2013 IEEE/RSJ International Conference on
Intelligent Robots and Systems (IROS)
November 3-7, 2013. Tokyo, Japan

Autonomous eFAST Ultrasound Scanning by a Robotic Manipulator using Learning from Demonstrations

George P. Mylonas*, *Member, IEEE*, Petros Giataganas*, Muzzafer Chaudery, Valentina Vitiello, *Member, IEEE,* Ara Darzi, Guang-Zhong Yang, *Fellow, IEEE*

Abstract— **We propose a learning-based controller to enable autonomous execution of the eFAST scanning by a lightweight robotic manipulator according to expert demonstrations. The benefits of this approach are two-fold. Firstly, the automatically acquired USS images can be sent to the expert radiologist from a remote location without the need for complex robotic tele-operation. Secondly, the application of learning by demonstration alleviates the complexity of robotic programming and allows extracting operator-specific knowledge *in situ* in a natural and intuitive way. The provision of incorporating force information can further improve the versatility of the system, allowing easy adaptation to different dynamic environments.**

I. Introduction

Intracavitary hemorrhage is potentially life threatening as it is considered to be non-compressible bleeding. This has been highlighted by military academics who state that in the past decade of war the majority of fatal hemorrhage has been due to truncal trauma (67.3%) [1]. In addition, identifying intracavitary bleeding in the pre-hospital care setting whether in the chest or abdomen remains a challenge with clinical examination alone being insufficient to diagnose bleeding. The Extended Focused Assessment with Sonography in Trauma (eFAST) scan has proven to be an effective diagnostic tool in experienced hands to identify bleeding. The traditional FAST scan examines for intraperitoneal fluid (perihepatic, perisplenic and pelvis) and this has recently progressed to the eFAST which also examines for pleural, pericardial fluid and pneumothoraces. Typically, the eFAST scan involves the radiologist placing the UltraSound Scanning (USS) probe at specific points on the thorax and interpreting the images [2, 3]. In spite of the advantages of this technique in terms of portability, accuracy and noninvasiveness, the need for an expert operator to generate meaningful images has limited its use in pre-hospital settings such as ambulances or combat zones. The concept of tele-ultrasound has recently been introduced to overcome this limitation [4]. USS images can be sent in real-time from a remote location to an expert sonographer, who guides untrained personnel to perform the scan. Nonetheless, the efficacy of this approach depends on the reliability of the communication with the expert, which can be affected by the cluttered and chaotic remote environment.

All authors are with the Hamlyn Centre for Robotic Surgery, Institute of Global Health Innovation, Imperial College London, SW7 2AZ, London, UK

e-mail: {george.mylonas; petros.giataganas11; muzzafer.chaudery10; v.vitiello07; a.darzi; g.z.yang}@imperial.ac.uk

* Corresponding author: George P. Mylonas, Hamlyn Centre for Robotic Surgery, Institute of Global Health Innovation, Imperial College London, St Mary's Campus, 3rd Floor, Room 5, Paterson Wing, South Wharf Road, London W2 1NY, (tel:+44(0)20 3312 5145)

To improve the accuracy of remote USS, a number of robotic tele-echography systems have been developed. Most efforts have been focused on the design of slave manipulators [5-7] and advanced control architectures such as visual servoing [8] and force-feedback tele-operation [9, 10] to optimize the remote manipulation of the USS probe and therefore the quality of the images. Other research groups have used custom-made [11] or industrial [12] lightweight robotic arms with a redundant structure to achieve the dexterity required by the scanning task. However, all of the above systems are still limited by the use of robotic tele-operation, which requires complex control strategies to avoid the instability introduced by time-delays and loss of data during communication between the master and the slave [13]. In addition, although previous studies have implemented force control of the robotic manipulator through variable stiffness of the joints to ensure safe interaction with the patient [11, 12, 14], the scanning trajectories are simple and the force exerted by the tip of the ultrasound probe on the tissue is only estimated on the basis of the torque measured at the joints.

To the authors knowledge, only one research group has previously used a “teaching” mode to autonomously execute ultrasound scanning using a robotic manipulator [15]. However, the system simply records an input trajectory, which is generated according to the force and torque measurements acquired by a sensor mounted at the end-effector of an industrial manipulator and directly handled by the operator. In addition, the resulting scanning trajectory can only be reproduced accurately in the same structured environment where the robot needs to be accurately registered to the target under examination. In this paper, we propose a learning-based controller to enable autonomous execution of the eFAST scanning by a lightweight robotic manipulator according to expert demonstrations. The benefits of this approach are two-fold. Firstly, the automatically acquired USS images can be sent to the expert radiologist from a remote location without the need for complex robotic tele-operation. Secondly, the application of learning by demonstration alleviates the complexity of robotic programming and allows extracting operator-specific knowledge *in situ* in a natural and intuitive way. The provision of integrating force information can further improve the versatility of the system, allowing easy adaptation to different dynamic environments.

978-1-4673-6357-0/13/$31.00 ©2013 IEEE

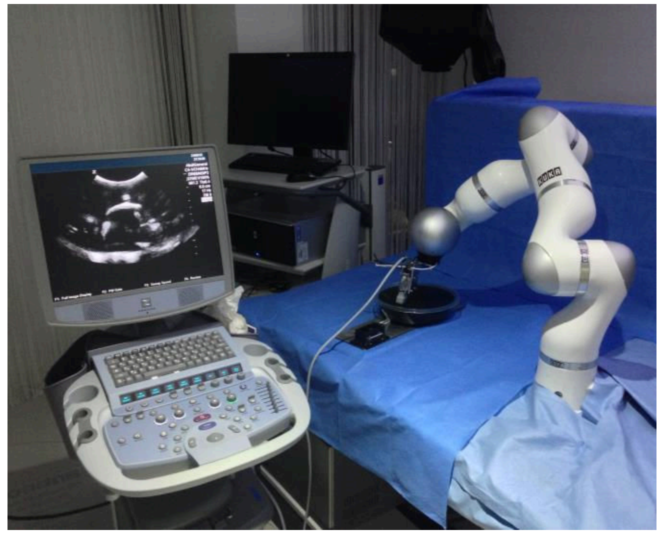

Fig. 1. On the left, the experimental setup used for the expert demonstration data collection is shown. The key components of the experimental setup are the medical ultrasound machine, an ultrasound phantom model and a handheld ultrasound transducer with force sensing and built-in 3D position markers. The respective optical-tracking camera is also shown on the background. On the right, the experimental setup used for the execution of the autonomous ultrasound routines learned by the proposed framework is shown. Here, the key component is the robotically controlled ultrasound probe which replaces the human operator.

II. System Description

The aim of the study is to demonstrate the feasibility of automating an eFAST ultrasound examination by means of a robot. For the first part of the task, learning from expert demonstrations is used. The last part involves the execution of the learned task by the robot as shown in Fig. 1.

The key components of the experimental setup are the medical ultrasound machine and the robot. For data acquisition, position and force-tracking of a hand-held ultrasound probe is used. For the execution of the learned demonstrations, the ultrasound probe is mounted on a robotic arm. For both the demonstration and automated execution phases, the same custom-made ultrasound phantom model is used.

A. *Ultrasound machine and probe*

For the ultrasound examination, a Zonare USS machine (Zonare Medical Systems, Inc. Mountain View, CA USA) was used as the imaging platform. The machine provides more than 100,000 dynamic channels per frame at more than 1000 frames per second and a total system dynamic range of 220 dB. The images are displayed on a 19” high resolution LCD monitor mounted on an articulating arm with 1280x1024 pixels resolution and a minimum of 400:1 contrast. For this study, the C4-1 Curved Array transducer (Zonare PN: Z119-00) was used which has a bandwidth of 1-4MHz and provides 80 degrees of viewing angle. The curved transducer has a number of applications such as abdominal, vascular and obstetrics imaging. In our experiments the settings were fixed at standard abdominal 2D B-mode at a depth of 6cm. Images are stored and processed in DICOM format.

For the hand held probe, the transducer is mounted onto an aluminum plate using a pair of sliding rails. This allows easy assembly and disassembly for quick conversion between the demonstration and the robotic tasks. Two force sensors (Honeywell FSG-15N1A) are appropriately mounted to provide the interaction force between the probe and the ultrasound phantom in 1 DoF along the longitudinal axis of the transducer. The whole assembly is packaged inside a plastic casing as seen in Fig. 2. A metallic boom is used to support a cubic structure that houses 12 infrared optical markers. These are used in combination with the Optotrak Certus system (NDI, Ontario, Canada) to provide and record the pose of the handheld probe at an average frequency of 40Hz. The force readings from the force sensor are amplified and digitized using an Atmel based microcontroller that connects to the data collection computer over a USB port.

For the part of the study involving the robot, the plastic

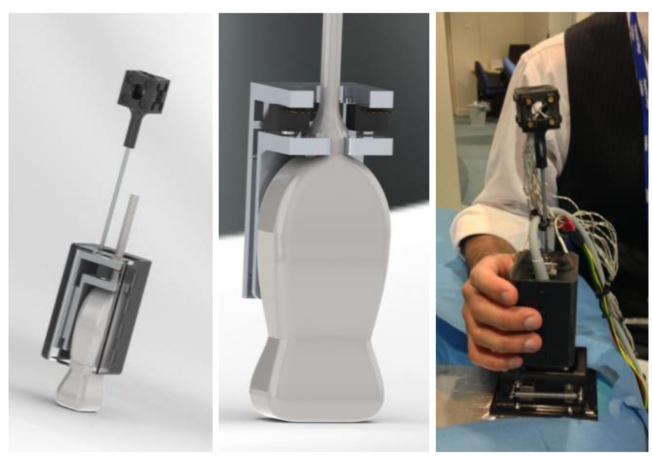

Fig. 2. The image on the left shows the hand-held ultrasound probe packaged inside a plastic casing. The integrated force sensors, shown as the black blocks in the centre picture, allow sensing of the forces exerted between the probe and the tissue. A cubic rigid body (shown on top of the attached rod) allows continuous tracking of the 3D position and pose of the probe. On the right the actual device held by the operator in its homing cradle.

casing is removed and the probe is mounted on the robot's end-effector using an L-shaped back-plate as seen in Fig. 3. The force sensing capability is still available.

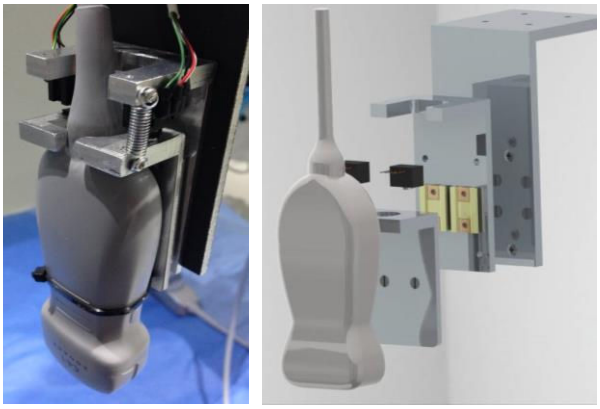

Fig. 3. The image shows the easy conversion of the hand-held probe into a robotically held one. The force sensors (black blocks) are still providing readings on the probe-tissue interaction forces.

B. Robot

For this study, a KUKA Light-Weight Robotic arm (KUKA Roboter GmbH, Augsburg, Germany) is used. This is a 7-DoF manipulator which is rapidly gaining wide acceptance among research institutions around the world. The redundant structure of the KUKA LWR offers small footprint, enhanced dexterity and flexibility. Its lightweight structure in addition to the integrated torque sensors and position sensing on each axis, allow increased safety. For this study, the robot is controlled using the Fast Research Interface (FRI) library and in Cartesian impedance control mode.

C. Examination phantom

To carry out the study, an ultrasound compatible phantom was made. Inside the phantom, a number of letters (A, C and D) are placed and play the role of ultrasound identifiable targets which are not visible to the naked eye. The letters are placed at random orientations, as shown in Fig. 4, in such a way that often considerable probe angles are required in order to be identified. The model was made in two parts using 500mls of hot water to which 10 grams of gelatin was added and stirred until dissolved. This was then diluted up to 1 liter with cold water and refrigerated for 12 hours to allow consolidation into a jelly consistency. The letters were made with plasticine and these were embedded on top of the phantom jelly. Another layer of gelatin and water was made up to 1 liter and poured over the 1st layer to encase the letters. Green food coloring dye was added to make the phantom opaque so that the letters were not visible and then it was refrigerated for another 12 hours before the phantom was ready for the experiment. In between experiments the phantom was kept refrigerated to ensure that it maintained the same consistency. Once ready for experiments, the phantom was mounted onto a metallic plate and was fixed into position based on a number of markers. Every effort was made to ensure the phantom was at the correct alignment at all times to maintain consistency between experiments. In order to provide a world frame-of-reference which is common and consistent between the manual and the robotic tasks, a probe homing base was also included as shown in Fig. 4. The base allows positioning of the probe at the same relative position with respect to the phantom model during both the manual and the robotic tasks.

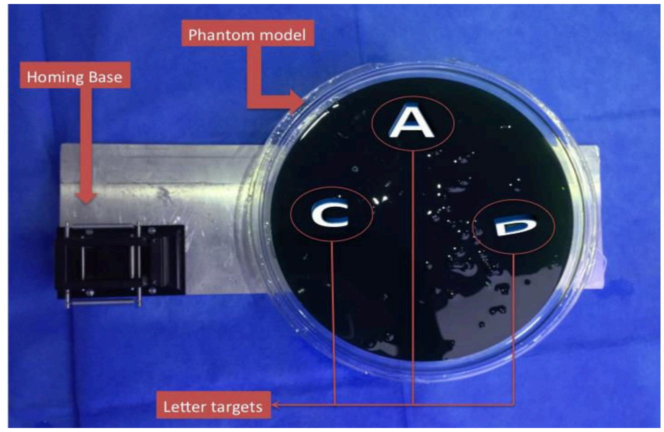

Fig. 4 The ultrasound phantom model as seen from the top. The general location of the letter targets is highlighted. On the left, the probe homing base can be seen.

III. METHOD

In order to demonstrate the feasibility of the proposed approach, the study is conducted in two parts; the expert demonstration and the automation parts. For the demonstration, an experienced sonographer is manually controlling an ultrasound probe to carry out a number of demonstrations of the ultrasound examination task. For the automation part, the ultrasound probe is mounted onto the robot that is used to execute the learned demonstrations. In both cases the same ultrasound phantom is used.

The general objective of the task is to scan the phantom and clearly identify all three letters which are embedded in it. During the demonstrations phase, the pose of the ultrasound probe is continuously tracked, along with the exerted force between the probe and the phantom. The recorded probe trajectories are then fed to the learning algorithm which is in charge of generating the learned demonstration. For the final part, the learned trajectories are mapped into robot trajectories and the ultrasound examination is autonomously carried out by the robot. The expert sonographer is also present during this last phase for providing the expert validation of the approach. It should be noted that although recorded, no force information is used with the learning algorithm at this stage.

A. Expert demonstrations

An expert demonstration starts with the handheld ultrasound probe positioned inside the homing base. The position of the probe is continuously recorder by means of the attached optical markers. The sonographer then moves the probe at the general location where each of the letter targets is located. By using visual feedback from the ultrasound screen the examination continues until the letter is clearly identifiable by the operator. Each target is acquired from roughly the same starting position which corresponds to the transducer touching the surface at the general area over the relevant letter. This ensures uniformity between experiments. The scan continues with the next letter until all three letters are clearly identified. For this study one sonographer performed each letter-scanning task twice.

B. Learning algorithm

The demonstration data are acquired through the tracking system as an expert operator performs the desired movements. These position data are imported to the learning algorithm which is based on Gaussian Mixture modeling (GMM) and Gaussian Mixture Regression (GMR). The method has first been used in [16] which is a modified version of Calinon's original algorithm [17]. The algorithm's goal is to encode the statistical characteristics of the task which is represented by the set of demonstrated trajectories. The output of the algorithm represents the learned generalized task. For the task specified above, during each of the n demonstrations the 3D position and the orientation in quaternion form of the ultrasound tip are recorded. Each training dataset is performed through T time units and consists of N data points. The dimensionality of the learned parameters for this specific case is 8, including the temporal information parameter. More specifically, the dataset is represented as $u = (u_{t,i}, u_{p,i}, u_{q,i})_{i=1}^{N}$ where the parameter $u_t \in \mathbb{R}$ corresponds to the temporal data, the parameter $u_p \in \mathbb{R}^3$ to the spatial data and the parameter $u_q \in \mathbb{R}^4$ to the orientation information.

The GMM algorithm models the demonstration in a stochastic way with a joint density estimate using a mixture of K Gaussian components of dimensionality D. Each Gaussian component corresponds to a single multivariate Gaussian distribution $\{\mu_K, \Sigma_K\}$ and the general GMM is defined by the probability density function as:

$$p(\mathrm{u}_j) = \sum_{\mathrm{k}=1}^{K} p(k) p(\mathrm{u}_j \mid k) \tag{1}$$

where the parameter $p(\mathrm{u}_j \mid k) = N(u_j; \mu_k, \mathrm{S}_k)$ is the conditional probability density function referring to the normal Gaussian distribution of the component k and the parameter $p(k) = \omega_k$ is a prior probability. The GMM encapsulates the correlations and the variations between the demonstrations. The K Gaussian components are then calculated iteratively through a batch learning algorithm based on the Expectation–Maximization (EM) algorithm [18], that optimizes the fit of the K Gaussian components to the training data by finding the maximum likelihood estimator. The maximization of the model likelihood is aided, before the derivation of the statistical modeling of the task, by a time-alignment pre-processing step based on Dynamic Time Warping [19]. Furthermore, the Bayesian information criterion [20] is incorporated in the proposed learning framework for the selection of the number of components as the latter significantly affect the bias-variance trade-off. A low bias gives an increased accuracy to the estimation, while a low variance provides smooth results in the estimation. The unbalance of these parameters can result to over-fitting or over-smoothing of the demonstrated datasets.

After modeling the demonstrated datasets, the GMR algorithm is implemented to reproduce the learned task in a continuous manner as it provides smooth trajectories while incorporating the essential constraints of the demonstrated tasks. These features are included in the joint density probability given by the GMM algorithm, which is then used as an input to estimate the regression function. In particular, at every time-step $\{\mathrm{u}_t\}$, which is used as a query point, the output generalized parameters $\{\hat{u}_i, \hat{\Sigma}_i\}$ of the GMR are:

$$\hat{u}_i = \sum_{k=1}^{K} \pi_k \hat{u}_k^{(i)}, \hat{\Sigma}_i = \sum_{j=1}^{K} \pi_k^2 \hat{\Sigma}_k^{(ii)} \tag{2}$$

where $i = \{p, \mathrm{q}\}$ is referring to the position and the quaternion data, π_k is the probability that the k Gaussian component is responsible for the query point $\{\mathrm{u}_t\}$ and the parameters $\{\hat{u}_k^{(\mathrm{i})}, \hat{\Sigma}_k^{(ii)}\}$ are the estimated conditional values. Thus, the aforementioned generalized parameters describe the reproduced motion with the covariance values $\hat{\Sigma}_k^{(ii)}$ corresponding to the variations/constraints of the task as can be seen in Fig. 6 on the top left. A more detailed description of the algorithm is presented in [16].

C. Autonomous execution

After the learning step, the generated trajectories are ready to be executed by the robot. The ultrasound probe is first mounted onto the robot's end effector. To make sure that the two coordinate frames are aligned, the probe's initial position is set at the homing base next to the ultrasound phantom, as shown in Fig. 5 and as it was the case during the demonstration task (shown in Fig. 2 right). The initial pose of the probe for each respective letter-target, is based on the obtained pose during the demonstration part of the study. By establishing a common frame of reference through the homing position, ensures that the robotic probe can autonomously move in between all points during execution of the robotic task. To ensure safe execution during the movement of the tool, the motion between points is calculated as a minimum jerk trajectory.

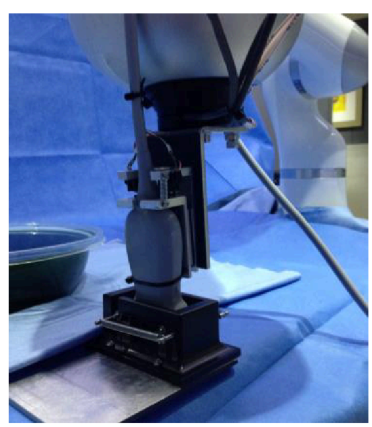

Fig. 5. The robotically held probe is shown at the homing position. The homing position is used to establish a common frame of reference between the manual demonstration and the autonomous robotic execution phases.

D. Validation

A qualitative validation of the autonomous robotic execution of the learned tasks is based on visual observation by the expert operator. During robotic execution, the same operator that was used to obtain the demonstration trajectories, is employed and asked to observe the ultrasound screen. When each target scan is completed, the expert operator is asked to comment on the perceived success of the automated task execution. More quantitative measures of the learning success are also obtained by comparing the demonstration and the learned trajectories.

IV. Experimental Results

Visual observation of the autonomous robotic execution task by the expert sonographer, has demonstrated the success of the proposed methodology. The series of photo snapshots shown at the bottom of Fig. 6 represent the ultrasound slices obtained during or at the end of the scanning evolution for

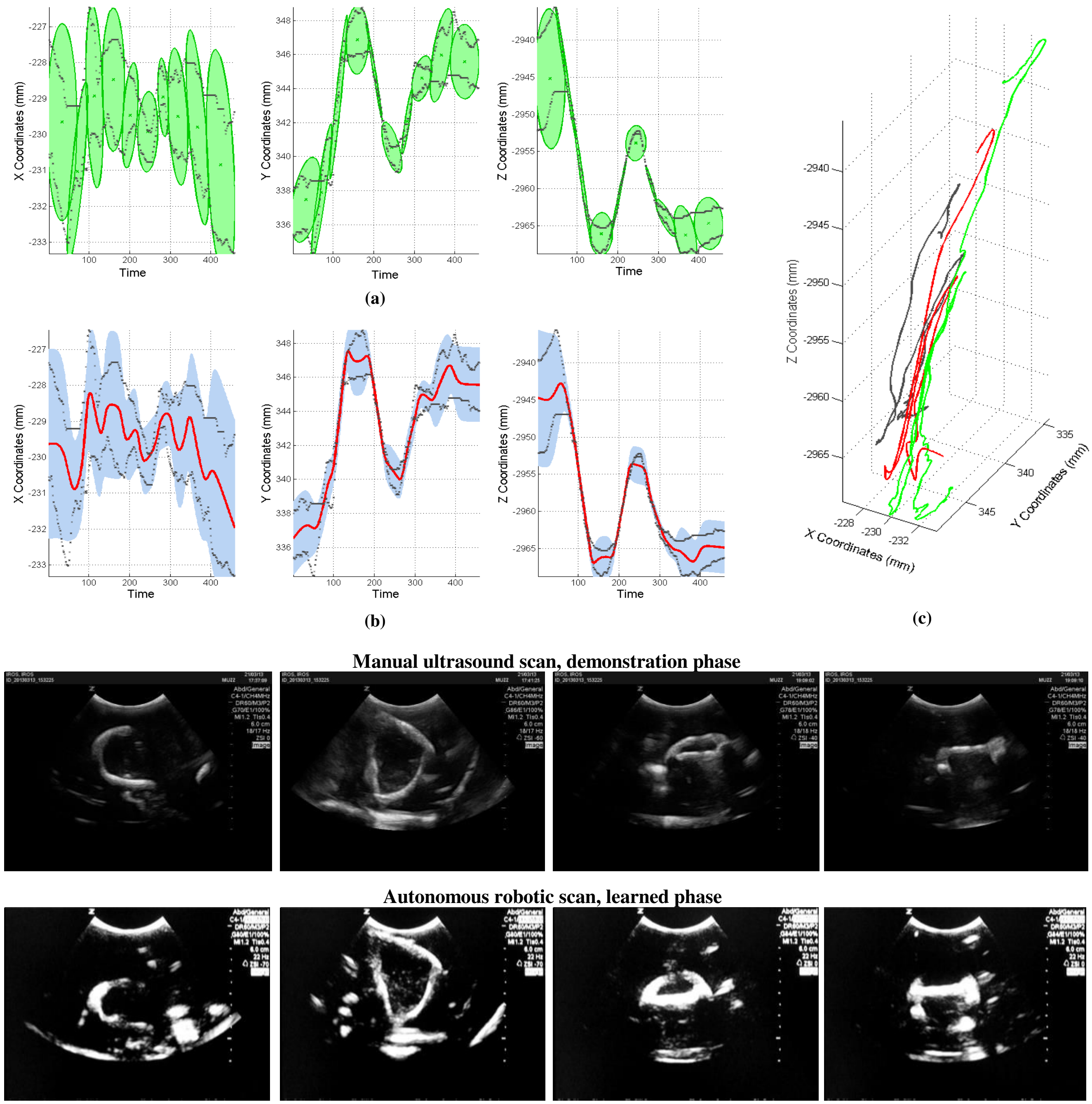

Fig. 6. (a) The GMM components referring to the spatial data used for modeling the set of demonstrated trajectories for letter D. (b) An illustration of the generated spatial trajectories from the GMR algorithm that encapsulate the variations between demonstrations. The shaded areas correspond to the task covariance. (c) The training (green and black lines) and generated (red line) 3D trajectories for letter D. At the bottom, the first row shows identified targets during the demonstration phase and the second row shows the respective targets identified during learned execution of the task by the robot alone. From left to right, the targets are letters C, D, upper part of letter A and lower part of letter A. The image quality difference between top and bottom rows is because of the different video acquisition medium used. The top row snapshots are acquired by the Zonare built-in video recorder, while the bottom row snapshots are from an external video camera. This is due to the maximum recording time restrictions imposed by the US machine and the long execution times of the learned trajectories by the robot.

each of the target letters. At the row on top are the images captured during the demonstration phase and at the row on the bottom are the respective captures during the autonomous robotic scan. In both cases, it is obvious that the targets are clearly identifiable. The last two snapshots on both rows correspond to parts of letter A, while the first to snapshots correspond to letter C and D respectively. The difference in image intensity between the top and bottom rows is because of the different video acquisition methods used due to the recording time restriction imposed by the US machine and the long execution times provided by the robot.

The plots on the top of Fig. 6 provide a quantitative measure of the task success.

V. Discussion and Conclusions

The clinical translation of this study would be to automate the eFAST scan by creating a machine learning algorithm that would allow the robot to independently move to and scan fixed points on the torso as the radiologist does in the FAST scan. The images could then be relayed across to the radiologist who can remotely interpret them. The remote presence application of this robot would mean that it could be of significant use in pre-hospital care medicine, remote environments, space and military medicine.

Although very promising, this is a preliminary study and future work will have to investigate a number of additions to the existing implementation. For this study no force information has been incorporated in the learning or execution phases. With ongoing work we are investigating probe and tissue interaction forces to further enhance autonomy and ensure safety. Learning manipulation-force patterns could improve ultrasound repeatability, allow the implementation of active safety constraints, carry out elastography-based examinations and even allow hemorrhage control. Another exciting extension would be the implementation of visual servoing by incorporating image semantics into the learning process. Ongoing work is also investigating real-time 3D model registration and fitting for automatic identification and localization of body areas. This, in combination with force sensing, will further improve autonomy and compliance to human body and environmental variability.

To our knowledge this is the first study of this kind, implementing a learning-based controller to enable autonomous execution of the eFAST scanning by a lightweight robotic manipulator according to expert demonstrations.

Acknowledgments

The authors would like to thank Dr James Clark for his helpful suggestions on the experimental setup.

References

[1] B. J. Eastridge, R. L. Mabry, P. Seguin, J. Cantrell, T. Tops, P. Uribe, *et al.*, "Death on the battlefield (2001-2011): implications for the future of combat casualty care," *J Trauma Acute Care Surg,* vol. 73, pp. S431-7, Dec 2012.

[2] S. V. Soundappan, A. J. Holland, D. T. Cass, and A. Lam, "Diagnostic accuracy of surgeon-performed focused abdominal sonography (FAST) in blunt paediatric trauma," *Injury,* vol. 36, pp. 970-5, Aug 2005.

[3] J. M. Hsu, A. P. Joseph, L. J. Tarlinton, L. Macken, and S. Blome, "The accuracy of focused assessment with sonography in trauma (FAST) in blunt trauma patients: experience of an Australian major trauma service," *Injury,* vol. 38, pp. 71-5, Jan 2007.

[4] K. S. Boniface, H. Shokoohi, E. R. Smith, and K. Scantlebury, "Tele-ultrasound and paramedics: real-time remote physician guidance of the Focused Assessment With Sonography for Trauma examination," *The American Journal of Emergency Medicine,* vol. 29, pp. 477-481, 2011.

[5] L. Nouaille, P. Vieyres, and G. Poisson, "Process of optimisation for a 4 DOF tele-echography robot," *Robotica,* vol. 30, pp. 1131-1145, 2012.

[6] S. Kumar, J. Marescaux, J. J. Banihachemi, E. Boidard, J. L. Bosson, L. Bressollette, *et al.*, "The Tele-Echography Robot: a Robot for Remote Ultrasonic Examination," in *Telesurgery*, S. Kumar and J. Marescaux, Eds., ed: Springer Berlin Heidelberg, 2008, pp. 91-99.

[7] R. Seip, A. Katny, W. Chen, N. T. Sanghvi, K. A. Dines, and J. Wheeler, "1C-1 Remotely Operated Robotic High Intensity Focused Ultrasound (HIFU) Manipulator System for Critical Systems for Trauma and Transport (CSTAT)," in *IEEE Ultrasonics Symposium*, 2006, pp. 200-203.

[8] P. Abolmaesumi, S. E. Salcudean, Z. Wen-Hong, M. R. Sirouspour, and S. P. DiMaio, "Image-guided control of a robot for medical ultrasound," *IEEE Transactions on Robotics and Automation,* vol. 18, pp. 11-23, 2002.

[9] C. Sousa, R. Cortesao, and L. Santos, "Computed torque posture control for robotic-assisted tele-echography," in *18th Mediterranean Conference on Control & Automation (MED)*, 2010, pp. 1561-1566.

[10] A. Capri, T. Slama, G. Charron, A. Fonte, N. Vincent, and P. Vieyres, "A mechatronic tele-operated system for echography using visual navigation assistance and a model based bilateral predictive control," in *IEEE International Symposium on Industrial Electronics (ISIE)*, 2007, pp. 3144-3149.

[11] P. Valdivia y Alvarado, C. Chu-Yin, and K. Hynynen, "Design of a manipulator system for hemorrhage detection and treatment using High Intensity Focused Ultrasound," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2009, pp. 4529-4534.

[12] T. Beyl, T. Brennecke, J. Raczkowsky, and H. Worn, "Ultrasound tomography using a light weight robot," *Proceedings of the Computer Assisted Radiology-25th International Congress and Exhibition - International Journal of Computer Assisted Radiology and Surgery,* vol. 6, pp. 305-366, 2011.

[13] P. F. Hokayem and M. W. Spong, "Bilateral teleoperation: An historical survey," *Automatica,* vol. 42, pp. 2035-2057, 2006.

[14] F. Pierrot, E. Dombre, E. Dégoulange, L. Urbain, P. Caron, S. Boudet, *et al.*, "Hippocrate: a safe robot arm for medical applications with force feedback," *Medical Image Analysis,* vol. 3, pp. 285-300, 9// 1999.

[15] M.-A. Janvier, L.-G. Durand, M.-H. R. Cardinal, I. Renaud, B. Chayer, P. Bigras, *et al.*, "Performance evaluation of a medical robotic 3D-ultrasound imaging system," *Medical Image Analysis,* vol. 12, pp. 275-290, 2008.

[16] P. Giataganas, V. Vitiello, V. Simaiaki, E. Lopez, and G.-Z. Yang, "Cooperative in situ Microscopic Scanning and Simultaneous Tissue Surface Reconstruction using a Compliant Robotic Manipulator," in *ICRA 2013*, 2013, pp. 5358-5363.

[17] S. Calinon, F. Guenter, and A. Billard, "On Learning, Representing, and Generalizing a Task in a Humanoid Robot," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on,* vol. 37, pp. 286-298, 2007.

[18] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum Likelihood form incomplete data via the EM algorithm," *Journal of Royal Statistical Society,* vol. Series B, vol. 39, pp. 1-38, 1977.

[19] P. Senin, "Dynamic Time Warping Algorithm Review," 2008.

[20] G. Schwarz, "Estimating the dimension of a model," *The annals of statistics,* vol. 6, pp. 461-464, 1978.