that the keypoint descriptor can be represented relative to this orientation. As shown in Figure 2(c), the gradient orientation in 3D space is represented by two angles: θ denotes the longitude while ϕ denotes the latitude. With the dominant orientation, the neighborhood region surrounding the keypoint is then rotated so that the dominant orientation points to the direction of $\theta^* = \phi^* = 0$. The rotation matrix is defined as:

$$R = \begin{bmatrix} \cos \theta^* \cos \phi^* & \sin \theta^* \cos \phi^* & \sin \phi^* \\ -\sin \theta^* & \cos \theta^* & 0 \\ -\cos \theta^* \sin \phi^* & -\sin \theta^* \sin \phi^* & \cos \phi^* \end{bmatrix} \quad (3)$$

Then $4 \times 4 \times 4$ sub-regions surrounding each keypoint are sampled in the rotated neighborhood. The gradient magnitude and orientation of each sample in the subregions are calculated. For each sub-region, the magnitude of the gradient, weighted by a Gaussian window centred at the keypoint, is added to the corresponding bin for the gradient orientation, where 8 bins (represent 360 degrees) are used for θ and 4 bins (represent 180 degrees) are used for ϕ . Finally, the feature vector of each keypoint is of $4 \times 4 \times 4 \times 8 \times 4 = 2048$ dimensions.

2.3 Ultrasound Volume Registration and Stitching

Once the 3D SIFT features are extracted from each volume, multiple ultrasound volumes then can be matched, and stitched into a 3D panorama. Given two ultrasound volumes A and B , a set of keypoints is first obtained for each volume. Usually, there exist thousands of keypoints, and they have to be validated before being used in the calculation of the transformation matrix. For one particular keypoint in volume A , its Euclidean distance between each keypoint in volume B is calculated. A ratio is assigned to each keypoint by comparing the distance of the closest neighbor to that of the second-closest neighbor. A keypoint is recognized as a stable feature only if its ratio is larger than a threshold. Since we apply the 3D SIFT algorithm in a single modality (ultrasound only) registration, the difference of the intensity between the two correctly matched keypoints is usually not very large. Therefore, we further eliminate false matching pairs by comparing the intensity difference between the two keypoints by using the histogram distance of greylevels within a 8×8 region around the keypoint.

We use RANSAC [12] to simultaneously solve the correspondence problem and estimate the transformation matrix related to every pair of volumes. The advantage of RANSAC is its ability to perform robust estimation with a high degree of accuracy even when outliers are presented. Once pairwise matches have been established between the ultrasound volumes, we then can obtain the global transformation matrix for each volume and stitch them into a panorama.

3 Results and Discussion

Data Acquisition. All ultrasound volumes in our experiments were acquired using a GE Voluson 730 ultrasound scanner with a dedicated 3D ultrasound probe (RAB2-5L Convex Volume Probe). Two sets of data were collected, where one included 5 ultrasound volumes obtained from a phantom (CIRS Model 057, mimicking human liver tissues) with different scan angles, another set included 3 ultrasound volumes obtained by scanning the liver of a patient. In our experiments, about 50% overlapping was used to guarantee sufficient number of keypoint pairs appear in neighboring volumes. In order to evaluate the precision of ultrasound panorama, the corresponding CT volume was acquired from a GE Lightspeed 16-slice multi-detector on the same phantom.

Stitching of Multiple Ultrasound Volumes. We have applied the proposed method on data sets of the phantom and clinical liver. Most of the parameters stated in the paper are specified according to the state-of-the-art 2D SIFT algorithms [4]. Figure 3(a) and 3(c) shows part of the matched features in phantom and clinical data sets respectively. Note that the data is actually a volume, a multi-planar reconstruction view is used to illustrate the matched feature point pairs. The cross represents a 3D SIFT feature, and each line links a pair of matched feature points. Figure 3(b) shows the resultant ultrasound panorama (one of the cross-sections is shown) obtained by stitching five phantom volumes. Figure 3(d) shows the ultrasound panorama obtained from the stitching of three

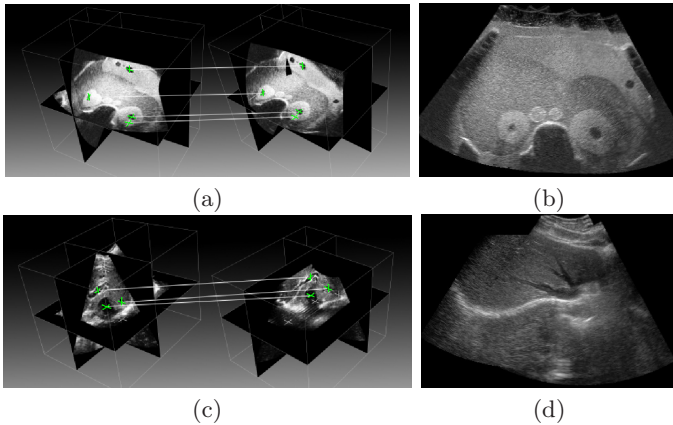


Fig. 3. (a), (c) 3D feature point matching in phantom and clinical data respectively. (b), (d) Stitched results of the phantom and clinical data sets.

clinical liver volumes. It takes about one minute to register two neighboring volumes ($256 \times 191 \times 248$) using a 3GHz Intel Pentium CPU and 3G RAM-based system.

Precision of Ultrasound Panorama. An important clinical application of ultrasound panorama is to provide a quantitative measurement of large organs. To verify the precision of the proposed method, we use the CT volume, obtained by scanning the same phantom, as a ground truth for evaluating the ultrasound panorama generated in our system. The CT volume covers the whole phantom. In both the ultrasound panorama and CT volume, distinguishable feature points were selected manually from the same anatomical position. Physical distance between each point pair, within the same modality, is calculated. Figure 4 shows the measurement of one line in the CT volume and that of another corresponding line in the ultrasound panorama. This physical distance provides a quantitative measurement of the distortion resulted from stitching multiple ultrasound volumes. Table 1 lists the measurement results. The length of three lines are measured in our experiments. The mean error is 2%, which is comparable to the 2.5% error in the results based on a position tracker [2].

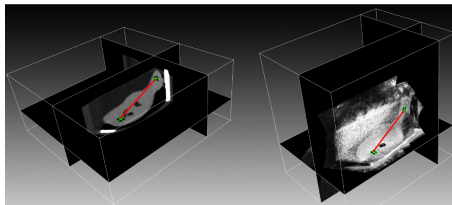
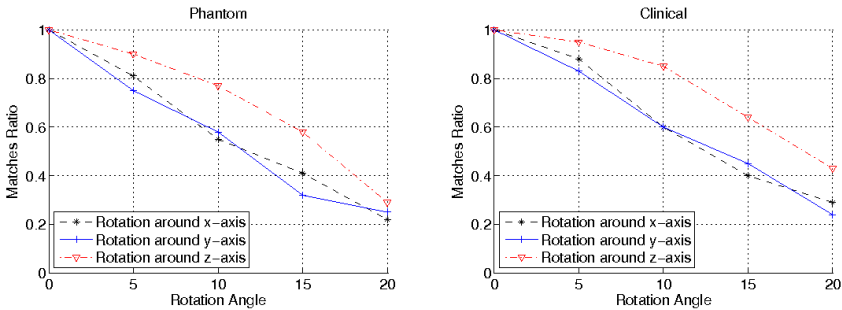


Fig. 4. Physical distance between pairs of distinguishable feature points (shown by crosses) from both CT (right) & US (left) are measured

Table 1. Precision Evaluation. Physical distances of three lines are measured respectively in CT volume and ultrasound panorama.

Line	Length in CT(mm)	Length in US(mm)	Error
1	138.84	141.34	1.8%
2	155.77	152.48	2.1%
3	130.93	128.23	2.0%

Stability of Feature Matching with Rotation. In this experiment, we evaluate the stability of feature matching under rotations. Given two ultrasound volumes A and B , which are first registered. B is resampled according to the transformation matrix between A and B to obtain a resampled volume B' . We rotate volume B' with various angles around x-axis, y-axis and z-axis respectively. The numbers of matched keypoints between A and the rotated B' are then calculated for each of these rotations. The result is normalized with respect to the number of matched keypoints under no rotation, to show the stability under rotation. In Figure 5, the left sub-figure shows the results from phantom data sets and the right one shows the results from clinical data sets. As shown in this figure, when the rotation angle is below 10° , the percentage of matched feature points is above 50%. When the rotation angle is around 15° , the ratio can still be maintained at around 40%. It is observed that when the volume is rotated around z-axis, the results outperform others. One possible reason is that the image resolution along z-axis is not as good as that along x-axis or y-axis.

**Fig. 5.** Stability of matched features for rotated ultrasound volumes. The volume is rotated by 0° , 5° , 10° , 15° , 20° around x-, y- and z-axis respectively.

4 Conclusion

In this paper, we have proposed a modified 3D SIFT based method for generating ultrasound panorama from multiple volume data sets. Ultrasound volumes are pre-processed to reduce the false detection and matching of keypoints obtained from 3D SIFT. Our approach permits the generation of panorama with high matching stability, because the proposed 3D SIFT feature has merit of invariance to rotation and provides robust feature matching across a substantial range of

addition of noise and change in illumination. We have tested the precision of our approach by comparing the stitched ultrasound volume with the CT volume, and results have shown promising feasibility of our method in application to rigid registration of ultrasound volumes. Although ultrasound is mostly used in soft-tissue imaging, the registration can be approximated as a rigid one in some cases, e.g. when patients are required to hold their breath during biopsy procedures. Besides, our method can serve as an essential step of calculating a global alignment [3] in registering non-rigid motions. We will improve the precision and apply our proposed method to non-rigid registration by integrating block based registration algorithm [2] to our method. We'll also improve the speed and stability according to characteristics of ultrasound. The long-term goal would be to develop a fully automatic 3D ultrasound panorama generation system with no need to use position tracker under routine clinical conditions.

Acknowledgement

The work described in this paper was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region (Project No. CUHK4461/05M). This work is also affiliated with the Virtual Reality, Visualization and Imaging Research Center at CUHK as well as the Microsoft-CUHK Joint Laboratory for Human-Centric Computing and Interface Technologies.

References

1. Gee, A., Treece, G., Prager, R., Cash, C., Berman, L.: Rapid registration for wide field of view freehand three-dimensional ultrasound. *IEEE Trans. Med. Imaging* 22(11), 1344–1357 (2003)
2. Poon, T., Rohling, R.: Three-dimensional extended field-of-view ultrasound. *Ultrasound in Medicine and Biology* 32(3), 357–369 (2003)
3. Wachinger, C., Wein, W., Navab, N.: Three-dimensional ultrasound mosaicing. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) *MICCAI 2007, Part II. LNCS*, vol. 4792, pp. 327–335. Springer, Heidelberg (2007)
4. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision* 60(2), 91–110 (2004)
5. Lazebnik, S., Schmid, C., Ponce, J.: Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In: *CVPR*, pp. II: 2169–2178 (2006)
6. Brown, M., Lowe, D.: Recognising panoramas. In: *Proceedings of International Conference on Computer Vision*, vol. 2, pp. 1218–1225 (2003)
7. Chen, J., Tian, J.: Rapid multi-modality preregistration based on sift descriptor. In: *Proceedings of International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1437–1440 (2006)
8. Moradi, M., Abolmaesoumi, P., Mousavi, P.: Deformable registration using scale space keypoints. In: *SPIE Medical Imaging*, vol. 6144, pp. 791–798 (2006)
9. Scovanner, P., Ali, S., Shah, M.: A 3-dimensional sift descriptor and its application to action recognition. In: Lienhart, R., Prasad, A.R., Hanjalic, A., Choi, S., Bailey, B.P., Sebe, N. (eds.) *ACM Multimedia*, pp. 357–360. ACM Multimedia, New York (2007)

10. Penney, G., Blackall, J., Hamady, M., Sabharwal, T., Adam, A., Hawkes, D.: Registration of freehand 3d ultrasound and magnetic resonance liver images. *Medical Image Analysis* 8, 81–91 (2004)
11. Zhang, Y., Brady, M., Smith, S.M.: Segmentation of brain mr images through a hidden markov random field model and the expectation maximization algorithm. *IEEE Trans. Med. Imaging* 20(1), 45–57 (2001)
12. Hartley, R., Zisserman, A.: *Multiple View Geometry in Computer Vision*, 2nd edn. Cambridge University Press, Cambridge (2003)