Hough Space Parametrization: Ensuring Global Consistency in Intensity-Based Registration

Mehmet Yigitsoy, Javad Fotouhi, and Nassir Navab

Computer Aided Medical Procedures (CAMP), TUM,
Munich, Germany
{yigitsoy,fotouhi,navab}@cs.tum.edu

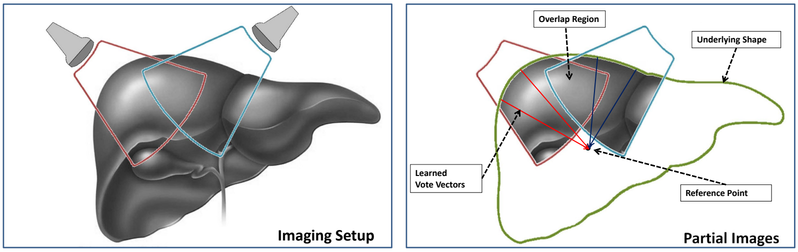
Abstract. Intensity based registration is a challenge when images to be registered have insufficient amount of information in their overlapping region. Especially, in the absence of dominant structures such as strong edges in this region, obtaining a solution that satisfies global structural consistency becomes difficult. In this work, we propose to exploit the vast amount of available information beyond the overlapping region to support the registration process. To this end, a novel global regularization term using Generalized Hough Transform is designed that ensures the global consistency when the local information in the overlap region is insufficient to drive the registration. Using prior data, we learn a parametrization of the target anatomy in Hough space. This parametrization is then used as a regularization for registering the observed partial images without using any prior data. Experiments on synthetic as well as on sample real medical images demonstrate the good performance and potential use of the proposed concept.

1 Introduction

Intensity based image registration is often challenging when images to be registered have insufficient amount of information in their overlapping region. Classical registration approaches are bounded to use only the overlapping region since they use intensity correspondences or statistical relationships. Especially, in the absence of relevant structures such as strong edges in this region, assessing the local image similarity alone becomes inadequate, leading to an ill-posed registration problem. It is often the case in ultrasound (US) imaging that the acoustic window of the transducer is limited; therefore, to capture a large field of view, several acquisitions are needed where partial images have to be then stitched together [13]. However, due to the typical imaging artifacts inherent to US, the information in the overlapping region is often not salient enough for an image based compounding.

Figs. 1(a)-1(b) show a case where two US images of the liver with little overlap need to be registered. Obviously, most intensity-based registration methods would fail here due to the limited information in the overlap to drive the registration process. On the other hand, there is a rich amount of information beyond the overlap which can support the registration.

P. Golland et al. (Eds.): MICCAI 2014, Part I, LNCS 8673, pp. 275–282, 2014.



(a) Imaging setup (b) Partial images

Fig. 1. (a) An illustration of the possible imaging setup where partial images may have limited overlap. (b) Given that the shape is known, dominant structures from non-overlapping regions can be utilized for regularization.

To alleviate this issue, acquisitions can be done such that there are large overlaps between partial images as done in [8] which obviously needs more time and effort. Similarly, methods based on matching extracted features [9] as well as patch-based voting schemes [7] require sufficient amount of overlap between images. Another typical approach is to augment the registration using previous data such as computed tomography (CT) scans [5] where each partial image is registered to the prior as well as to other partial images. For this approach to work, however, a previous scan of the same patient has to be readily available, which is often not the case. In the absence of previous scans, other prior data such as anatomical atlases are used for registration. Although it is possible to have good registration results using atlases, an explicit use of the intensity distribution of an atlas biases registration towards the chosen reference, making it inflexible and necessitating additional regularization. Recently, a method seeking for a consistent alignment of structures beyond overlapping region was proposed in [14]. The method, however, provides only local smoothness; thus, it cannot be generalized as a global regularization constraint.

All the previously mentioned approaches will have difficulties when the overlap size is small, non-salient or no prior data is given, therefore, they cannot ensure global consistency. In this work, we propose not to use prior data explicitly, but we learn a global parametrization $\mathcal{P}$ of the anatomy in question. To this end, we use Generalized Hough Transform (GHT) [1] to learn $\mathcal{P}_H$ of the target anatomy in the Hough space. GHT has the favorable property of being robust to partial occlusions and noise which we exploit here to enable global regularization for intensity based registration. Coupled with local similarity, $\mathcal{P}_H$ will serve as a regularization to ensure global consistency while registering partially overlapping images.

2 Methods

Generalized Hough Transform: Hough Transform was originally proposed for lines and edges, but later it was extended to other analytical shapes. Finally, GHT was introduced as an extension where any arbitrary shape can be used as the prior for object detection [1]. GHT parametrizes a shape with the offsets of each shape element from a reference point. This parametrization is stored in a look-up table, so-called R-Table. This table is later used for detection where elements in a target image vote for the hypotheses that might have generated them. The peaks in the Hough space created by accumulating such votes correspond to parameters (such as the chosen reference point) of possible target object.

There is an ever increasing number of applications of GHT and its voting scheme in computer vision including object classification [6], detection [2] and tracking [3] to name a few. However, it has fewer applications in intensity based image registration where there is a large potential of usage. In [12], GHT was used to estimate the initial pose of intraoperative images where possible poses of target object are learned a priori. [10] addresses the initialization of intensity based rigid registration by using standard Hough Transform on gradient fields. To the best of our knowledge, GHT has not been used for global regularization of intensity based registration.

Global Consistency Measure: A Hough space parametrization $\mathcal{P}_H$ is learned from prior data such as previous scans, statistical shape models or atlases. Following the traditional Hough-like voting framework, using $\mathcal{P}_H$, the voting elements[1] in partial images vote for the hypotheses in a Hough space. Finally, a global consistency measure (GCM) is inferred from the resulting distribution. Various types of voting elements can be used ranging from edge pixels to more complex features such as keypoints or image patches [2]. Here, for the proof of concept, only edge pixels are considered as voting elements which can be extracted via standard edge detection methods such as Canny.

Let $\mathbf{I} = \left\{I_i : \Omega_i \to \mathbb{R}, \Omega_i \subset \mathbb{N}^N\right\}_{i=1}^n$ be n partial images to be registered, N being the image dimension. Furthermore, let Ω_c be the common spatial domain of partial images and $\mathbf{T} = \left\{T_i : \mathbf{x}_c \mapsto \mathbf{x}_i, |\mathbf{x}_i, \mathbf{x}_c \in \mathbb{R}^N\right\}_{i=1}^n$ be the corresponding transformations parametrized by $\mathbf{p} = \{p_i\}_{i=1}^n$, which, when optimal, will bring partial images into spatial alignment. Finally, we define Hough images $\mathbf{H} = \left\{H_i(\mathbf{x}|I_i, p_i, \mathcal{P}_H) : \Omega_h \to \mathbb{R}, \Omega_h \subset \mathbb{N}^M\right\}_{i=1}^n$ where M is the dimensionality of Hough space.

In this work, the sum of pairwise distances between the maxima in Hough images is considered as the GCM. Using the same Hough space size for each partial image, the goal is to bring the strongest hypotheses into a cluster. We define GCM on the joint Hough space as

$$GCM(\mathbf{I}, \mathbf{p}, \mathcal{P}_H) = \frac{2}{n(n-1)} \sum_{i=1}^{n} \sum_{j=i+1}^{n} \left\| \arg\max_{\mathbf{h}_i} H_i(\mathbf{h}_i) - \arg\max_{\mathbf{h}_j} H_j(\mathbf{h}_j) \right\| \quad (1)$$

where $H_k(\mathbf{h}_k) = H_k(\mathbf{h}_k|I_k, p_k, \mathcal{P}_H)$, $\mathbf{h}$ is a hypothesis and $\arg\max$ locates the strongest hypothesis in the Hough image. The goal here is then to minimize the

[1] The elements do not have to be necessarily in the overlapping region which is one of the key idea behind this work.

GCM by clustering the peaks. This measure will be then coupled with a local intensity-based similarity term.

Registration: We define the local similarity measure (LSM) in terms of intensities as

$$LSM(\mathbf{I}, \mathbf{p}) = \frac{2}{n(n-1)} \sum_{k=1}^{|\Omega_c|} \sum_{i=1}^{n} \sum_{j=i+1}^{n} \xi(I_i(T_{p_i}^{-1}(\mathbf{x}_k)), I_j(T_{p_j}^{-1}(\mathbf{x}_k))) \tag{2}$$

with T_{p_i} being the parametrization of T_i by p_i. Finally, we pose the alignment of all partial images as an optimization problem such that optimal transformations $\mathbf{T}$ optimizes an energy $\mathcal{E}$. Optimal parameters $\mathbf{p}$ then can be estimated via the following equation

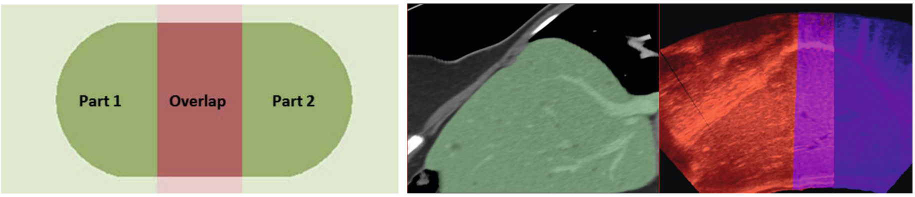
$$\hat{\mathbf{p}} = \arg\min_{\mathbf{p}} \mathcal{E}(\mathbf{p}|\mathbf{I}, \mathcal{P}_H) \quad \text{with} \quad \mathcal{E}(\mathbf{p}|\mathbf{I}, \mathcal{P}_H) = LSM(\mathbf{I}, \mathbf{p}) + e^{GCM(\mathbf{I}, \mathbf{p}, \mathcal{P}_H) - \rho}. \tag{3}$$

LSM serves as a data fidelity term evaluated in the overlapping region, whereas the exponential term is a global regularization evaluated in the Hough space. ρ is a constant controlling the amount of regularization based on the uncertainty in GCM. The choice of ρ depends on the resolution of the Hough space as well as on the image content. For large Hough images, there is more uncertainty regarding the maxima in the Hough space, therefore, inconsistent local alignments should be penalized less. Similarly, object deformations cause dispersion in Hough space, thus, leading to an increased uncertainty.

The data term $\xi(\cdot)$ can be chosen according to the modalities being registered. In this paper, we use Normalized Cross Correlation (NCC). We also assume that the optimal transformations $\mathbf{T}$ are only rigid, leading to 3D Hough images. For the optimization of Eq. 3, Nelder-Mead Simplex algorithm as part of the NLopt package [4] is used.

3 Experimental Validation and Results

We have conducted experiments on synthetic and real images. For all experiments, we used 500 iterations for the optimizer and set ρ to 10 and 50 for synthetic and real experiments respectively. Through the synthetic experiments, we analyzed the robustness of our approach to the size of overlap and to the varying degrees of imaging noise. For this purpose, a binary image with 400x190 pixels size containing a shape was used as a template to learn the object. Two partially overlapping images are extracted from the same shape as the observed images (cf. Fig. 2(a)). Note that if the shape information is not utilized, a registration method will be insensitive to the horizontal translations in the overlapping region to a certain extent. Therefore, this synthetic experiment will also show the need for regularization when there is no global optimum in terms of intensity correspondences. The goal in this experiment is to reconstruct the same geometry starting from a randomly chosen initial relative positioning of partial images within a specified range.



(a) Synthetic image (b) US-CT image pair

Fig. 2. (a) An arbitrary shape and sample selection of partial images for experiments. (b) Used US-CT pair with CT segmentation overlayed on CT (left) and registered partially overlapping US images (right) where the overlapping region is highlighted.

Starting from a full overlap of objects in the partial images, we varied the overlap size by increments of 10% of the image width till we get a -60% overlap, which is a gap of size 60% of the image width. For each overlap size, we added uniform noise by varying its maximum relative to the the image dynamic range. For each overlap size and noise level, we applied 20 combinations of initial rigid transformations with translations and rotations chosen from $[-200, 200]$ pixels and $[-50, 50]$ degrees respectively. For evaluation, we warped the original partial images before adding noise using the optimal transformations and compared to the corresponding part of the original full image using Dice score. For comparison, we have conducted the same experiment without using the GCM. Results using GCM shown in Figs. 3(a)-(b) compared to Figs. 3(c)-(d) without GCM indicate the robustness of the method to varying degrees of noise as well as its good performance even in the presence of a gap. It is clear that usage of GCM improves the capture range of cost function and avoids undesired local optima.

In order to demonstrate the performance on real images, we performed an experiment where we took a pair of slices, each having 512x384 resolution with a pixel spacing 0.45mm, from a co-registered US-CT pair corresponding to the liver area. Then, we used the segmented CT for learning the parametrization $\mathcal{P}_H$ and cut the US image into two partially overlapping sub-images as shown in Fig. 2(b) which were used to reconstruct the original US image. The size of the overlap was about 15% of the original image size. We applied random initial rigid transformations in a range around the optimum where 50 combinations of translations and rotations were randomly sampled from $[-100, 100]$ pixels relative to the optimum and $[-30, 30]$ degrees respectively. For the evaluation of each case, we warped the segmentations of partial images and compared with the segmentation of the uncut US image using Dice score. Mean, median and STD values with and without using GCM were recorded as 0.95, 0.99, 0.07 and 0.48, 0.45, 0.34, respectively. Results shown in Fig. 4(a) support our previous observations in terms of robustness. Moreover, we calculated the scores with respect to changing uncertainty parameter ρ in Eq. 3 by varying it from 0 to 100. Fig. 4(b) shows that large values of uncertainty lead to a degradation of performance which is expected due to the reduced amount of regularization. This is also valid for small values resulting in a very strict regularization, thus, making it sensitive to the errors in detecting strongest hypotheses.

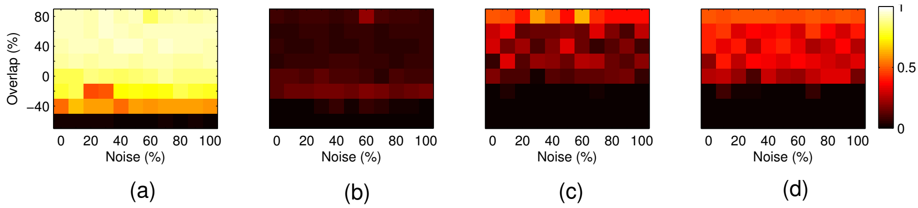


Fig. 3. Performance evaluation against noise and the amount of overlap using synthetic images. (a)-(b) Dice score and its standard deviation (STD) using GCM. (c)-(d) Dice score and STD without using GCM. A negative overlap value indicates a gap between images. It is clear from (a) that the method is robust to noise and can tolerate even gaps between the images. Whereas, without GCM, Dice score is very low with a high STD even when the overlap is sufficient and it is not possible to register with a gap.

For further evaluation in terms of Target Registration Error (TRE), we used 16 pairs of landmarks manually extracted from the overlapping region. The same experiments were repeated by using US and CT segmentations as well as their slightly deformed versions respectively for learning $\mathcal{P}_H$. As seen from Fig. 4(c), the best median TRE (8.46 pixels) was obtained by using CT for learning followed by using US (13.38 pixels). There were small degradations in each case when their deformed versions were used for learning, indicating the tolerance of the method to small deformations. A median error of 127.88 pixels was obtained when GCM was not employed. Obviously, most of the registrations without using GCM failed due to the small size of the overlap and the sensitivity to the initial parameters indicating a very limited capture range of the cost function. The slightly worse performance when using US compared to CT is due to the speckle and shadows in US images leading to false edges, thus, more uncertainty.

4 Discussion

The proposed concept differs from model based segmentation and registration methods in that we do not make any explicit use of prior data. We only learn a parametrization to employ it later for registering partial images which differentiates it also from the model-to-image alignment methods [11]. Therefore, once the parametrization is learned, the full intensity distribution of the prior image is not required for registration.

GCM term does not depend on the modality at hand as long as features required for the GHT can be extracted from the images. This makes the proposed framework suitable also for multi-modal applications. Theoretically, as prior data, it is possible to use 1) a different modality, 2) an image of the slightly deformed target anatomy, 3) a statistical shape model for learning the parametrization. This is one of the key features of the proposed concept allowing flexibility in model based reconstruction. Note that we are not making any comparison to the atlas based registration techniques. Here, we propose an alternative concept with its own advantages.

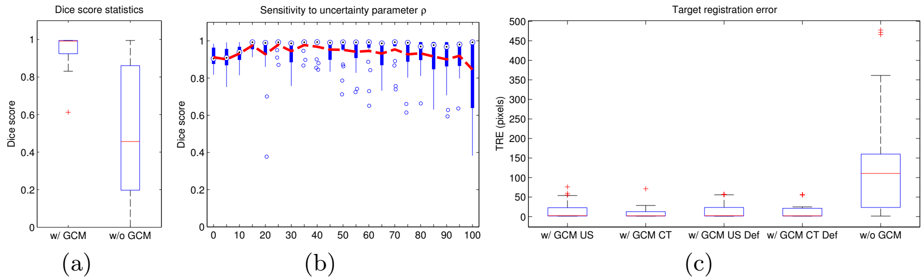


Fig. 4. Evaluation on US images. (a) Dice scores with and without using GCM. (b) Sensitivity analysis with respect to the changing uncertainty parameter (x-axis) in Eq. 3. Dashed line represents the mean value. (c) TREs when US(w/ GCM US), CT(w/ GCM CT), deformed US (w/ GCM US Def) and CT (w/ GCM CT Def) are used for learning. The last one is when GCM is not used at all.

We did not employ any sophisticated techniques for finding modes in the Hough space, emphasizing the simplicity of the proposed concept. Nevertheless, it is possible to augment the proposed Hough space parametrization approach by using advanced voting and mode finding techniques such as the ones used in computer vision applications. Here, a minimal implementation of the concept and its key features are presented through a proof-of-concept study.

The classical GHT is invariant to rotations and isotropic scaling of the object and tolerates small deformations which was confirmed by our experiments using slightly deformed segmentations for learning. However, it still cannot handle large deformations. This limitation, however, can be relaxed by employing large training datasets in the learning phase in a decision trees framework [3]. As future work, we will extend the proposed method by posing it as a regression problem and by solving it in a random forest framework. This will allow us to employ higher order features as voting elements [2]. Finally, using higher order transformations such as affine and deformable as well as mosaicing in 3D will be the immediate extensions of the proposed concept.

5 Conclusion

In this work, we have proposed a novel global regularization term using Generalized Hough Transform (GHT). It is used as a Global Consistency Measure (GCM) in intensity-based registration when the local information in the overlap of partial images is insufficient or corrupted. A Hough space parametrization of the target anatomy is learned from prior data such as previous scan, atlas or statistical shape model. This parametrization enables a global consistency voting through local information. The proposed concept is fully parallelizable and suited for reconstructing sparsely sampled scenes. Through experiments on synthetic and real images, it was shown that using GCM improves the registration quality when the partial images have less in common in their overlap.

References

1. Ballard, D.H.: Generalizing the hough transform to detect arbitrary shapes. Pattern Recognition 13(2), 111–122 (1981)
2. Gall, J., Lempitsky, V.: Class-specific hough forests for object detection. In: Decision Forests for Computer Vision and Medical Image Analysis, pp. 143–157. Springer (2013)
3. Godec, M., Roth, P.M., Bischof, H.: Hough-based tracking of non-rigid objects. Computer Vision and Image Understanding 117(10), 1245–1256 (2013)
4. Johnson, S.G.: The nlopt nonlinear-optimization package, `http://ab-initio.mit.edu/wiki/index.php/NLopt` (accessed February 21, 2014)
5. Kutter, O., Wein, W., Navab, N.: Multi-modal registration based ultrasound mosaicing. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) MICCAI 2009, Part I. LNCS, vol. 5761, pp. 763–770. Springer, Heidelberg (2009)
6. Leibe, B., Leonardis, A., Schiele, B.: Combined object categorization and segmentation with an implicit shape model. In: Workshop on Statistical Learning in Computer Vision (ECCV) (May 2004)
7. Ourselin, S., Roche, A., Prima, S., Ayache, N.: Block matching: A general framework to improve robustness of rigid registration of medical images. In: Delp, S.L., DiGoia, A.M., Jaramaz, B. (eds.) MICCAI 2000. LNCS, vol. 1935, pp. 557–566. Springer, Heidelberg (2000)
8. Øye, O., Wein, W., Ulvang, D., Matre, K., Viola, I.: Real time image-based tracking of 4d ultrasound data. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012, Part I. LNCS, vol. 7510, pp. 447–454. Springer, Heidelberg (2012)
9. Schneider, R.J., Perrin, D.P., Vasilyev, N.V., Marx, G.R., del Nido, P.J., Howe, R.D.: Real-time image-based rigid registration of three-dimensional ultrasound. Medical Image Analysis 16(2), 402–414 (2012)
10. Shams, R., Barnes, N., Hartley, R.: Image registration in hough space using gradient of images. In: 9th Biennial Conference of the Australian Pattern Recognition Society on Digital Image Computing Techniques and Applications, pp. 226–232. IEEE (2007)
11. Toews, M., Wells III, W.M.: Efficient and robust model-to-image alignment using 3d scale-invariant features. Medical Image Analysis 17(3), 271–282 (2013)
12. Varnavas, A., Carrell, T., Penney, G.: Fully automated initialisation of 2D-3D image registration. In: 2013 IEEE 10th International Symposium on Biomedical Imaging (ISBI), pp. 568–571. IEEE (2013)
13. 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)
14. Yigitsoy, M., Navab, N.: Structure propagation for image registration. IEEE Transactions on Medical Imaging 32(9), 1657–1670 (2013)