Image-Based Registration in Canonical Atlas Space

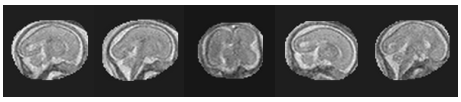
Benjamin Hou[1], **Bishesh Khanal**[1,2], **Amir Alansary**[1], **Steven McDonagh**[1], **Alice Davidson**[2], **Mary Rutherford**[2], **Jo V. Hajnal**[2], **Daniel Rueckert**[1], **Ben Glocker**[1], **Bernhard Kainz**[1]
[1]Imperial College London, [2]Kings College London
bh1511@imperial.ac.uk

Abstract

Limited capture range, and the requirement to provide high quality initialization for optimization-based 2D/3D image registration methods, can significantly degrade the performance of 3D image reconstruction and motion compensation pipelines. Challenging clinical imaging scenarios, which contain significant subject motion such as fetal in-utero imaging, complicate the 3D image and volume reconstruction process. In this paper we present a learning based image registration method using Convolutional Neural Networks (CNNs) to predicting 3D rigid transformations of arbitrarily oriented 2D image slices, with respect to a learned canonical atlas co-ordinate system. Only image slice intensity information is used to perform registration and canonical alignment. We extensively evaluate the effectiveness of our approach quantitatively on simulated Magnetic Resonance Imaging (MRI), fetal brain imagery with synthetic motion and further demonstrate qualitative results on real fetal MRI data where our method is integrated into a full reconstruction and motion compensation pipeline. Furthermore, we utilise Monte Carlo Dropout for the purpose of establishing a prediction confidence metric.

1 Introduction

Reconstructing a 3D volume from misaligned and motion corrupted 2D images is a challenging task. The process involves labor intensive pre-processing steps, such as manual landmark matching, exhibiting both inter and intra-observer variance. Traditional automatic intensity-based slice-to-volume reconstruction methods [2] involve solving the inverse problem of super-resolution from slice acquisitions, $y_i = D_i B_i S_i M_i x + n_i; \quad i = 1, 2, \ldots, N$. However, arbitrary subject motion (e.g. Fig 1) can invalidate slice alignment assumptions that are based on the scanner co-ordinate system, and manual intervention may be necessary. Manual correction of slice-to-volume registration often becomes unfeasible in practice due to the magnitude of image data involved.



(a) #12 (b) #13 (c) #14 (d) #15 (e) #16

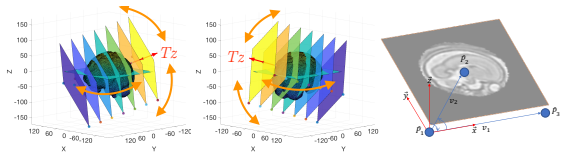
Figure 1: Sequential scan slices from a sagittal image stack of a fetus with extreme motion, it can be observed that the fetus has rotated its head 90°, causing slice #14 to be a coronal view.

The optimization methods employed in this domain typically do not guarantee a globally optimal registration solution from arbitrarily seeded slice alignment. The function that maps each 2D slice to its correct anatomical position in 3D space may be subject to local minima and the requirement for small initial misalignment typically improves result quality. Previous work have attempted to make this optimization robust by introducing appropriate objective functions and outlier rejection strategies based on robust statistics [2, 5]. Despite these efforts, good reconstruction quality still depends on having good initial alignment of adjacent and intersecting slices.

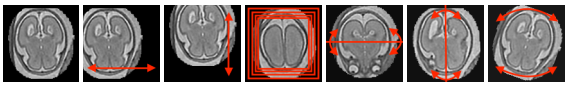
1st Conference on Medical Imaging with Deep Learning (MIDL 2018), Amsterdam, The Netherlands.

2 Method

As shown in [3], to fully evaluate and assess the performance of 2D/3D registration via a learning based approach, we incorporated it into a full 3D reconstruction pipeline featuring three modular components; (*1*) Approximate Organ Localization, (*2*) Canonical Slice Orientation Estimation, and (*3*) Intensity-Based 3D Volume Reconstruction. Organ localization (*1*) is concerned with localization of a learned RoI. This can be achieved through automatic methods or rough manual segmentation. For 3D Volume Reconstruction (*3*) we use a modified iterative SVR method [4], which additionally allows for compensation of any remaining small misalignments between single slices, caused by prediction inaccuracies. Here, we focus on (*2*) Canonical Slice Orientation Estimation, denoted as *SVRnet* [3].



(a) Front View (b) Rear View (c) Anchor Points



(d) Orig (e) T_x (f) T_y (g) T_z (h) R_x (i) R_y (j) R_z

Figure 2: Slice Generation

The motion of a rigid body in 3D has six Degrees of Freedom (6 DoF), three parameters for translation (T_x, T_y, T_z) and three for rotation (R_x, R_y, R_z). To model the movement of each slice in 3D space, we divide the parameters into two categories; in-plane transformation T_x, T_y and R_z and out-of-plane transformation T_z, R_x and R_y (see Fig. 2(d-j)). If each DoF is allowed ten interval delineation, this would result in 10^6 slices per organ volume. Automatic segmentation methods define the RoI on a slice by slice basis throughout the 3D volume. The desired RoI (*e.g.*, segmented brain) is masked and center aligned within each slice. This vastly decreases the valid range for in-plane motion parameters T_x and T_y. Similar to [6], we reduce the number of slices required to create training and validation data sets by simplifying the sample space, such that it is constrained by the parameters: T_z, R_x, R_y and R_z. We can further discount a portion of slices that yield little or no content at the extremities of the T_z range, in the considered volume. We generate the training dataset by sampling 2D slices within a 3D reconstructed brain volume in regular intervals throughout each DoF. For validation, random sampling is performed to mimic the natural continuous movement of the fetus.

We also introduce a novel labelling system where the rotation and translation components of the labels are combined together. Any three non co-linear points in a 3D Euclidean space form a plane, while their order defines the orientation. We therefore call them *Anchor Points* (see Fig. 2(a-c)). Three Anchor Points can be defined anywhere on the 2D slice, as long as they are not identical or co-linear and the relative in-plane locations are consistent throughout all slices in the data set. For simplicity, we defined the three Anchor Points, which represent the pose of each slice, to be located at the center and the bottom left and right corners. The loss function of the network is therefore:

$$\textbf{Loss} = \alpha \left\| \hat{P}_1 - P_1 \right\|_2 + \beta \left\| \hat{P}_2 - P_2 \right\|_2 + \gamma \left\| \hat{P}_3 - P_3 \right\|_2 \tag{1}$$

Gal et al. [1] recently showed that dropout layers in Neural Networks can be interpreted as a Bayesian approximation to probabilistic models. We leverage this technique to take epistemic uncertainty into consideration, in order to gauge alignment prediction confidence in real-world test cases, as quality of 3D reconstruction depends on precision of slice alignment. Each slice is passed through the network N times, where the mean of all N predictions is the final pose and the variance of all N predictions is the uncertainty. This is calculated on the $SE(3)$ manifold using [7].

3 Experiments and Results

We test SVRnet on a case, which our clinical partners dismissed as impossible to reconstruct. Both, extensive manual and automatic reconstruction attempts have failed for this case. With no ground truth to compare to, reconstruction quality can only be validated qualitatively. Fig. 3a, b and c show the raw scan stacks, and the degree of motion corruption. In a case like this, excessive motion can cause ambiguity. Fig. 1 shows a sequential sagittal stack of slices where the fetus has turned its head almost $90°$, causing a coronal slice to be in a scan stack that is assumed sagittal. This unexpected slice

does not fit in the stack, and is normally rejected by robust statistics implemented in SVR algorithms. Rejecting too many slices will cause a lack of scan data, while accepting too many slices will cause a corrupt reconstruction volume as seen in Fig. 3d. Fig. 3d shows a SVR-based reconstruction attempt, using [4], without SVRnet initialization.

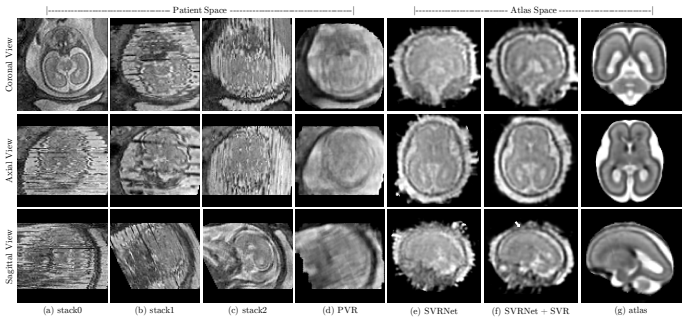
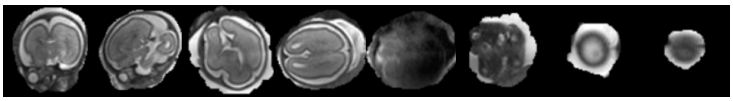


Figure 3: Reconstruction attempt of a heavily motion-corrupted fetal brain scan at approx. 20 weeks GA.

Fig. 3e shows Gaussian average of all slices that are predicted and realigned to atlas space by SVRnet. This was further refined with four iterations of Slice to Volume Registration (SVR) [4], with slice location initialized by SVRnet in Fig. 3e. It can be seen that reconstructed volume is now in canonical atlas space, similar to the training atlas Fig. 3f.

The decision of whether or not to include a slice in subsequent reconstruction depends on the prediction confidence and the robustness of the chosen reconstruction algorithm for (*3*). Prediction confidence can be thresholded and if the reconstruction algorithm is very robust, like [4], we can make multiple predictions per slice and let the reconstruction algorithm handle further outlier rejection, which allows for a greater margin of error (see Fig. 3). Fig 4 shows prediction confidence of various test slices. As SVRnet is trained on the middle 60% of the fetal brain volume, slices closer to the edge (Fig 4f-h) have a high variance (high uncertainty). Fig 4e shows a corrupt slice due to signal dropout, the variance of this slice is also high. Slices with a high variance should be marked for rejection.



(a) 1.92 (b) 2.46 (c) 3.82 (d) 9.66 (e) 29.31 (f) 30.92 (g) 36.98 (h) 43.34

Figure 4: Monte Carlo predictions of a unitless Geodesic distance variance metric for each slice. (a)-(d) represent confident predictions. (e)-(h) represent less confident predictions which are discarded for subsequent volume reconstruction.

Here, we have shown that SVRnet is able to predict slice transformations relative to a canonical atlas coordinate system, using only the intensity information in the image. This allows motion compensation for highly motion corrupted scans, *e.g.*, MRI scans of very young fetuses. It allows the incorporation of any images that have been acquired during examination, thus relaxing the requirement for temporal scan-plane proximity and widening the capture range. Our work also leverages the computational framework to do statistics on $SE(3)$ Lie groups, performing Bayesian Inference and Monte Carlo dropout sampling for outlier rejection.

References

[1] Gal, Y., et al.: Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In: Proc. ICML. Proc. Machine Learning Research, vol. 48, pp. 1050–1059. PMLR (20–22 Jun 2016)

[2] Gholipour, A., et al.: Robust super-resolution volume reconstruction from slice acquisitions: application to fetal brain MRI. IEEE Trans. Med. Imag. 29(10), 1739–1758 (2010)

[3] Hou, B., et al.: 3D Reconstruction in Canonical Co-ordinate Space from Arbitrarily Oriented 2D Images. IEEE Trans. Med. Imag. PP(in print, doi.: 10.1109/TMI.2018.2798801), 1–13 (2018)

[4] Kainz, B., et al.: Fast Volume Reconstruction from Motion Corrupted Stacks of 2D Slices. IEEE Trans. Med. Imag. 34(9), 1901–13 (2015)

[5] Kuklisova-Murgasova, M., et al.: Reconstruction of Fetal Brain MRI with Intensity Matching and Complete Outlier Removal. Medical Image Analysis 16(8), 1550–60 (2012)

[6] Miao, S., et al.: A CNN Regression Approach for Real-Time 2D/3D Registration. IEEE Trans. Med. Imag. 35(5), 1352–1363 (2016)

[7] Miolane, N.: Geomstats: Computations and statistics on manifolds with geometric structures. (Feb 2018), `https://github.com/ninamiolane/geomstats`