Achieving High Multi-modal Registration Performance Using Simplified Hough-Transform with Improved Symmetric-SIFT

Md. Tanvir Hossain, Shyh Wei Teng, Guojun Lu
Gippsland School of Information Technology
Monash University
Churchill, Victoria 3842, Australia
Email: {Tanvir.Hossain, Shyh.Wei.Teng, Guojun.Lu}@monash.edu

***Abstract*—The traditional way of using Hough Transform with SIFT is for the purpose of reliable object recognition. However, it cannot be effectively applied to image registration in the same way as the recall rate can be significantly lower. In this paper, we propose an alternative implementation of Hough Transform that can be used with Improved Symmetric-SIFT for multi-modal image registration. Our experimental results show that the proposed technique of applying Hough Transform can significantly improve the key-point matching as well as registration accuracy by utilizing aggregated information from key-points throughout the input images.**

I. Introduction

The problem of image registration is to align two given images where the transformation between the images is unknown. Rotation, scaling and translation are the three fundamental types of transformation that are commonly seen in general registration problems. Multi-modal image registration, on the other hand, is a special kind of registration problem where the input images might be captured by different types of sensors. The visual difference in the images caused by the difference in the types of imaging sensors is an added complexity to the standard registration problem. Common examples of multi-modal image registration include Computed Tomography (CT) $\longleftrightarrow$ Magnetic Resonance Imaging (MRI), CT $\longleftrightarrow$ Positron Emission Tomography (PET), CT $\longleftrightarrow$ Single Photon Emission Computed Tomography (SPECT), PET $\longleftrightarrow$ MRI, PET $\longleftrightarrow$ Ultrasound (US) and Electro-Optical (EO) $\longleftrightarrow$ Near Infrared (NIR).

Multi-modal registration techniques found in the literature can broadly be categorised into two groups – Global and Local. Most global techniques [1], [2], [3], [4], [5], [6] are based on some form of statistical measures e.g., Mutual Information (MI) [7] and correlation. These measures are generally computed based on the entire content of the given images and indicate how good the alignment of the images is in a particular test iteration. At each test iteration the alignment is re-estimated and the measurement is re-computed. Finding the correct transformation, therefore, becomes an optimisation problem which can be computationally very expensive to solve. Moreover, based on the range of possible transformations the search space for this optimisation problem can be too large to explore. This is why many of them [8], [9], [10], [11], [12], [13] require an initial estimation to be provided manually in order to narrow down the search space. Yet they may fail if the initial estimation is not close enough to the correct values. Global techniques also tend to be affected by truncation and outliers.

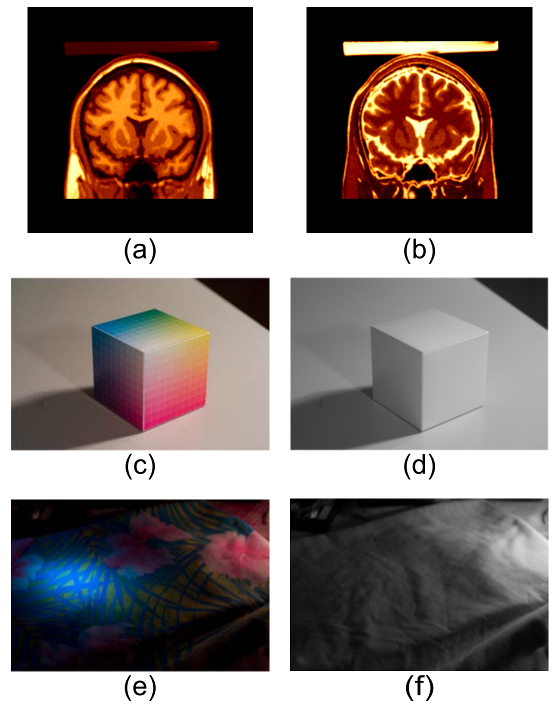

Fig. 1. A few pair of multi-modal images. (a,b) T1 vs T2 weighted coronal section MR images, (c,d) and (e,f) Electro-optical vs Infra-red images

In contrast, local description techniques are more invariant to affine transformations and are less affected by the presence of outliers, occlusion, clutter, truncation and low overlap of images. Despite the strength of using local description techniques, very few have been studied so far for multi-modal registration.

Improved Symmetric-SIFT (ISS) has been proposed in [14] that can be used to effectively register multi-modal images.

It is one of the very few local description techniques [15], [16], [17] that are invariant to multi-modality. The ancestors to these techniques is SIFT [18], [19]. Though SIFT is one of the most popular local description techniques, it is not designed to cater for variations in modality. This is why SIFT does not work well for multi-modal images, though its performance is remarkably good for single modal images. In order to make up for this deficiency, techniques like Symmetric-SIFT, MI-SIFT and ISS have been proposed recently. We found ISS to be the most mature in terms of its effectiveness.

One of the important steps in local registration techniques is 'Descriptor Matching'. Nearest Neighbour (NN) matching, Similarity matching and Nearest Neighbour Ratio (NNR) matching are some of the commonly used matching algorithms. However, these matching algorithms compare the descriptors on a one-to-one basis and do not make use of any information from other descriptors. Should additional spatial information (such as location, scale and orientation of the descriptors) be considered at the time of matching, better accuracy could be achieved. A more advanced technique that exactly does the later part is Hough Transform [20], [21], [22], as demonstrated in [19] where Hough Transform has been used with SIFT for object recognition. Though this technique has great potential, it has not yet been used with ISS. Neither it can be directly used in ISS as there are some associated problems that we shall describe in Section III. In this paper we not only demonstrate how these problems can be overcome, we also show how Hough Transform can be seamlessly integrated into ISS by harnessing its inherent information. As a result, higher descriptor matching and registration accuracy is achieved.

The rest of this paper is organised as follows. Section II gives an overview on SIFT, Symmetric-SIFT, ISS and Hough Transform. The implications of directly using Hough Transform with ISS is discussed in Section III. Section IV presents our proposed solution. Section V describes the experimental results and we finally conclude this paper in Section VI.

II. Background

In this section we will give an overview on SIFT, Symmetric-SIFT and ISS. This will help in understanding the working principle of ISS as well as the problem which we described in Section III.

A. SIFT

SIFT [19] is a key-point detection and description technique in image registration, retrieval and object recognition. It has become one of the most popular techniques used in image processing and computer vision due to the robustness of its key-points and the distinctive nature of its descriptors.

SIFT descriptors are built on key-points that are identified by applying Difference of Gaussian (*DoG*) in scale space. Its detection process attributes each key-point with a scale parameter which is further used to determine the area of the key-regions. A key-region is the portion of the image surrounding a given key-point whose content is encoded to build the descriptor. The use of the scale parameter, thus allows to build the descriptors in a scale-invariant way. Rotation invariance, on the other hand, is achieved by building the descriptors relative to their dominant orientations. Dominant orientation is the direction in which most of the gradients of a particular key-region are oriented. For a given key-region, this can be computed by finding the peak in an orientation histogram that summarizes all the gradient directions found therein. Finally, the descriptor is built on a 4-by-4 spatial grid where each cell in the grid consists of an 8-bin orientation histogram. All gradients within a cell are quantized into one of these 8 bins resulting in a descriptor having (4×4×8 or) 128 dimensions. Descriptors from a given image pair can then be compared to derive the key-point matching set.

$$D = \left[\begin{array}{cc|cc} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \\ \hline h_{31} & h_{32} & h_{33} & h_{34} \\ h_{41} & h_{42} & h_{43} & h_{44} \end{array}\right] \qquad D_r = \left[\begin{array}{cc|cc} h_{44} & h_{43} & h_{42} & h_{41} \\ h_{34} & h_{33} & h_{32} & h_{31} \\ \hline h_{24} & h_{23} & h_{22} & h_{21} \\ h_{14} & h_{13} & h_{12} & h_{11} \end{array}\right]$$

(a) (b)

$$f(D) = \left[\begin{array}{cc|cc} h_{11}+h_{44} & h_{12}+h_{43} & h_{13}+h_{42} & h_{14}+h_{41} \\ h_{21}+h_{34} & h_{22}+h_{33} & h_{23}+h_{32} & h_{24}+h_{31} \\ \hline h_{31}\sim h_{24} & h_{32}\sim h_{23} & h_{33}\sim h_{22} & h_{34}\sim h_{21} \\ h_{41}\sim h_{14} & h_{42}\sim h_{13} & h_{43}\sim h_{12} & h_{44}\sim h_{11} \end{array}\right]$$

(c)

Fig. 2. (a) and (b) represents the histogram arrangement in D and D_r respectively. (c) shows the final merged descriptor.

However, there are a couple of important properties of multi-modal images that the original SIFT does not take care of. These properties are, namely, Gradient Reversal and Region Reversal [14]. The pattern of intensity levels may dramatically vary in multi-modal images due to variation in sensor types. Because of this, gradients at corresponding locations of a pair of multi-modal images may have opposite orientations. This property is called Gradient Reversal. Region Reversal, on the other hand, is a consequence of Gradient Reversal when rotation normalisation by dominant orientation is attempted. An image region may remain misaligned (by 180°) even after rotation normalisation and such misalignment of a region is called Region Reversal. The presence of Gradient Reversal and Region Reversal [14], cause SIFT to fail detect correspondences between multi-modal images. This is why SIFT is not suitable for multi-modal image registration.

B. Symmetric-SIFT

Symmetric-SIFT [15] is a variation of SIFT with the added capability of being modality invariant. It is very similar to the original SIFT with some important differences in the descriptor building process. One difference is that, it forces all the gradients to remain within the range $[0, \pi)$ according to the following rule.

$$\forall\ \theta \in (\pi, 2\pi] : \theta = \theta - \pi \tag{1}$$

This enables Symmetric-SIFT to address the Gradient Reversal problem. In order to cater for Region Reversal, on the other hand, it builds a pair of intermediate descriptors for every key-region. These intermediate descriptors have their special bins rotated by 180° (See Figures 2 (a) and (b)). They are finally merged to produce the actual Symmetric-SIFT descriptors (Figure 2 (c)). The following equation shows the merging function used in this process.

$$f(D) = [D^u + D^u_r \quad D^l \sim D^l_r] \tag{2}$$

Here D^u and D^u_r are the upper halves of the two intermediate descriptors D and D_r respectively. D^l and D^l_r, on the other hand, are the lower halves of D and D_r respectively.

As in this technique, both the Gradient Reversal and Region Reversal problems are taken care of, unlike SIFT, Symmetric-SIFT can be used for multi-modal registration. However, as identified in [14], there are problems associated with the merging function (Equation 2) of Symmtric-SIFT that have adverse effect on its distinctiveness. The authors in [14], therefore, came up with a different approach (called ISS) to solve the above mentioned problem.

C. ISS

ISS is basically a two-phase modality invariant description building technique that is aimed to overcome the descriptor merging problem in Symmetric-SIFT. The first phase is similar to Symmetric-SIFT. However, the second phase of ISS produces the final set of descriptors by considering an appropriate global rotation. This rotation parameter is estimated in between the two phases of ISS by analysing the outcomes of the first phase. Consideration of the rotation parameter is critical to ISS, as this allows the technique to build descriptors in a much simpler and less error prone way. We will also see in Section IV-B that this pre-estimated rotation parameter also makes ISS a suitable candidate for Hough Transform.

As presented in [14], the performance of ISS is clearly better than both SIFT and Symmetric-SIFT. In this paper we will show how even better performance can be achieved by applying Hough Transform with ISS. It may be noted that, though ISS provides some positive aspects for easy integration of Hough Transform, it cannot be done straight away in its traditional form. The reasons are explained in further detail in Section III.

D. Hough Transform

1) Overview: Hough Transform is a systematic voting procedure which can be used to identify shapes in a given image. What more interesting is, this voting mechanism is performed in parameter space, i.e., the Hough Space. As long as a shape has a well defined or at least an approximate parametric representation, it can be identified by Hough Transform. The number of parameters needed to represent a shape determines the extent of the Hough Space. For example, a line parametrized into Normal Form according to Equation 3 has two parameters and, therefore, the respective Hough Space is two-dimensional.

$$x \cos\theta + y \sin\theta = r \tag{3}$$

Now, the following properties will clarify our understanding on how the Cartesian Space (or the X-Y Space) is related to our example Hough Space (i.e., the r-θ Space):

- Each point (x_i, y_i) in X-Y Space corresponds to a sinusoidal curve in the r-θ Space.[1]
- If two or more points in the X-Y Space belong to the same line L, their corresponding sinusoidal curves in the Hough Space will all intersect at a single point P.
- The r-θ parameters at P defines the line L in X-Y Space.

To detect a line L in X-Y Space, Hough Transform computes the value of r for each point (x_l, y_l) that lies on L and for each of the quantised value of θ using Equation 3. The value of r and θ at each iteration can thus be used to vote for in the Hough Space. If there are N_L points on line L and N_θ different values for θ in the quantised Hough Space, then there would be a total of $N_L \times N_\theta$ votes. Finally, the peak bin in the Hough Space is determined. Ideally there should be N_L votes in this peak bin and the corresponding (r, θ) values when substituted in Equation 3 would define the line L in question.

2) Application of Hough Transform in SIFT: Hough Transform [20], [21], [22] has traditionally been used with SIFT [19] to detect reliable matches for the purpose of object recognition. In SIFT, Hough Transform has been used to identify clusters of SIFT key-points that share a common transform. Scale, rotation and (x, y) translations are the four parameters that have been used in that application of Hough Transform. The four-dimensional Hough Space has 12 bins for orientation (30 degrees each), a factor of two for scale and a quarter times the maximum training image dimension for location. The technique is applied on the matching set obtained from the NNR matching of SIFT descriptors. In order to cater for boundary effects, a total of 16 bins are populated for each of the key-point matching pair, i.e. an additional bin populated on each of the four dimensions. Peaks detected in the Hough Space indicate the likelihood of the existence of the query object(s). The application of Hough Transform allows to eliminate potential false matches even if the ratio of false to true matches is significantly high.

III. Disadvantages of Using Hough Transform with ISS in its Traditional Form for Image Registration

A. Over-Sensitivity to Scale and Lower Recall

The traditional way of using Hough Transform in SIFT is highly sensitive to variations in scale. In this section we will explain this problem in greater detail. Besides, we will also address, why this is less of a concern when Hough Transform is used with SIFT for object recognition, however, a major issue for image registration.

We have already mentioned about the four parameters required to build the Hough Space in Section II-D2. The scaling

[1]This is because, for all possible lines that can be drawn through a given point in $X - Y$ Space, there would be a distinct corresponding point in the r-θ space and all these points in the $r - \theta$ Space eventually form a sinusoid.

factor and orientation parameters used in Hough Transform are those derived during the scale-space analysis and orientation assignment phases. However, the translation parameter has to be geometrically derived using the information already at hand.

It can be shown that, the derived translation parameters can be greatly affected by the amount of noise present in the pre-detected scale. The extent of the error inferred into the translation estimation is linearly proportional to the noise in scale as well as the distance of the key-point from origin. This is explained in Equation 4 below,

$$(\tau \sim \tau_H) \propto \sqrt{(\sigma \sim \sigma_{DoG}) \times R}\,, \tag{4}$$

where, τ and σ are the ground truth translation and scaling, respectively, τ_H is the amount of translation identified by Hough Transform, σ_{DoG} is the scale detected by the Difference of Gaussian technique [18] and R specifies how far the point under consideration is from the origin.

To better quantify the amount of possible error, here we give an example. In a typical 500 by 500 pixel image, over 140 pixel error in translation can occur for 20% error in scale estimation. In other words, for any given rotation and scale factor, the derived translation can fluctuate by over 140 pixels if the scale factor varies by 20% only. Moreover, as the scaling bins are separated by a factor of 2, theoretically (in worst case) it is possible to have 100% error in scale detection, which can produce more than 700 pixel error in translation for the same 500 by 500 pixel image.

To summarize, the more the error in scale, the wider the error range in the derived translation will be. And the wider the error range in translation, the higher is the probability of drifting into an incorrect translation bin in the Hough Space. On the other hand, as the number of bins in the Hough Space is very high, the probability of having a false positive (i.e., incorrectly casting a vote to the right bin) is very low. This is why the precision of the matching set obtained by Hough Transform in SIFT is good enough, whereas the recall can be significantly low.

The application of Hough Transform in SIFT is to enhance reliability in key-point matching for the purpose of object recognition. It is important to note here that, for object recognition, recall is of less concern as long as at least three matches can be detected with absolute precision. To get a better registration result, on the other hand, it is necessary to have better rates for both precision and recall.

B. *Higher Memory Requirement*

The memory requirement of Hough Transform can be very high, too. This is because, the Hough Space has as many dimensions as the number of parameters that we want to consider. According to what is proposed in [19], it is necessary to consider four transformation parameters. Two of them are for the horizontal and vertical translations and the other two are for scaling and rotation. Considering the number of bins used to track each of these parameters, the size of the Hough Space can contain as many as $12 \times 17 \times 21 \times 21$ bins. These include 12 orientation bins ($[0, 2\pi)$, 30 degrees each), 17 scale bins and 21 translation bins for both horizontal and vertical dimensions ($[-10w, 10w]$, where w is one-fourth of the maximum image dimension).

In Section IV we present a modified approach to applying Hough Transform which significantly improves both precision and recall in key-point matching by overcoming the above mentioned problems.

IV. Proposed Technique

In the last section we have described why Hough Transform cannot be directly applied to ISS in its traditional form. In Section IV-B we will show how we overcome these problems. Prior to that, in Section IV-A, we propose to make use of an alternate approach to building orientation histogram in ISS.

A. *Using Alternate Weighting Strategy to Populate Orientation Histogram Bins*

We have seen in Section II-A that SIFT descriptors are collections of orientation histograms. These histograms are the core building blocks of SIFT that captures intensity patterns from the image. However, as depicted in [23], adding magnitudes to the orientation histograms of SIFT may not appropriately correspond to the actual visual appearance of an image region. The presence of a gradient actually indicates the presence of a possible edge. Gradient magnitude, on the other hand, indicates how sharp the associated edge could be. It has been shown that, should the histogram bins accounted for gradient occurrences instead, more accurate matching can be obtained. In this paper we, therefore, modify ISS by making use of gradient occurrence histograms to characterize intensity patterns. We call this modified technique as ISS-O for future reference.

B. *Incorporating Hough Transform into ISS*

1) Addressing the recall and memory requirement issues: We have discussed about Hough Transform in Section III. In this section we will see, how this technique can be fitted within ISS in a more simplified fashion. Considering ISS to be a two-phase technique, we propose to integrate Hough Transform in its second phase. The incorporation of Hough Transform into ISS is outlined in Figure 3.

The Hough Space as used in [19], is a four-dimensional space based on the four basic transformation parameters – rotation, scale and translation in either vertical or horizontal direction. We, however, use only the translation parameters to build our Hough Space. A good estimation of orientation is already available from ISS upon completion of its initial phase.[2] It is the last step of the first phase of ISS as depicted in Figure 3. A better scale, on the other hand, can be acquired during the second phase of the proposed technique by analysing the scales in the NNR matching set. To explain

[2]In ISS, this orientation is used for rotation normalisation and for building the final set of descriptors.

this process, let us assume the output from the NNR matching be

$$\{D_{i1} \rightarrow D_{j1}, D_{i2} \rightarrow D_{j2}, D_{i3} \rightarrow D_{j3}, \\ ..., D_{in} \rightarrow D_{jn}\},$$

where i and j denote the indices of the key-points from the two input images *I1* and *I2* respectively and n is the total number of matches identified. Let $\{\sigma_{i1}, \sigma_{i2}, \sigma_{i3}, ..., \sigma_{in}\}$ be the associated scale factors for $\{D_{i1}, D_{i2}, D_{i3}, ..., D_{in}\}$ and $\{\sigma_{j1}, \sigma_{j2}, \sigma_{j3}, ..., \sigma_{jn}\}$ be the scale factors for $\{D_{j1}, D_{j2}, D_{j3}, ..., D_{jn}\}$.

Now assuming the difference in scale between *I1* and *I2* is σ, we can say,

$$\sigma_j / \sigma_i = \sigma$$

However, as in a real scenario, the σ values tend to vary, it is necessary to pick a central value for the scale. We, therefore, compute $median(\sigma)$ and consider it to be the scale difference σ_D between the two given images *I1* and *I2*. In this way, we make the scale factor constant, too.

To further ensure that the estimated scale factor is a very close approximation to the ground truth, we only consider those matches that have a higher NN ratio.

The orientation and scale parameters are not directly used in our proposed Hough Space, however, they are used to derive the translation instead. This reduction in Hough Space dimension not only reduces the memory requirement for the computation (explained in Section III-B), the use of an appropriate and constant scale factor also eliminates the errors in Hough Transform caused by minor inaccuracies in scale (as described in Section III-A).

2) Achieving more accurate matching using weighted voting in Hough Space: We also use a custom weighting parameter W to weight the votes in Hough Space. We have mentioned in Section II-D2 that the input to the Hough Transform is the matching set found from NNR matching of the descriptors. Now let us assume, R is the maximum ratio that returns a match M from the NNR matching. We set W linearly proportional to R when voting for M in the Hough Space. Matches obtained against higher NN ratio are indeed very distinctive matches. The proposed weighting strategy, thus emphasises votes from stronger matches and reduces the impact of votes from relatively weaker matches, making the voting procedure more rational.

With the application of Hough Transform in the proposed technique, the descriptor matching is no longer a simple point-to-point matching. The transform forces to consider the relationship among key-points in terms of rotation, scaling and translation. The experimental results presented in Section V further demonstrates the efficacy of applying our proposed technique.

We combine both the approaches outlined in IV-A and IV-B and refer the final solution as ISS-OH (or, ISS with improved Orientation histograms and simplified Hough transform).

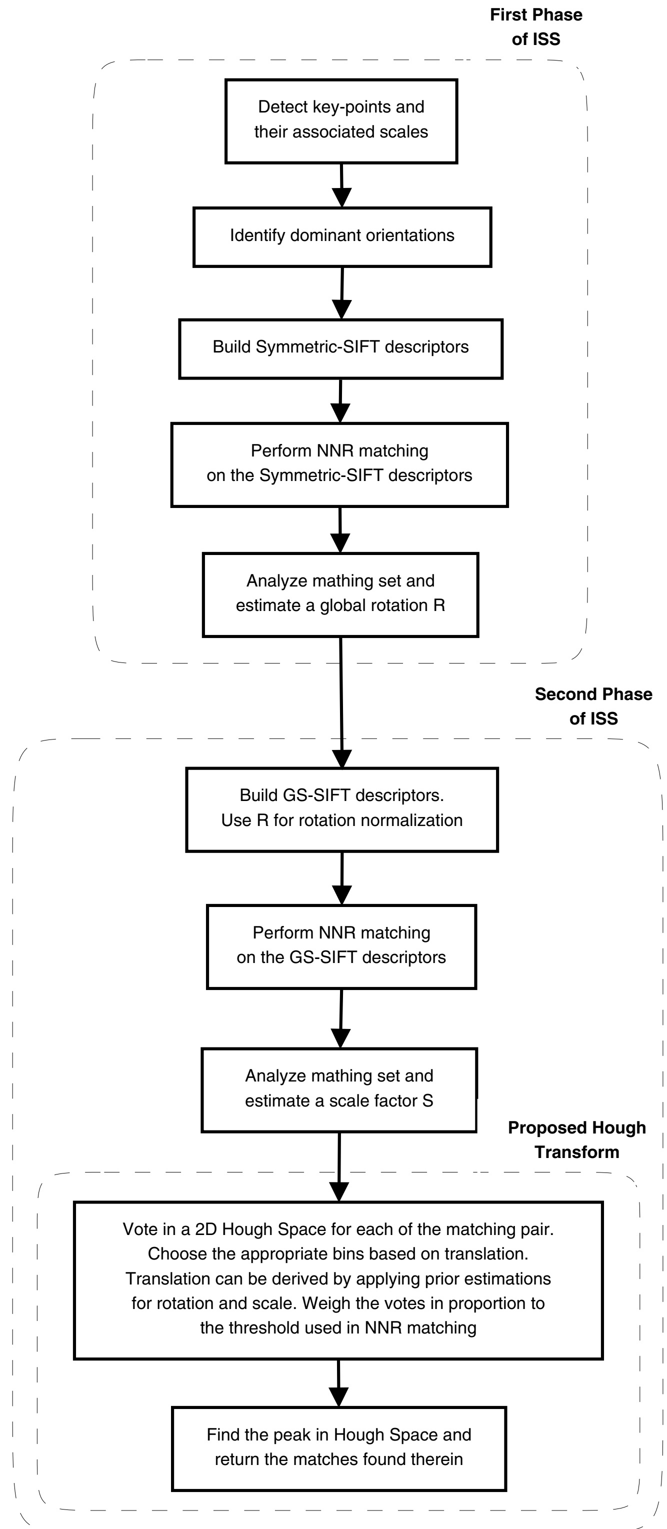

Fig. 3. Shows how Hough Transform is fitted within ISS. Out of the four pose parameters, only translation is used in the transform. Values determined at an earlier stage for orientation and scale are used to track for translation.

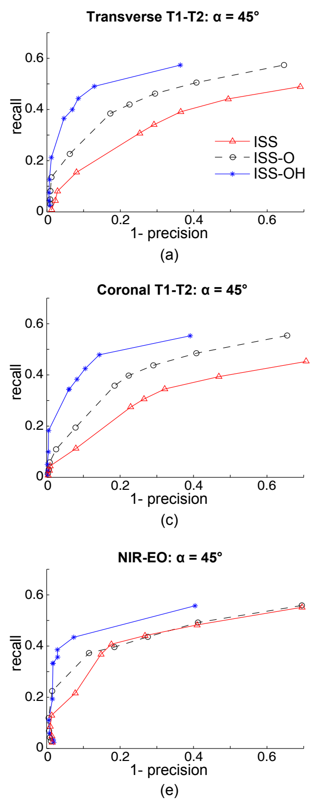

Fig. 4. Comparison between ISS, ISS-O and ISS-OH. *Recall* versus *1-Precision* curves for 45 ° and 60 ° rotational differences.

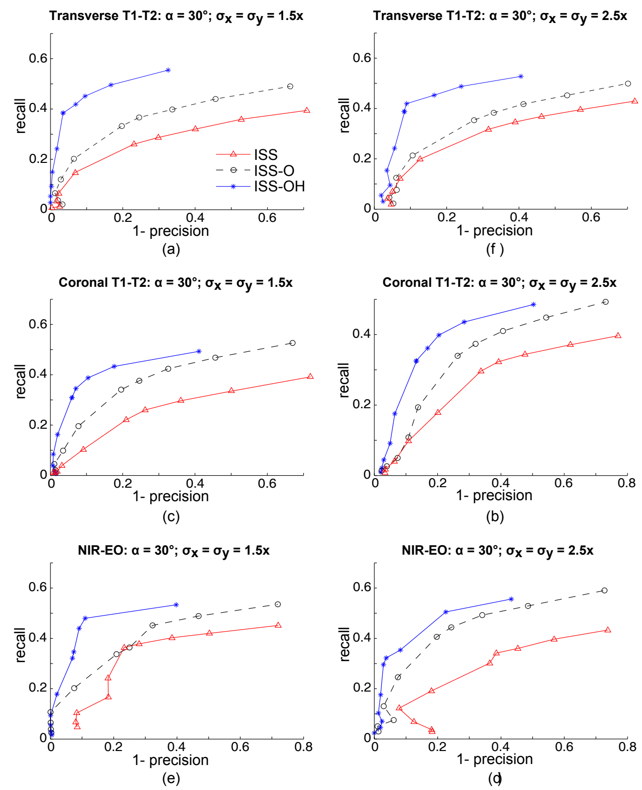

Fig. 5. Comparison between ISS, ISS-O and ISS-OH. *Recall* versus *1-Precision* curves for rotation and uniform scale transformations.

V. Experimental Results

We evaluate our technique with *Recall* versus *1-Precision* curves [24] where

$$Recall = \frac{\text{number of true matches}}{\text{number of actual correspondences}} \tag{5}$$

and

$$1 - Precision = \frac{\text{number of false matches}}{\text{number of total matches}} \tag{6}$$

Our test datasets constitute of transverse and coronal T1-T2 weighted brain MRIs collected from McConnell Brain Imaging Centres online brain image data store [25] as well as Near Infra-Red (NIR) versus normal Electro-Optical (EO) images from several other sources [17], [24], [26], [27]. There are 101, 87 and 18 image pairs, respectively in these image datasets. Whereas our brain image dataset prominently observes Gradient Reversal between the two modalities, the NIR vs EO image dataset observes it occasionally. It is important to note that, no dataset specific parameter tuning is required in any of our test cases which proves the versatility of the proposed technique.

In [14], the performance of ISS (Improved Symmetric-SIFT) has been evaluated against Symmetric-SIFT [15]. According to the authors, the performance of Symmetric-SIFT is the best among the very few multi-modal registration techniques and, therefore, has been used as the benchmark technique. Hence, in this paper we compare our proposed techniques (ISS-O and ISS-OH) against ISS.

Figures 4 (a)-(f) present the performance comparison between ISS, ISS-O and ISS-OH in terms of *1-Precision* and *Recall*. We had carried out experiments with a wide range of rotational and scale differences with consistent results. However, here we present results for images having 45 and 60 degrees of rotation only. Later, Figures 5 (a)-(f) show the results where the input images have both rotation and uniform scaling. Finally, in Figures 6 (a)-(f) we show another set of results. In this case, unlike the previous case we even used non-uniform scaling (unequal scaling in X and Y directions).

From the experimental results, it can be summarized that, ISS-O shows clear improvement over original ISS. ISS-OH, on the other hand, performs the best among the three techniques compared. Note that there are a few curves having unusual bends at their lower parts (eg. the red curve in Figure 6 (f)) indicating an increase in precision with an increase in recall. This happens because in these cases the number of key-point matches identified (corresponding to the lower part of the curves) are very low and the number of correct matches in the ground truth is not a large number either. As a result, at the lower part of some of the curves, the measurement of precision and recall (itself) may not be stable enough. However, as more matches are identified, the more stable the statistic becomes. Hence this particular phenomenon tends to disappear at the later part of the curves.

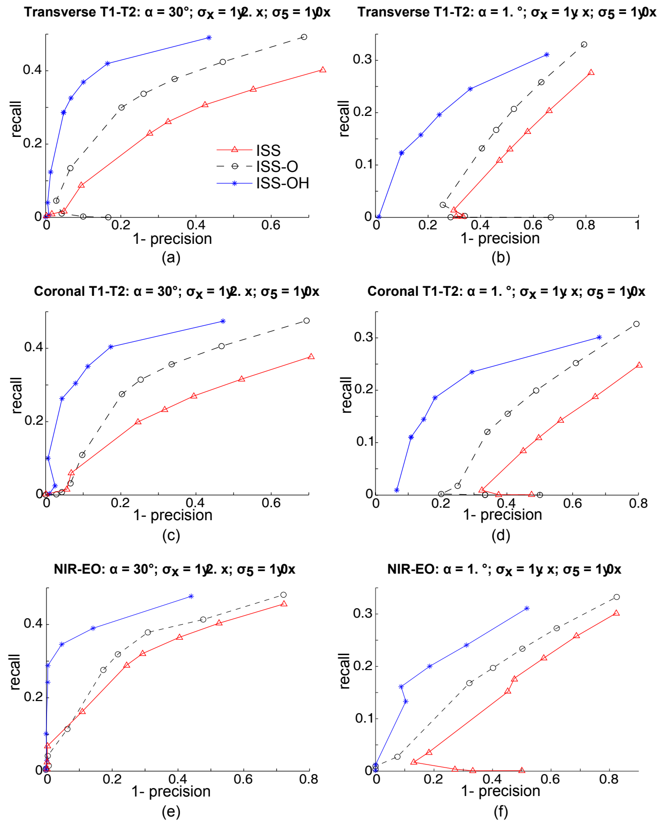

Fig. 6. Comparison between ISS, ISS-O and ISS-OH. *Recall* versus *1-Precision* curves for rotation and non-uniform scale transformations.

VI. Conclusion

In this paper we show the suitability and benefit of applying Hough Transform into ISS. The exclusion of the rotation and scale parameters from the transform remarkably reduces the size of the Hough Space. We also identify problems of using Hough Transform in its traditional way. The proposed technique simplifies the application of Hough Transform and defines mechanisms to avoid systematic errors caused by the minor and natural inaccuracies in DoG scale estimation. The experimental results that we present in this paper, prove the merit of the proposed technique and demonstrates higher precision and recall in key-point matching which is essential for improved registration accuracy.

References

[1] J. Pluim, J. Maintz, and M. Viergever, "Image registration by maximization of combined mutual information and gradient information," *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 103–129, 2000.

[2] P. Anuta, "Spatial registration of multispectral and multitemporal digital imagery using fast fourier transform techniques," *IEEE Transactions on Geoscience Electronics*, vol. 8, no. 4, pp. 353–368, 1970.

[3] J. Orchard, "Efficient least squares multimodal registration with a globally exhaustive alignment search," *IEEE Transactions on Image Processing*, vol. 16, no. 10, pp. 2526–2534, 2007.

[4] A. Roche, G. Malandain, X. Pennec, and N. Ayache, "The correlation ratio as a new similarity measure for multimodal image registration," *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, p. 1115, 1998.

[5] D. Russakoff, C. Tomasi, T. Rohlfing, and C. Jr, "Image similarity using mutual information of regions," *Computer Vision-ECCV 2004*, pp. 596–607, 2004.

[6] P. Viola and W. Wells III, "Alignment by maximization of mutual information," *International Journal of Computer Vision*, vol. 24, no. 2, pp. 137–154, 1997.

[7] A. Papoulis and R. Probability, *Stochastic processes*. McGraw-Hill, New York, 1991, vol. 3.

[8] J. Boes and C. Meyer, "Multi-variate mutual information for registration," *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 606–612, 1999.

[9] C. Meyer, J. Boes, B. Kim, P. Bland, K. Zasadny, P. Kison, K. Koral, K. Frey, and R. Wahl, "Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin-plate spline warped geometric deformations," *Medical Image Analysis*, vol. 1, no. 3, pp. 195–206, 1997.

[10] C. Meyer, J. Boes, B. Kim, and P. Bland, "Probabilistic brain atlas construction: Thin-plate spline warping via maximization of mutual information," *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 631–637, 1999.

[11] C. Meyer, J. Boes, B. Kim, P. Bland, G. Lecarpentier, J. Fowlkes, M. Roubidoux, and P. Carson, "Semiautomatic registration of volumetric ultrasound scans," *Ultrasound in Medicine & Biology*, vol. 25, no. 3, pp. 339–347, 1999.

[12] C. Meyer, B. Moffat, K. Kuszpit, P. Bland, T. Chenevert, A. Rehemtulla, and B. Ross, "A methodology for registration of a histological slide and in vivo mri volume based on optimizing mutual information," *Molecular Imaging: Official Journal of the Society for Molecular Imaging*, vol. 5, no. 1, p. 16, 2006.

[13] B. Kim, J. Boes, K. Frey, and C. Meyer, "Mutual information for automated multimodal image warping," *Visualization in Biomedical Computing*, pp. 349–354, 1996.

[14] M. Hossain, G. Lv, S. W. Teng, G. Lu, and M. Lackmann, "Improved symmetric-sift for multi-modal image registration," *International Conference on Digital Image Computing Techniques and Applications (DICTA)*, pp. 197–202, dec. 2011.

[15] 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, 2009.

[16] R. Ma, J. Chen, and Z. Su, "MI-SIFT: mirror and inversion invariant generalization for SIFT descriptor," *Proceedings of the ACM International Conference on Image and Video Retrieval*, pp. 228–235, 2010.

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

[18] D. Lowe, "Object recognition from local scale-invariant features," *ICCV*, p. 1150, 1999.

[19] D. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.

[20] P. Hough, "Method and means for recognizing complex patterns," Dec. 18 1962, uS Patent 3,069,654.

[21] W. Grimson, "Object recognition by computer. 1990."

[22] D. Ballard, "Generalizing the hough transform to detect arbitrary shapes," *Pattern Recognition*, vol. 13, no. 2, pp. 111–122, 1981.

[23] M. Hossain, S. W. Teng, G. Lu, and M. Lackmann, "An enhancement to sift-based techniques for image registration," *International Conference on Digital Image Computing Techniques and Applications (DICTA)*, pp. 166–171, dec. 2010.

[24] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1615–1630, 2005.

[25] "Brainweb: Simulated brain database," brainWeb. [Online]. Available: http://mouldy.bic.mni.mcgill.ca/brainweb/

[26] Y. S. Kim, J. H. Lee, and J. B. Ra, "Multi-sensor image registration based on intensity and edge orientation information," *Pattern Recognition*, vol. 41, no. 11, pp. 3356–3365, 2008.

[27] C. Fredembach and S. Susstrunk, "Illuminant estimation and detection using near-infrared," *Proceedings of SPIE*, vol. 7250, p. 72500E, 2009.