Semi-Autonomous Navigation for Robot Assisted Tele-Echography using Generalized Shape Models and Co-Registered RGB-D cameras

Lin Zhang*, Su-Lin Lee, Guang-Zhong Yang, *Fellow, IEEE*, and George P. Mylonas, *Member, IEEE*

***Abstract*— This paper proposes a semi-autonomous navigated master-slave system, for robot assisted remote echography for early trauma assessment. Two RGB-D sensors are used to capture real-time 3D information of the scene at the slave side where the patient is located. A 3D statistical shape model is built and used to generate a customized patient model based on the point cloud generated by the RGB-D sensors. The customized patient model can be updated and adaptively fitted to the patient. The model is also used to generate a trajectory to navigate a KUKA robotic arm and safely conduct the ultrasound examination. Extensive validation of the proposed system shows promising results in terms of accuracy and robustness.**

I. INTRODUCTION

A large and growing body of literature is dedicated to investigations on robot assisted tele-echography. Most of the studies are based on slightly different flavors of a master-slave setup. In [1], a parallel uncoupled robot positioned on the patients torso, was used to control an ultrasound transducer. For surveillance of the slave side, live video streams captured by webcams were transferred to the master side via ISDN links. Another master-slave system using a custom slave manipulator was presented in [2]. Here, audio and video tele-conferencing facilities were utilized for surveillance. Other studies, like the one presented in [3], are investigating remote Focused Assessment with Sonography for Trauma, commonly known as FAST, which is a standardized ultrasound examination used to detect and monitor hemorrhage. The examination concentrates on four body areas on the patient; right and left abdomen, pericardium and pelvis. In [4] a robotic arm was used to combine 2D ultrasound images and reconstruct a 3D model of the scanned anatomy in real-time. In [5], two robotic arms were used along with a haptic manipulator for haptic feedback at the master side. Additionally, a Kinect camera was used for rough body skeleton tracking but without the ability to accurately determine and track the exact location of the human skin. What becomes apparent in existing literature investigating remote examinations like FAST, is that they all assume that paramedics or emergency assisting personnel are always available at the slave side to initialize the required body locations for the robot to scan.

In order to accurately determine and track a required location on the human skin, several studies have attempted to build deformable models and flexible templates. In these methods, the model is parameterized so that the variation of its shape and pose can be controlled in several modes. Statistical shape modeling (SSM) is a method that has been widely applied on medical imaging segmentation and anatomical structure reconstruction such as the brain, liver, bones, atria and heart chambers [6][7]. In [8] a 3D deformable model was introduced in order to provide force feedback to the medical expert during tele-echography. However, the model required patient specific manual calibration and registration could not be updated on-line and according to body changes. In the study presented here and in order to generate a customized model of the patient we deploy SSM by extending [6] to 3D and build an on-line updated deformable model that is able to conform the patients motion and movement based on points collected using two Kinect sensors. This method allows adapting the shape model in real time by tuning a number of parameters.

In the study presented here, we present a master-slave navigation system for robot assisted tele-echography which is capable of semi-automating ultrasound scanning. Based on a generic human model, our approach generates a customized model that is able to conform to breathing motion and body movement of any patient. For generating the patient specific model we deploy SSM by extending [6] to 3D. The model is deformable and can be updated on-line, according to the 3D points collected using two Kinect sensors. Adaptation of the generic shape model is achieved by tuning a number of parameters as will be described later.

An interface at the master side for visualization and control is also developed. Ultrasound examination routines, like FAST, can be preprogrammed on the generic model and automatically adapted to any human body in real time. Additionally, based on a generic body model, sonographer-specific skills and examination habits could be encoded [9], and then autonomously reproduced in the patient specific domain. To the authors knowledge, this is the first study that attempts to address the navigational aspects pertinent to robot assisted tele-echography.

II. SYSTEM DESCRIPTION

The main objective of the proposed system is to create a customized model that is compliant to the patients position and physiological motion. With the customized model, navigation of the robotic ultrasound examination can be planned accordingly. In our setup the slave side is responsible for collecting 3D information using RGB-D cameras while the master side is responsible for processing the 3D information

All authors are with The Hamlyn Centre for Robotic Surgery, Institute of Global Health Innovation, Imperial College, SW7 2AZ, London, UK
email:{lin.zhang11; su-lin.lee; g.z.yang; george.mylonas}@imperial.ac.uk
*Corresponding author: Lin Zhang, Hamlyn Centre, 3rd Floor, Patterson Wing, St. Mary's Hospital, W2 1NY, London, UK
Some materials presented here are from Lin Zhangs MRes thesis.

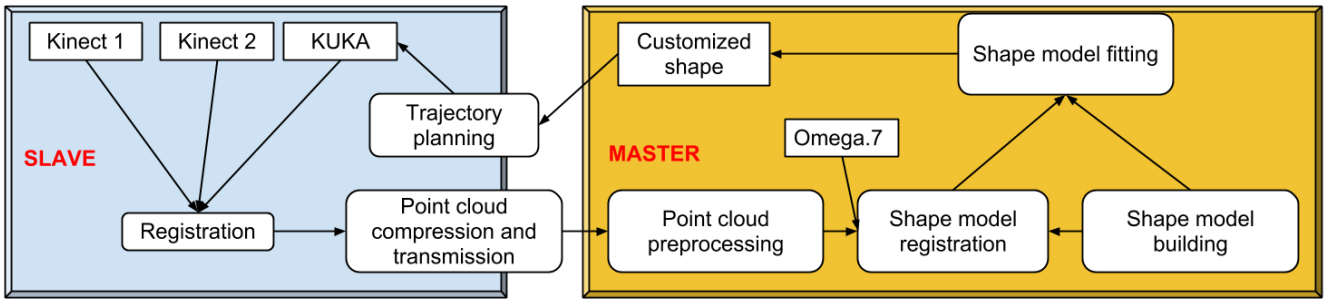


Fig. 1: A schematic of the system describing the processes and components on both slave and master sides. For this study, the two sites are connected via a Local Area Network.

and for creating the customized model for navigation. The two sides communicate via a LAN or WAN connection.

The main hardware components of the system are two RGB-D cameras, a slave robotic arm and a master robotic manipulator.

A. The Framework

As shown in Fig. 1, the proposed framework consists of two main parts; the slave side for patient point-cloud registration and processing, and the master side for model fitting, trajectory planning and visualization. At the slave side, two RGB-D sensors are placed opposite to each other and facing toward the human target torso (Fig. 2a). The reason for using two sensors is that the occlusions from the target itself and other objects appearing in the scene, like the robot, may lead to failing detection of the target. One-off registration allows registering the point clouds produced by the two RGB-D sensors into the robot coordinate frame. For data communication between the slave and the master, two separate network connections are used, one for the robot control and the other for the point cloud transmission. The real-time point cloud produced by the RGB-D sensors is relatively dense and its processing is computationally expensive. It also contains some degree of redundant information, which is irrelevant to the shape model fitting process. For this reason, the points are downsampled and compressed before are transferred to the master side. In this context, a modified octree based point cloud compression method is applied using [10], which is capable for temporal compression as well. Both point clouds from the sensors are downsampled into one-tenth of their original size.

At the master side, the compressed cloud is first decompressed and the bed surface is cropped-out, resulting in a region-of-interest containing just the target torso. As the beds relative position to the RGB-D sensors is fixed in this scenario, the region-of-interest can be defined when setting up the system. Next, a trained statistical shape model is registered with the target point cloud using the point-to-point Iterative Closest Point (ICP) algorithm [11] and initial positions given by the user, from which a rough alignment between shape model and point cloud is built. The registered shape model is subsequently fine fitted to the target cloud and a customized model is created. Based on this customized model, a circular trajectory planning algorithm is applied and the trajectory is sent to the slave robot which is then able to safely approach the desired examination site(s) on the human body.

B. RGB-D Sensor

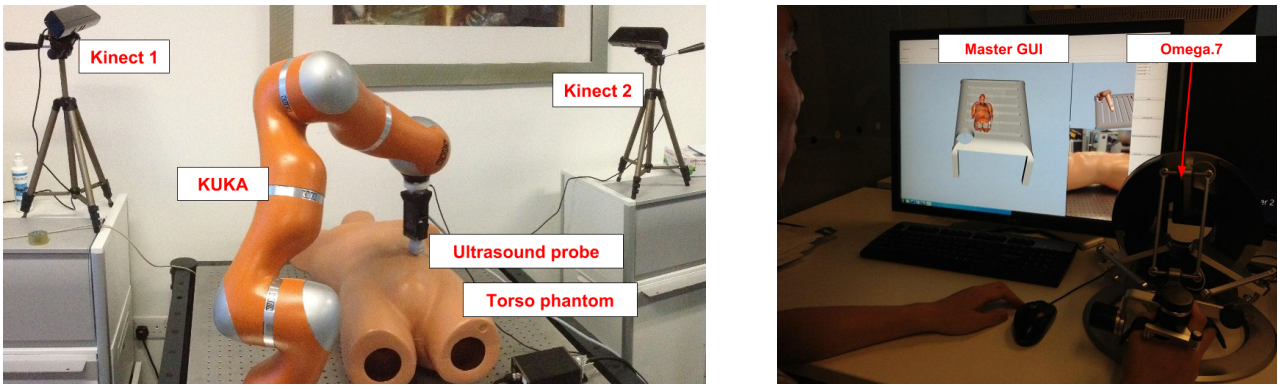
In our setup, two Microsoft Kinect sensors (first generation) are used. They are able to capture real-time depth and color images simultaneously at a frame rate of 30 fps and at a maximum resolution of 640x480 pixels. The depth image can be converted to a point cloud which is a set of 3D points describing the observed scene. The Kinect provides for wide working distances in a range of 0.8m to 4m (default mode) or 0.4m to 3m (near mode). It has an angular field-of-view of 43 vertically and 57 horizontally. Both Kinect sensors are interfaced on a single computer using two of the available USB ports controlled by separate USB controllers.

C. Robot

A 7 degrees-of-freedom (DoF) KUKA Light-Weight Robotic arm (KUKA Roboter GmbH, Augsburg, Germany) is used for echographic examination (Fig. 2). Embedded torque sensors into all joints allow collision detection with the environment. It is capable of a maximum 7 kg payload while it weights about 16 kg. For this study, the robot is controlled in Cartesian impedance control mode via the Fast Research Interface (FRI) provided by the manufacturer. An ultrasound probe is rigidly mounted at the robot end-effecter, interfaced through an ATI 6DoF force/torque sensor (ATI Industrial Automation, USA) for accurate force feedback during tele-echography. A 7 DoF Omega.7 (Force Dimension, Switzerland) haptic manipulator (Fig. 2b) is used for remote manipulation and interaction with the developed user interface.

III. METHODOLOGY

At the slave side, the point cloud data produced by the two Kinects is registered with the robot coordinate frame. At the master side, statistical shape modeling is used to create a deformable torso model. A customized model is created based on the deformable model and the point cloud data. A trajectory planning algorithm is applied for robot navigation and a visualization tool is developed. Each of these components is described in detail at the following sections.



(a) Slave setup **(b)** Master setup

Fig. 2: a) The system setup at the slave side including two Kinect sensors and a KUKA robotic arm holding an ultrasound probe. The arrangement of the Kinect sensors allows for both to image the torso in the middle. b) The master setup comprising an Omega.7 master manipulator and a GUI showing the status of the model, the KUKA robot and the torso at the slave side.

A. Point Cloud Registration

To produce a complete view of the human torso, registration between the point clouds captured from the two Kinect sensors is required. For this, we use three markers with different colors whose position and hue can be detected by both Kinect sensors. Each marker is represented by a cluster of points and its geometric center is considered as the position of this markers. To identify the markers based on their colors, different thresholds are set in the HSV color space.

In addition to the Kinect-to-Kinect registration, the point clouds are also register with the robot frame for trajectory planning and navigation. The position of the markers in this frame can be retrieved directly by placing the robot end-effecter onto each one of the markers. Both registrations are only performed once, after the robot and the Kinect sensors are setup and fixed.

B. Statistical Shape Modeling

The statistical shape model can be used to analyze and fit a new shape with the existing model. For this paper, a training set of 17 human models (9 females and 8 males) was used to build a statistical shape model of the torso [12]. Fig. 3 shows examples of the 3D meshes making up the models. As only the torso needs to be tracked, the shapes in the training set were cropped so that they are close to the final target shape. Since point correspondences were already established across the training set, this was a matter of removing the same points and facets in each mesh. A shape in 3D can be arranged as a column vector:

$$x = [x_1, x_2, \cdots, x_n, y_1, y_2, \cdots, y_n, z_1, z_2, \cdots, z_n]^T \quad (1)$$

where n is the number of points defining the torso. Therefore, the training set can be represented as a $3n$-by-k matrix in which each column is a shape and is the size of the training set.

To make two different shapes comparable, all training shapes must be aligned into a common coordinate system. For this paper, Procrustes Analysis [13] is used to align the

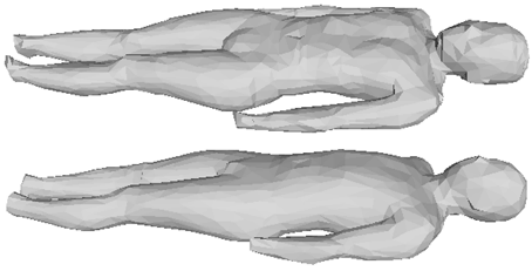


Fig. 3: Two training shapes produced from MRI scan (upper: female, lower: male).

training shapes. The method iteratively minimizes the sum of distance of each shape to the current mean of all shapes until the result has converged. The mean shape can be calculated as:

$$\bar{x} = \frac{1}{k} \sum_{i=1}^{k} x_i \quad (2)$$

In order to extract main variations within the training data, principal component analysis (PCA) was applied to a covariance matrix of the training shapes, defined as:

$$\Sigma_x = \frac{1}{k-1} \sum_{i=1}^{k} (x_i - \bar{x})(x_i - \bar{x})^T \quad (3)$$

The principal variations of the training shapes were found by calculating the eigenvectors P and the corresponding eigenvalues b of the covariance matrix. As $3n$ variables exist in each shape while only k ($k \ll 3n$) samples are available for training, the number of principal components with non-zero variance cannot exceed $k-1$. The size of the eigenvalues indicates the amount of variance in the corresponding modes of variation; the largest eigenvalues representing the most significant variation is shown in Table I. The first few rows of the table also show that main variations in the training shapes are available. In this paper, 10 modes of variation are used retaining 95% of the shape variation, the rest was treated as noise. To this end, the t largest eigenvalues can be chosen so that certain proportion (f_v) of total variance (V_T)

is retained. If the eigenvalues are sorted in descending order, then the above process can be formulated as [6]:

$$\sum_{i=1}^{t} b_i \geq f_v V_T \tag{4}$$

Afterwards, a shape y can be created using [6]:

$$y = \bar{x} + Pl, \tag{5}$$

where l is a column vector containing t shape parameters and P is the remaining t eigenvectors. If all entries of l are zero, then the created shape is equal to the mean shape in the model. By tuning different values in b, the shape will have different variations from the mean shape at the corresponding axes. As shown in Fig. 4, the main modes of variation of the human torso (leg cropped) model can be interpreted as: (a) global changing of body size, (b) local variation of chest and (c) local variation of breast.

For a new shape y_{new}, its parameters can be estimated using [6]:

$$l_{new} = P^T (y_{new} - \bar{x}) \tag{6}$$

Another functionality of SSM is to constrain the new created shape so that it is similar to those in the original training set. To this end, we can apply limits of $\pm c\sqrt{b_i}$ to the i-th element of shape parameter l in which b_i is the corresponding eigenvalue calculated from the training shapes. The constant c is a factor to control the allowable variation in each axis. It should not be very large or this may result in an arbitrary shape that is different from the range of the training set.

TABLE I: Eigenvalue of the Covariance Matrix Derived from a Training Set of the Human Torso

Eigenvalue	$(b_i/V_T) \times 100\%$
b_1	27%
b_2	16%
b_3	12%
b_4	8%
b_5	7%
b_6	5%

C. Customized Model Generation

After the shape model of the torso is generated, it is used to create a customized model according to the target point cloud of the torso. First, the mean shape of the trained model is roughly aligned to the target torso in the scene, based on manual intervention. Next, points of the mean shape are registered with the points collected from the target torso using ICP, leading to an initial alignment between the shape model and the target points.

To produce a customized model which adapts to the current target, the mean shape is updated by its correspondences in the target points captured by the sensors. For each point in the mean shape, its correspondence is determined by the point in the target cloud that has the nearest distance to its normal. To this end, the normal of each point in the shape model is calculated based on nearby face vertices. The update is achieved by replacing each point with its correspondence. Using such an updating strategy, the shape is able to fit convex as well as concave target clouds. Since the back of the torso is occluded and invisible to both Kinects, in order to improve robustness of the model only points from the front of the torso are updated.

The new updated shape can be deformed and must be constrained by limiting its shape parameters with the SSM. To this end, the new shape is aligned to the shape model using Procrustes Analysis [13] and subsequently projected to the shape parameters using (6). In this paper, the limit for the shape parameters is set to $\pm 3\sqrt{b_i}$ which enables a wide range of variations while still being able to restrict the shape reasonably. So far, we have generated a well fitted shape which is customized for the target torso in the scene. This customized model will be updated iteratively based on the live point cloud arriving from the slave side using the same procedure.

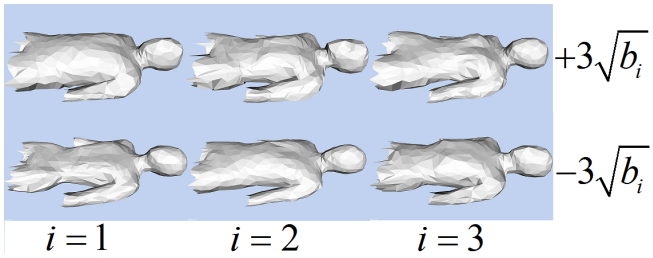


Fig. 4: Variability of a statistical model of the human torso built from 17 training sets (9 females and 8 males): in the left column the principal mode with the largest variance l_1 is varied between $\pm 3\sqrt{b_1}$ while in the second and third column the parameter is varied between $\pm 3\sqrt{b_2}$ and $\pm 3\sqrt{b_3}$ respectively.

D. Trajectory

For navigating the robot to conduct an ultrasound scan, it is necessary to create a smooth trajectory which is able to avoid collision with the human torso. For this, the method presented in [14] is applied, which allows creating a circular trajectory only based on start, middle and end waypoints. The starting point is the current robot position which is read from the slave side. The goal point can either be a pre-determined anatomical position on the model (e.g., one of the standard FAST scan points) or selected by the user using haptic manipulator and the visual interface which is presented in the next section. The selection of the middle point is important as it defines the trend of the trajectory. In this context, the middle point is always located above the torso model so that the trajectory never intersects with the torso or the table within the scene. After determining the three waypoints, we can define a unique circle and interpolate the rest of the points on the circular trajectory.

E. Visualization

In order to provide the user with a tool for surveillance and interaction, the user interface is developed based on the visualization API in PCL [15]. In addition to displaying the

shape model and point cloud received from the slave, the interface also displays a rendered ultrasound probe and a sphere cursor whose positions are updated according to the current positions of the KUKA and the Omega.7. Based on the position of the sphere cursor in the virtual environment, the user can select the initial position for the coarse shape alignment and the goal point for the robot. Finally, the resulting planned trajectory can be displayed within the interface as well.

IV. Experimental Methodology

In order to assess the performance of the proposed system, two experiments are considered. The first experiment is used to validate the deformability of the virtual shape model (VSM). For this, a respiration task is carried out using a human subject whose customized model is used for the validation. We use the NDI Polaris Vicra tracker to record the actual breathing motion of the subject. The systems stated accuracy is 0.35mm within its valid working volume. The first step of the experiment is to register the tracker with the robot coordinate frame, based on 18 calibrated correspondences. Next, an optical marker that can be detected by the tracker is attached to the torso of the subject, as shown in Fig. 5. In the experiment, the positions of the marker and its approximately corresponding point in the customized model are recorded. As the model is updated at each iteration using the real-time points representing the subject, the performance of the shape model can be evaluated by comparing the spatial variation between the actual motion (from tracker) and the estimated motion (from shape model). In total, we have carried out 3 trials for three anatomical positions (left and right chest and belly) that are mostly affected by breathing.

A second experiment is performed for evaluating the accuracy and robustness of the overall navigation system. For this, a series of robot navigation tasks are considered. For the creation of our customized model, a torso phantom is used as shown in Fig. 6. A layer of plasticine is applied on the torso. Using the methods discussed earlier, a customized model of the torso phantom is created and a trajectory to an anatomical position on the torso is generated. The robotically controlled probe is then navigated according to the generated trajectory. At the end of the executed trajectory, the minimum distance between the probe and the surface is recorded and used as an error measure, while the closest point on the torso is marked. If the probe comes in contact with the torso, the contacting point is imprinted on the plasticine layer and recorded. Additionally, the force applied to the probe is recorded by the force sensor. For each position on the torso, three trials are carried out for different poses of the phantom. The Polaris Vicra system is used to measure the orientation between the poses. The poses used for the experiment are: normal lie low, 12 degrees deviated to left and 14 degrees deviated to right. In accordance with the FAST examination, a total of three anatomical positions (right abdomen, pericardium and pelvis) are tested and the results are recorded for later analysis.

V. Experimental Results

A. VSM validation

A right chest example is shown in Fig. 7 where the position data from the optical tracker and the ones from the corresponding shape model are manually aligned in time. The baselines of the two datasets are aligned according to their means. By comparing the position of the optical marker and the corresponding point on the shape model, we found that the shape model is able to track all respiration cycles along the z-axis perpendicular to the table, with only a few millimeters of error on average. The results for the left chest are similar to the right chest while the shape model is less sensitive to variation on the belly. As shown in Fig. 8, the true amplitude of variation on the z-axis for the belly is about 20mm while only about 6mm can be captured by the shape model. In addition, the shape model is not very capable in capturing variation on the plane parallel to the table, as the signal to noise ratio on the x and y axes is very small. There are a number of reasons that could explain this asymmetry of our system. First, only a limited number of shapes were available for training and the modes of variation exhibited by the shapes are not sufficient. Second, the main variations expressed by the shape model are primarily on the z-axis, as shown in the Fig. 4. Therefore, any deformation occurring on x and y axes will be constrained by the shape model. Lastly, the shape-updating strategy using normals may exhibit preference in selecting correspondences in the z direction, since the point normals are usually pointing in the z axis. The noise on x-y plane may be the result of the iterative shape registration which can be slightly jittery. To minimize this problem, we can set a threshold for the transformation at the ICP step for every iteration. Transformations between adjacent frames that are lower than the a threshold can be ignored.

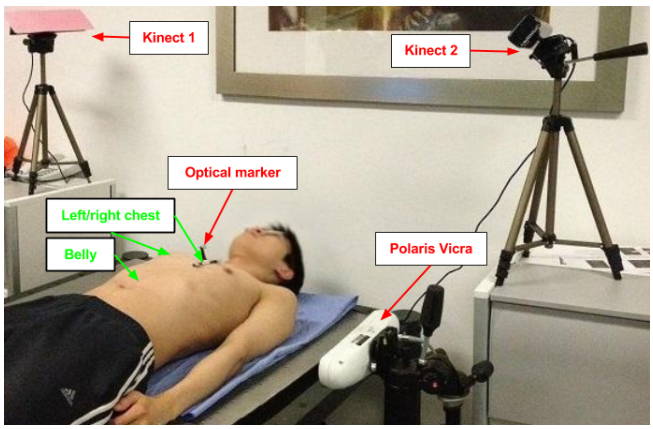


Fig. 5: Experimental setup for validation of the shape model. An optical marker attached to the subject is detected by the Polaris Vicra system. Two Kinects generate point cloud of the subject and a customized model is created correspondingly.

Since the data were not collected at the same time, manual alignment of the two datasets is required. To evaluate accuracy of the respiratory compliance of the shape model, we compare the amplitude in the data between shape model and tracker. The amplitude is defined as the difference of

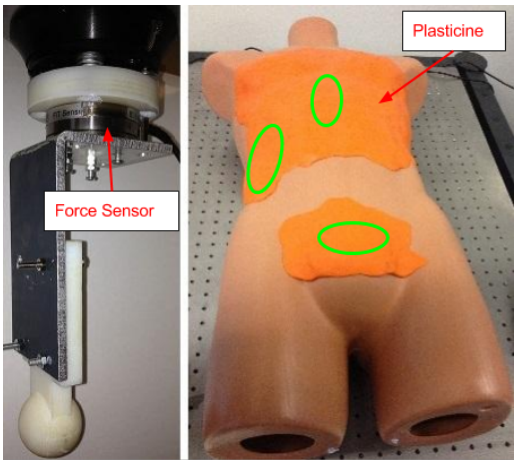


Fig. 6: The image shows the equipment used for validation. A 6-DOF force sensor is attached between the probe and the end-effecter of the robot. A layer of plasticine is applied on the torso to avoid hard collisions of the robot with the hard-plastic phantom and also for validation purposes. The green circles indicate the target areas of FAST scan for the validation.

two adjacent extreme points in the dataset, corresponding to the fully exhaling or fully inhaling breathing motions. Fig. 9 shows the amplitude on the right chest for 9 respiratory cycles. As can be seen from the figure, the overall trend of the amplitude derived by the shape model is correct. The shape model can fairly accurately track the phase and amplitude of the breathing cycle, with a maximum amplitude error of 11.4mm. Overall, the shape model amplitude is lower than the actual one which may be caused by the constrained parameters of the shape model. It is also worth noting that the depth error reported by the Kinect is around 8mm for our setup [16].

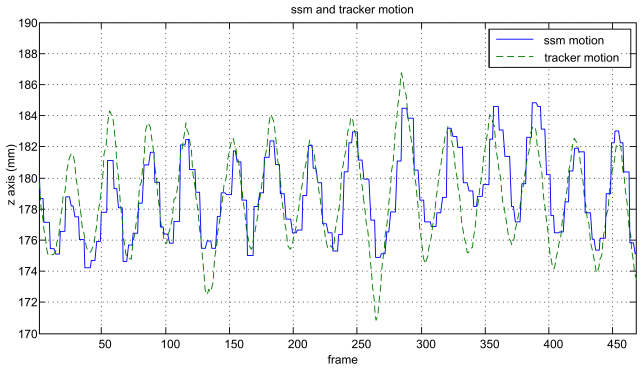


Fig. 7: Comparison of the motion in the z-axis of the optical marker on the subjects right chest and the motion captured by the shape model. The baselines of two datasets are aligned for better visualization.

B. Overall System Validation

For all trials, the final position of the probe tip was close to the surface of the phantom. The error is the distance between the tip of the probe and the surface and ranges from 0 to 2.7cm as shown in Table II. The corresponding surface point is marked. For same target positions, the relative error indicates the displacement between the marked points produced by the first and the other two poses. The relative error shows the stability of the system in different pose configurations. Based on visual observation of the model fitting,

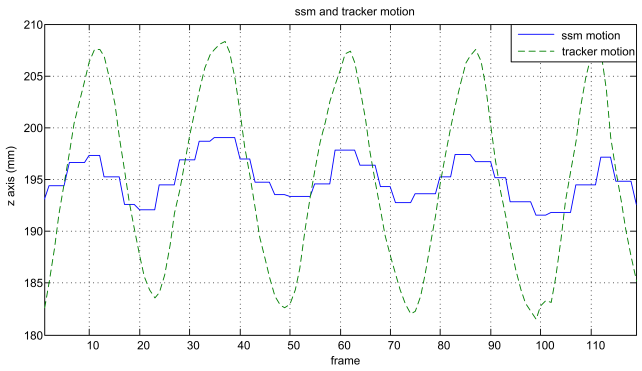


Fig. 8: Comparison of the z-axis motion of the optical marker and the shape model, demonstrating the lower sensitivity of the model on the belly.

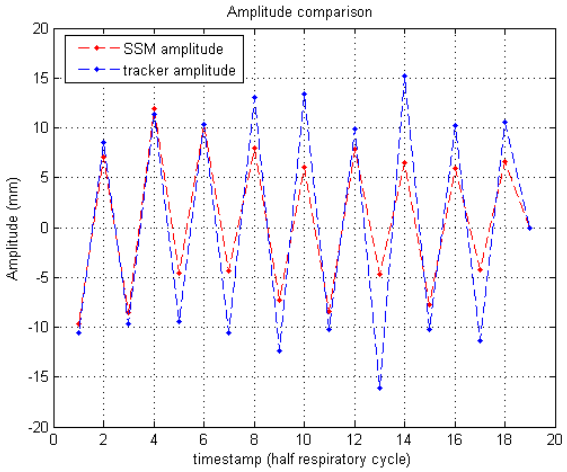


Fig. 9: Comparison of amplitude of z-axis motion on right chest for 9 cycles.

the customized model cannot perfectly adapt to the point cloud in some local areas which results in different variance of error at different anatomical positions. A maximum error of 2.7cm is still acceptable for the sonographer to initiate the examination.

The probe touched the surface of the pelvis area for all poses of the phantom, which resulted in an exerted force of around 10N. It is worth noting that even when plasticine is used, the surface is still harder than a human subject where the force would be lower and within reasonable safety limits.

Throughout the experiments, the occlusion problem is successfully addressed due to the use of a dual Kinect setup. The customized model keeps been correctly updated while the robot arm is moving across the body targets. For more results on the customized model and trajectory, please refer to the supplementary video of this paper.

VI. DISCUSSION AND CONCLUSIONS

The main objective of the proposed framework is to create a patient specific model which can adaptively fit to a real-time updated point cloud of the patient. The model is used as a guide for trajectory planning during robot assisted tele-echography. The experimental results show that the framework generates a customized model able to accurately fit the subjects body. The current performance is approximately

TABLE II: System Validation Results

Trial		Error (cm)	Force (N)	Relative Error (cm)
Right abdomen	***Pose 1***	2	N/A	0
	Pose 2	1.1	N/A	0.3
	Pose 3	1.2	N/A	0.8
Pericardial	***Pose 1***	2.7	N/A	0
	Pose 2	1.5	N/A	1
	Pose 3	2.5	N/A	1.2
Pelvis	***Pose 1***	0	8.1	0
	Pose 2	0	1.2	3
	Pose 3	0	11.3	1.5

0.2 seconds for each iteration while the ideal real-time performance should be around 10 to 15 Hz in order to capture smooth motion of the patient. Although the shape fitting process is not taking place in real-time yet, the model is able to track respiratory motion of the subject with relatively high spatial accuracy.

The proposed system is promising but there are a number of limitations to be overcome before it can be adopted in practice. To address the mentioned real-time problem, we need to parallelize the processes at both slave and master sides. For example, the point cloud received by the master only has an update rate of 7 Hz which is mainly the result of the sequential processing of the point cloud at the slave side. Moreover, the process of correspondence search and update during model fitting is also time-consuming for serial computation. The reason why the data need to be processed at the master side is that the user at the master side needs to be able to overlook and approve the result of the registration and shape model fitting for safety. The second reason is that processing at the slave site is not desirable for the sake of keeping the remote setup as basic and low cost as possible. This is important when multiple remote sites are required (i.e., a number of ambulances).

The current model fitting algorithm can only deal with whole body movement as whole human shapes were used to build the shape model. Therefore, if the target shape is very different from the training set (e.g., people in different poses, very young subjects, people with disabilities), the shape model may fail to register correctly. To address this issue, we can introduce a dividable human model which represents the body as a set of articulated body parts, such as in [9]. For each body part, we can train a deformable shape model based on the statistical shape model. During registration, each body part is registered with the target point-cloud while constraints that ensure kinematically reasonable postures and the connection to neighbor body part are applied subsequently. Moreover, the skeleton provided by the Kinect can also be used to initialize postures of each body part. The current model may also be improved by including further models of subjects while at full exhalation and full inhalation.

Another deficiency of the system is that it cannot differentiate clothes with wrinkles from skin. Although the echography requires direct contact with the patients skin, the rest of the torso can be partially covered by clothes. To address this problem to some extent, computer vision techniques could be exploited to differentiate between clothing and skin.

To our knowledge this is the first study that attempts to address semi-autonomous navigation of a robot for performing a remote ultrasound examination and repositioning of the robot without direct human intervention. It is a preliminary study that represents a paradigm shift towards automation in ultrasound examination and further validation is required.

ACKNOWLEDGMENTS

The authors would like to thank Yang Hu for his assistance and discussion during the system implementation and Konrad Leibrandt for helping with the experimental setup.

REFERENCES

[1] A. Vilchis, J. Troccaz, P. Cinquin, K. Masuda, and F. Pellissier, "A new robot architecture for tele-echography," *Robotics and Automation, IEEE Transactions on*, vol. 19, no. 5, pp. 922–926, 2003.

[2] H. Hashizume, N. Koizumi, S. Warisawa, and M. Mitsuishi, "Automatic control switching according to diagnostic tasks for a remote ultrasound diagnostic system," in *Biomedical Robotics and Biomechatronics, 2006. BioRob 2006. The First IEEE/RAS-EMBS International Conference on*. IEEE, 2006, pp. 1141–1148.

[3] M.-J. Su, H.-M. Ma, C.-I. Ko, W.-C. Chiang, C.-W. Yang, S.-J. Chen, R. Chen, and H.-S. Chen, "Application of tele-ultrasound in emergency medical services," *Telemedicine and e-Health*, vol. 14, no. 8, pp. 816–824, 2008.

[4] F. Conti, J. Park, and O. Khatib, "Interface design and control strategies for a robot assisted ultrasonic examination system," in *Experimental Robotics*. Springer, 2014, pp. 97–113.

[5] H. Monnich, P. Nicolai, T. Beyl, J. Raczkowsky, and H. Worn, "A supervision system for the intuitive usage of a telemanipulated surgical robotic setup," in *Robotics and Biomimetics (ROBIO), 2011 IEEE International Conference on*. IEEE, 2011, pp. 449–454.

[6] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models-their training and application," *Computer vision and image understanding*, vol. 61, no. 1, pp. 38–59, 1995.

[7] T. Heimann and H.-P. Meinzer, "Statistical shape models for 3d medical image segmentation: A review," *Medical image analysis*, vol. 13, no. 4, pp. 543–563, 2009.

[8] A. Vilchis Gonzales, P. Cinquin, J. Troccaz, A. Guerraz, B. Hennion, F. Pellissier, P. Thorel, F. Courreges, A. Gourdon, G. Poisson *et al.*, "Ter: a system for robotic tele-echography," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2001*. Springer, 2001, pp. 326–334.

[9] G. Mylonas, P. Giataganas, M. Chaudery, V. Vitiello, A. Darzi, and G.-Z. Yang, "Autonomous efast ultrasound scanning by a robotic manipulator using learning from demonstrations," in *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*. IEEE, Nov 2013, pp. 3251–3256.

[10] J. Kammerl, N. Blodow, R. B. Rusu, S. Gedikli, M. Beetz, and E. Steinbach, "Real-time compression of point cloud streams," in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*. IEEE, 2012, pp. 778–785.

[11] P. J. Besl and N. D. McKay, "Method for registration of 3-d shapes," in *Robotics-DL tentative*. International Society for Optics and Photonics, 1992, pp. 586–606.

[12] S.-L. Lee, K. Ali, A. Brizzi, J. Keegan, Y. Hao, and G.-Z. Yang, "A whole body statistical shape model for radio frequency simulation," in *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*. IEEE, 2011, pp. 7143–7146.

[13] C. Goodall, "Procrustes methods in the statistical analysis of shape," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 285–339, 1991.

[14] Y. Bosheng, "Implementation of arc interpolation in a robot by using three arbitrary points," *JOURNAL-HUAZHONG UNIVERSITY OF SCIENCE AND TECHNOLOGY NATURE SCIENCE EDITION*, vol. 35, no. 8, p. 5, 2007.

[15] R. B. Rusu and S. Cousins, "3d is here: Point cloud library (pcl)," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1–4.

[16] Y. Liu. (2011) Precision of the kinect sensor. [Online]. Available: http://wiki.ros.org/openni_kinect/kinect_accuracy