2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

An Approach for Live Motion Correction for TRUS-MR Prostate Fusion Biopsy using Deep Learning

Aditya Bhardwaj*, Praful Mathur*, Tejal Singh*, Venkata Suryanarayana, Yuri Son, Srinivas Rao Kudavelly, Sangha Song and Hokyung Kang

Abstract— **TRUS-MR fusion guided biopsy highly depends on the quality of alignment between pre-operative Magnetic Resonance (MR) image and live trans-rectal ultrasound (TRUS) image during biopsy. Lot of factors influence the alignment of prostate during the biopsy like rigid motion due to patient movement and deformation of the prostate due to probe pressure. For MR-TRUS alignment during live procedure, the efficiency of the algorithm and accuracy plays an important role. In this paper, we have designed a comprehensive framework for fusion based biopsy using an end-to-end deep learning network for performing both rigid and deformation correction. Both rigid and deformation correction in one single network helps in reducing the computation time required for live TRUS-MR alignment. We have used 6500 images from 34 subjects for conducting this study. Our proposed registration pipeline provides Target Registration Error (TRE) of 2.51 mm after rigid and deformation correction on unseen patient dataset. In addition, with a total computation time of 70ms, we are able to achieve a rendering rate of 14 frames per second (FPS) that makes our network well suited for live procedures.**

Clinical Relevance— **It is shown in the literature that systematic biopsy is the standard method for biopsy sampling in prostate that has high false negative rates. TRUS-MR fusion guided biopsy reduces the false negative rate of the sampling in prostate biopsy. Therefore, a live TRUS-MR fusion framework is helpful for prostate biopsy clinical procedures.**

I. Introduction

Prostate cancer is the most common form of cancer among global male population [1]. The standard diagnostic procedure for detecting and grading of prostate cancer involve ultrasound guided biopsy. The biopsy samples are taken systematically by dividing the prostate into equal segments and taking tissue samples close to uniform sampling across the prostate volume. The main drawback of the systematic biopsy is the high false-negative rate. Localization of the region of interest for biopsy is useful in reducing the false-negative rate of biopsy sampling, thereby eliminating the chances of repeated biopsy.

To reduce the false-negative rate of the systematic biopsy, fusion biopsy is becoming very popular now a days [2]. In fusion biopsy, pre-operative MR study is performed on the subject and potential lesion regions which are primarily the region of interest (ROI) are marked on the MR images. During the clinical procedure for the fusion biopsy, the ROI from pre-operative MR t2w scan is mapped on the intra-operative ultrasound images to guide the clinician for potential areas of biopsy. To map the pre-operative MR scan onto the live ultrasound during the procedure, a base level rigid transform is computed between TRUS and MR using the electromagnetic (EM) tracker present on the system. This initial registration is bound to change over time due to various factors like patient motion and deformation of the prostate due to the probe pressure. Due to this motion, the alignment between MR and TRUS has to be updated during the biopsy.

To overcome misalignment, live correction is needed, which include rigid motion and deformation compensation. Tharindu De Silva et. al.[3] describes a 2D/3D TRUS image based registration using normalize cross correlation metric and powel optimizer for rigid correction. They achieve an accuracy of 3.18 ± 1.6 mm on clinical data, simulated with various probe deformations, with run-time of 1.1s. Sheng Xu et. al.[4] describes an approach to perform real-time 2D/3D TRUS image-to-image registration using intermediate 2.5D to 3D TRUS with accuracy of 2.4 ± 1.2 mm on phantom data. Shihui Zhang et. al.[5] describes a technique to perform a 2D/3D image-to-image multi-modal registration, without requiring a 3D TRUS by directly registering 2D TRUS with 3D MR using mutual information metric on feature image (gradient). This technique achieve an accuracy of 2.52 ± 0.46 mm on clinical data but is not real-time (~35 seconds) and does not support deformation correction. In one of our previous works, we have proposed a workflow for fast rigid (~120ms) and deformation (+40ms) compensation [6]. Even at 8 FPS using frame-interleaving implementation, it does not fulfill the need of live motion correction for the doctor to see registered MR and TRUS images during live biopsy. In this work, we have designed a novel comprehensive framework for fusion based biopsy using an end-to-end deep learning network for performing both rigid and deformation correction with minimal latency (70ms / 14 FPS).

II. Methodology

In this section, we present a registration framework to estimate rigid translation transform for TRUS and MR $(T_{US}^{ct}, T_{MR}^{ct})$, rotation $\overrightarrow{(Rot_{MR})}$ and deformation $\overrightarrow{(D_{MR})}$ displacement fields for multi-modal registration of 2D TRUS and 3D MR. We have used our proposed network for live registration using feature based matching. Fig. 1 presents end-to-end architecture of proposed deep learning (DL) based registration pipeline, which includes feature generation (a),

* Authors with equal contribution.
Aditya Bhardwaj, Praful Mathur, Tejal Singh, Venkata Suryanaraya, Srinivas Rao Kudavelly are with Samsung R&D Institute, Bangalore, India (Phone: +91-80-3341-3000; Fax: +91-80-3341-3001; E-mail: s.kudavelly@samsung.com)

Yuri Son, Sangha Song, Hokung Kang are with Samsung Electronics, Seoul, South Korea

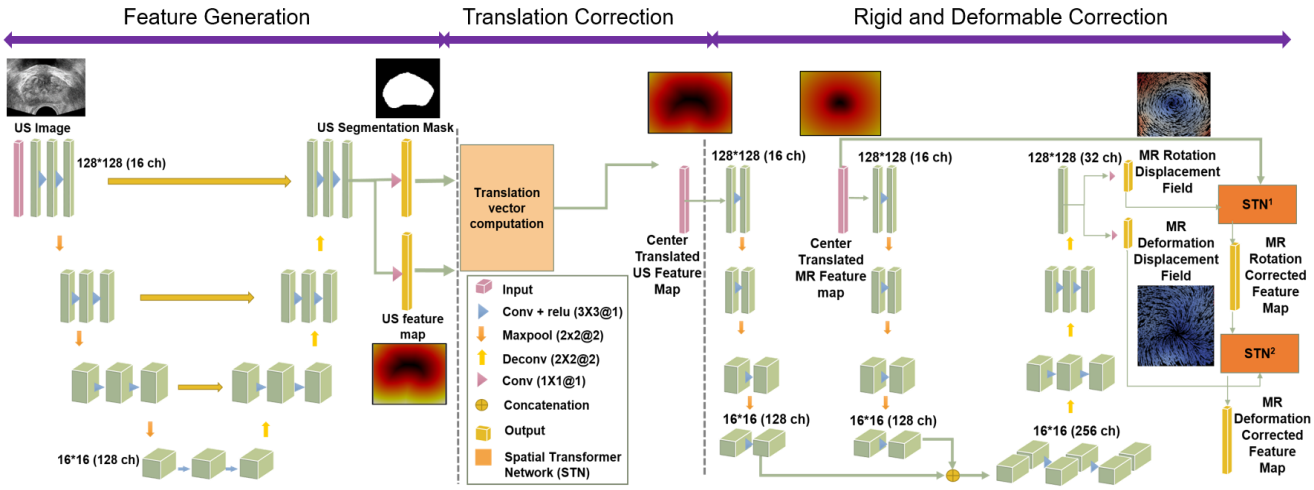

Figure 1. Complete registration pipeline (a) Feature Generation (b) Translation Correction (c) Rotation and Deformation Correction

translation correction (b) and rotation and deformation correction (c).

A. *Feature generation*

Feature based image-to-image registration is proven to be robust for multi-modal image registration [7]. In this work, we have used signed Euclidean distance map obtained from segmentation mask of prostate as feature map. For pre-operative 3D MR volume, we trained a 3D U-Net based segmentation network using segmentation mask of prostate marked by an expert radiologist as ground truth. The precision and correctness of 3D MR segmentation masks obtained as network's output is of utmost importance, as it will directly affect the quality of registration. We have used Signed Maurer algorithm in ITK to obtain the 3D MR feature map (F_{MR}^{3D}). Then EM transform is used to re-slice F_{MR}^{3D} to obtain a 2D MR feature map (F_{MR}^{2D}) in TRUS space.

For intra-operative 2D TRUS image, we have modified our encoder-decoder network proposed in [6] to compute the segmentation of prostate and 2D TRUS feature map (F_{US}^{2D}) simultaneously as shown in Fig. 1(a). The network is trained in a supervised manner using TRUS segmentation masks marked by clinician and feature maps obtained using Signed Maurer algorithm in ITK as ground truth. DL regression of the distance map reduces the feature computation time compared to [6] and is essential for live TRUS-MR feature based registration. For training the feature generator network, the losses used are binary cross entropy on segmentation predictions and mean squared error (MSE) on distance map predictions with equal weightage.

B. *TRUS-MR Translation Correction*

In order to compensate for translation, we translate the MR and trans-rectal ultrasound (TRUS) feature maps such that the centroid of prostate aligns with the image center. For computing the translation vector (T_{US}^{ct}) for TRUS feature map and (T_{MR}^{ct}) for MR feature map we first compute the centroid of prostate using segmentation mask and then find its offset with image center. Center translated 2D MR feature map (F_{MR}^{ct}) is obtained in prior on CPU and is used as input for component (Fig. 1(c)) of the proposed network. TRUS feature map is center translated in the network as shown in Fig. 1(b) to obtain center translated TRUS feature map (F_{US}^{ct}) using a Spatial Transformer Network (STN) [8].

C. *Rotation and Deformation Fields Regression*

We have designed a two-encoder one-decoder network for obtaining rotation ($\overrightarrow{\text{Rot}_{\text{MR}}}$) and deformation ($\overrightarrow{\text{D}_{\text{MR}}}$) displacement field vectors. The network architecture is shown in Fig. 1(c). It takes center translated MR feature map (F_{MR}^{ct}) and center translated TRUS feature map (F_{US}^{ct}) as inputs to generate ($\overrightarrow{Rot_{MR}}$) and ($\overrightarrow{D_{MR}}$) as outputs.

In order to increase the efficacy of training and faster convergence, the training network includes two STN layers for loss computation as shown in Fig. 1(c). STN[1] takes as input rotation displacement field ($\overrightarrow{Rot_{MR}}$) as produced by decoder and MR feature map (F_{MR}^{ct}) to produce rigid corrected MR feature map (F_{MR}^{ct+rot}). STN[2] takes as input deformation displacement field ($\overrightarrow{D_{MR}}$) and STN[1]'s output (F_{MR}^{ct+rot}) to produce rigid and deformation corrected MR feature map ($F_{MR}^{ct+rot+deform}$). For training the rotation and deformation regression network, the losses used are defined as follows:

1. Rotation Displacement Field Regression Loss

We have defined a custom loss which is a weighted combination of two losses as given in (1). First component is the supervised MSE loss on rotation displacement field vector. Second component is L2 norm of gradients of displacement field vector as shown in (2) that is added as a regularizer to avoid large and arbitrary displacements.

$$\text{Loss}=1*\text{MSE}+5*\text{L2} \tag{1}$$

$$\text{L2}=\left(\frac{\partial(\overrightarrow{\text{Rot}_{\text{MR}}})}{\partial x}\right)^2+\left(\frac{\partial(\overrightarrow{\text{Rot}_{\text{MR}}})}{\partial y}\right)^2 \tag{2}$$

2. Deformation Displacement Field Regression Loss

We have defined a custom loss which is a weighted combination of two losses as given in (3). First component is L2 norm of gradients of displacement field as given in (4). Second is the bending energy loss as given in (5) to ensure smoothness of deformation displacement field.

$$Loss = 5 * L2 + 5 * BE \quad (3)$$

$$L2 = \left(\frac{\partial(\overrightarrow{D_{MR}})}{\partial x}\right)^2 + \left(\frac{\partial(\overrightarrow{D_{MR}})}{\partial y}\right)^2 \quad (4)$$

$$BE = \left(\frac{\partial(D_{MR}^{x})}{\partial y}\right) + \left(\frac{\partial(D_{MR}^{y})}{\partial x}\right) \quad (5)$$

3. STN[1] and STN[2] loss

MSE loss is computed between the STN[1]'s output i.e. rotation corrected MR Feature map (F_{MR}^{ct+rot}) and its ground truth obtained by iterative ITK registration algorithm. For STN[2], we compute the MSE loss between STN[2]'s output i.e. (rigid + deformation) corrected MR feature map ($F_{MR}^{ct+rot+deform}$) and center translated TRUS Feature map(F_{US}^{ct}). As deformation displacement field regression is unsupervised, STN[2]'s loss plays a vital role in ensuring proper deformations.

III. Experimental Setup

A. Dataset and Preprocessing

Experiments were conducted on a dataset collected from 34 patient biopsies [6]. The data was collected in collaboration with Narayana Healthcare in accordance with the guidelines established by their Institutional Review Board (IRB). During pre-operative image acquisition process, t2w 3D MR volumes were acquired from multi-vendor (GE, Philips) systems with varying imaging parameters (like in-plane resolution - 0.351 mm to 0.664 mm, slice thickness - 0.700 mm to 7.000 mm). During intraoperative image acquisition process, 2D TRUS image (size: 640 x 704, spacing: 0.162 mm isotropic) and corresponding baseline alignment (EM transform) at 10 samples per second were recorded for entire duration of biopsy. For training the proposed registration pipeline, 4400 TRUS images from 23 patient biopsy data were used. Test set comprises of 2100 TRUS images from remaining 11 patients with annotations for TRE. The dataset is resampled to 128x128 dimension before passing into the network.

TABLE I. HYPER-PARAMETERS

Hyper-parameters	**Values**
Weights initialization	Xavier
Learning rate	1×10-4
No. of epochs	6500
Optimizer	Adam
Batch size	128
Activation Function	ReLU (hidden layers) Sigmoid (TRUS segmentation out) Linear (TRUS feature out) Linear (Rotation field out) Linear (Deformation field out)
Dropout	0.5

In order to increase the robustness of network and imitate high rotation angles between TRUS and MR during practical scenarios, we augmented rigid aligned MR (obtained via [6]) distance map by randomly rotating in range of -30 to +30 degrees.

B. Implementation Details

The network experiments were performed using tensorflow framework. Initially, TRUS feature generation network and rotation + deformation regression networks were trained separately for 3000 epochs. Then both of these networks were integrated and trained further until convergence. The hyper-parameters used during training of integrated pipeline are shown in Table 1.

During inference, the network pipeline takes live TRUS image and center translated MR feature map as input, and produces T_{US}^{ct} and composed (rotation + deformation) displacement field as output. The composed displacement field $(\overrightarrow{C_{MR}})$ is computed by transforming $(\overrightarrow{Rot_{MR}})$ using $(\overrightarrow{D_{MR}})$ and adding $(\overrightarrow{D_{MR}})$ as per (6).

$$(\overrightarrow{C_{MR}}) = (\overrightarrow{D_{MR}})\,[(\overrightarrow{Rot_{MR}})] + (\overrightarrow{D_{MR}}) \quad (6)$$

Fig. 2 shows$(\overrightarrow{Rot_{MR}})$, $(\overrightarrow{D_{MR}})$ and $(\overrightarrow{C_{MR}})$ respectively. The 3D MR volume is resliced using EM transform to obtain 2D MR image (I_{MR}^{2D}). I_{MR}^{2D} is then center translated using prior computed T_{MR}^{ct} and rotation + deformation corrected using $\overrightarrow{C_{MR}}$ to obtain $I_{MR}^{ct+rot+deform}$. Then $I_{MR}^{ct+rot+deform}$ is translated back using inverse of T_{US}^{ct}. In this way, TRUS-MR image registration is performed for each live TRUS image. With the proposed framework, we were able to achieve rigid and deformation registration of TRUS and MR in 70ms on Intel i7 2.6 GHz CPU and NVIDIA GeForce 920MX 2 GB GPU.

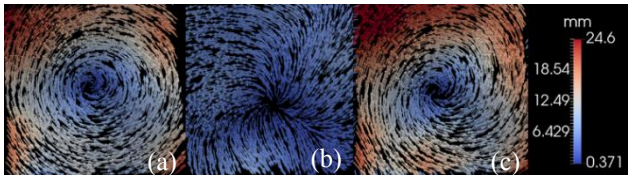

Figure 2. (a) Rotation Displacement Field (b) Deformation Displacement Field (c) Composed Displacement Field

IV. Results

For evaluating and quantifying the performance of proposed registration framework, we measured Dice score and Target Registration Error (TRE) metrics. Dice score is computed as per (7) between TRUS prostate segmentations (X) with corresponding MR prostate segmentations (Y). Dice score before applying the correction is compared with Dice post rigid and deformation correction. TRE is computed in 3D MR space as in (7), where P_{MR} is anatomical landmark on original 3D MR and $\widetilde{P_{MR}}$ is corresponding point mapped from live TRUS (P_{US}) onto 3D MR. A total of 110 anatomical landmark pairs were marked by expert radiologist in 2D TRUS images and 3D MR from 11 patient's data. These landmark pairs were marked in visible locations such as the corpora amylacea, urethra, cysts, ejaculatory ducts, Benign Prostatic Hyperplasia (BPH) nodules, calcifications regions on both TRUS and MR. TRE values are compared before and after rigid and deformation correction.

$$\text{Dice} = 2*\frac{|X \cap Y|}{|X|+|Y|}\ ,\ \text{TRE} = \left\|P_{MR}\text{-}\widetilde{P_{MR}}\right\| \quad (7)$$

The comparison of Dice score prior and post correction are shown in Table 2. Compared to unregistered TRUS-MR prostate Dice, rigid registration gave a significant improvement in Dice score from 0.73 to 0.85. Applying composed (rigid and deformation) displacement field on unregistered MR further improved the Dice to 0.93. To evaluate the robustness of our network towards high degree of rotations, we have artificially rotated MR in range (-30 to +30) degree and compared the Dice score between MR and TRUS post rigid and deformation correction.

TABLE II. DICE COMPARISON

Unregistered	**Rigid**	**Rigid + Deformation**	**Rigid + Deformation (augmented)**
0.73 ± 0.142	0.85 ± 0.12	**0.93 ± 0.18**	0.91 ± 0.23

The examples of rigid and deformation correction as applied on patient data using our network are shown in Fig. 3. Red contour represents the prostate boundary of uncorrected MR and green contour represents the prostate boundary of MR after rigid + deformation correction. Observe how corrected contour (green) matches the outline of prostate on TRUS compare to uncorrected (red). The comparison of TRE scores and run-time of other approaches is presented in Table III.

TABLE III. TRE COMPARISON WITH OTHER APPROACHES

Method	Clinical/ Phantom	Baseline TRE (mm)	Rigid TRE (mm)	Rigid + Deformation TRE (mm)	Run-Time (ms)
Ours	Clinical	8.73 ±5.1	2.79 ± 1.6	2.51 ±1.2	70
Aditya et al. [6]	Clinical	8.89 ±5.106	3.36 ±1.8	2.98 ±1.51	112 (rigid)
De Silva et al. [3]	Clinical Simulated	6.89 ±4.1	3.18 ±1.6	N/A	1100
Shihui et al.[5]	Phantom	N/A	2.4 ±1.2	N/A	350
Sheng xu et al. [4]	Clinical	N/A	2.52 ±0.46	N/A	3500

V. CONCLUSION

In this paper, we have proposed a deep learning based registration framework for both rigid and deformation error correction during live fusion biopsy. Conventional non DL based registration approaches are time consuming and not suitable for live applications. The inference time of our solution is 70 milliseconds for all steps combined (feature generation, rigid and deformation correction), giving it the capability to run at 14 FPS for live applications. Furthermore, our feature map based registration solution does not depend on the modalities used for fusion procedures. With real-time capability and high rigid and deformation correction accuracy, our solution can be used clinically for cognitive fusion biopsies. The present solution is limited to 2D registration and is not capable of compensating for out of plane motion. In future, we would like to extend our proposed registration framework to perform 3D registration and compensate for out of plane motion as well. Our solution use case can also be extended for work on live needle tracking in prostate fusion biopsy [9].

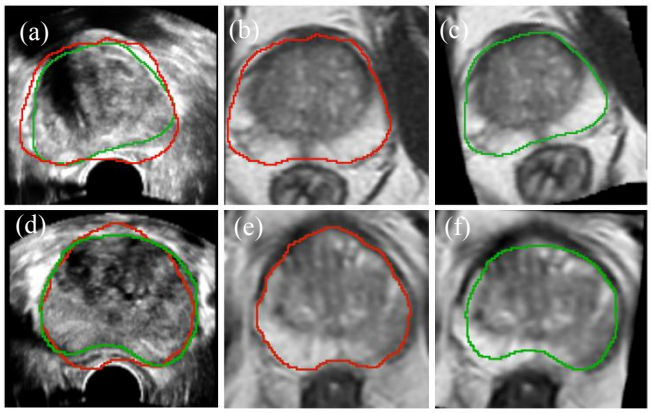

Figure 3. Results of correction on two patient data (row). (a)(d) shows TRUS image overlaid with uncorrected MR prostate contour (red) and rigid + deformation corrected MR prostate contour (green). (b)(e) shows unregistered MR image with overlaid prostate contour. (c)(f) shows registered MR image with overlaid prostate contour.

ACKNOWLEDGMENT

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2018.html

[2] Rabbani F, Stroumbakis N, Kava BR, Cookson MS, Fair WR. Incidence and clinical significance of false-negative sextant prostate biopsies. J Urol 1998;159:1247–1250. [PubMed: 9507846]

[3] De Silva, Tharindu & Fenster, Aaron & Cool, Derek & Gardi, Lori & Romagnoli, Cesare & Samarabandu, Jagath & D Ward, Aaron. (2013). 2D-3D rigid registration to compensate for prostate motion during 3D TRUS-guided biopsy. Medical physics. 40. 022904. 10.1118/1.4773873.

[4] S. Xu, J. Kruecker, B. Turkbey, N. Glossop, A. K. Singh, P. Choyke, P. Pinto, and B. J. Wood, "Real-time MRI-TRUS fusion for guidance of targeted prostate biopsies," Comput. Aided Surg. 13 (5), 255–264 (2008)

[5] Zhang, Shihui & Jiang, Shan & Yang, Zhiyong & Liu, Ranlu. (2015). 2D Ultrasound and 3D MR Image Registration of the Prostate for Brachytherapy Surgical Navigation. Medicine. 94. e1643.

[6] Bhardwaj, Aditya, et al. "Rigid and deformation corrections in real-time using deep learning for prostate fusion biopsy." Medical Imaging 2020: Image-Guided Procedures, Robotic Interventions, and Modeling. Vol. 11315. International Society for Optics and Photonics, 2020.

[7] Oliveira, Francisco PM, and Joao Manuel RS Tavares. "Medical image registration: a review." Computer methods in biomechanics and biomedical engineering 17.2 (2014): 73-93.

[8] Balakrishnan, Guha, et al. "Voxelmorph: a learning framework for deformation medical image registration." IEEE transactions on medical imaging 38.8 (2019): 1788-1800.

[9] Mukhopadhyay, Soumik, et al. "Deep learning based needle tracking in prostate fusion biopsy." Medical Imaging 2021: Image-Guided Procedures, Robotic Interventions, and Modeling. Vol. 11598. International Society for Optics and Photonics, 2021.