Teleoperated Multimodal Robotic Interface for Telemedicine: a Case Study on Remote Auscultation

Sara Falleni[1], Alessandro Filippeschi[1], Emanuele Ruffaldi[1], and Carlo Alberto Avizzano[1]

***Abstract*— The remote examination is becoming more and more important as the population is aging and experts lack as ever before. We propose a novel system which is suitable for remote examination and in particular for remote auscultation. The system is located at two sites, at the patient site a robot holds a stethoscope which is placed on the patient while an RGB-D sensor streams a video of the scene. At the doctor site, the doctor interacts with a haptic interface that allows s/he to move the stethoscope while receiving haptic feedback when the stethoscope is in contact with the patient and looking at the remote scene on a screen. The doctor listens to the noise from the stethoscope thanks to a diaphragm and a headset where the audio stream from the patient site is played. After presenting this novel system, we show its effectiveness by means of experiments that involve auscultation-like tasks. We show the usability of the system to place the stethoscope, the usability to hear correctly the noise of the heart as well as the overall quality of the streamed audio signal.**

I. INTRODUCTION

The increasing aging of population in developed countries will demand very soon to perform more and more medical examinations and interventions. This situation will result in a lack of physicians that is already occurring and will grow in the following years. Telemedicine is the best solution to maintain a high-quality health care service in areas with a shortage of specialists and hospitals. In the last few years, the possibility to perform tele-auscultation, especially in addition to ultrasonography to obtain a more accurate medical evaluation, has raised the attention of engineers and physicians.

During auscultation, the doctor listens to the internal sounds of the body using a stethoscope. Auscultation is performed to examine the circulatory and respiratory systems namely heart and breath sounds, as well as the gastrointestinal system namely bowel sounds. Auscultation requires the stethoscope to be placed on different parts of the patient, which can either lay on the bed or be seated. Contact forces of the stethoscope are usually as small as a few Newtons, i.e. what it is required to keep the contact of the instrument with the patient's skin.

In the last few years, many systems have been developed to manage the tele-auscultation question. In particular, in each work ([1], [2], [3]) a digital stethoscope is used to obtain digital audio. Then in both [1] and [3], the audio streaming is supported by the video streaming (with Mixed Reality technology in [3]) while in [2] a short message technology is used to manage the stethoscope position.

The problem has been faced in European project ReMeDi that aims at achieving a complete tele-examination system for the cardio and abdominal ultrasonography (USG), palpation, and auscultation. Our tele-auscultation system was developed in the ReMeDi context and includes state of the art technologies for haptic interfaces, visualization systems, and audio stream.

This paper presents the architecture, the control strategy and the communication system for virtual remote auscultation. Given the aforementioned auscultation requirements, our approach combines visual, haptic and sonorous cues and allows the doctor to freely move the stethoscope over the patient even in difficult sites. A preliminary evaluation of the system is also proposed to check for its usability and to investigate the role of haptic feedback in teleauscultation. To perform an examination as truthful as possible and to facilitate the positioning of the stethoscope, a force feedback has been implemented using a model mediated technique. In this way every time the stethoscope is in contact with the human body, the doctor has the feeling of pressing on the skin. Finally, a best-effort-communication is developed to let the doctor hearing lounge, heart, and bowel sound directly from patient's side.

After an overall presentation of the system, the components will be detailed. A particular focus will be given to the possibility to perform tele-auscultation with force feedback. The selected model-mediated paradigm allows a transparent teleoperation also in presence of unpredictable delays. Then a demonstrative experiment will be presented to show the effectiveness of the system.

II. STATE OF THE ART

Starting from the possibility to exchange video, audio and data streaming, telediagnosis systems are now focusing on two important aspects. On one side the important thing is that the doctor is able to perform a remote diagnosis without any specialize help like in [4] where the sonography probe is held on the patient using a portable robotic arm held in turn by a non-specialized assistant. In this case, the probe mimics the movements made by the sonographer's hand on a dummy probe at the expert center. On the other side the examination must be as similar as possible to the real one and the work described in [5] is an example of this concept. Indeed in [5] the importance of adding haptic feedback to a Virtual Reality system is apparent for a needle insertion or a palpation task. Analogously, haptic feedback

*This work was supported by the ReMeDi European Project, grant number 610902

[1]Authors are with Percro Laboratory, TeCIP Institute, Scuola Superiore Sant'Anna, 56121 Pisa, Italy n.surname@santannapisa.it

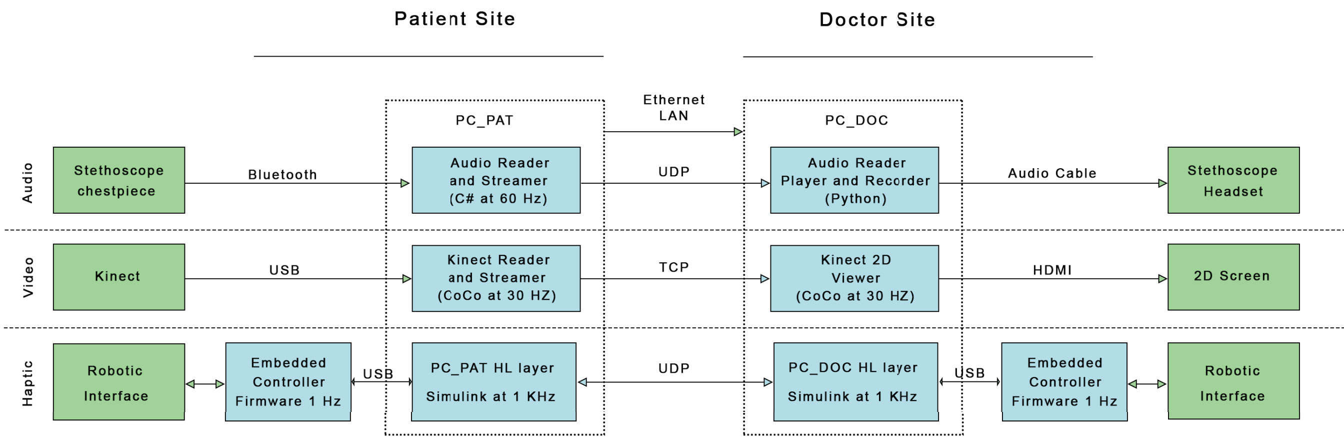

Fig. 1. The system architecture. Green blocks refer to hardware components, whereas blue ones are software components. The figure defines the high-level controllers at the doctor and patient sites. All elements inside blue dashed lines are run in either PC-pat or PC-doc. All communications over Ethernet are UDP communications. Audio, Video and Haptic layers are separated as highlighted by dashed lines

is crucial for palpation in [6] and in [7] for ultrasonography. Regarding the teleauscultation, the two important features are the audio-video real-time streaming and the correctness of the stethoscope positioning. While the audio-video streaming is the starting point, in [2] a Short Messaging Service is used to send the doctor the stethoscope position while in [3] a Mixed Reality system is used allowing the doctor to indicate the desired positioning of the stethoscope. In both Doctor and Patience side, the scene is seen, and some pressure sensors are used to identify the pressure the assistant is exerting on the patient.

Taking a cue from both telediagnosis and teleauscultation systems, we developed a multimodal system which allows three sensory channels to be used at a time. In fact, the system features audio streaming from an electronic stethoscope, a visual feedback realized with a Virtual Reality system like previous teleauscultation systems. Differently from other teleauscultation systems, the one proposed here provides the doctor with haptic feedback. Indeed, this system features a 3-Dof robot at the patient site that is teleoperated by the doctor who acts on a 3 DoFs haptic interface. This solution enriches the doctor's experience and allows s/he to evaluate the contact of the stethoscope on the patient's skin without the mediation of an assistant or the patient her/himself. In this way, the doctor can perform a more realistic exam having the feeling of pressing the stethoscope

III. The Teleauscultation system

A. *Overview of the System*

This teleauscultation system operates between two locations, namely the *patient site* where the patient is examined, and the *doctor site* where the doctor carries out the examination.

At the patient site, a 3 Degrees of Freedom (DoFs) robotic interface [8] holds a purposely designed end-effector aimed at placing a Stethoscope on the patient skin. An RGB-D sensor is positioned aside over the haptic interface to gather and transmit images of the scene where the examination takes place. The haptic interface is equipped with an embedded controller [9] that allows for accurate force rendering. The embedded controller communicates via USB to a PC (PC-pat). The PC-pat performs audio-video acquisition, compression, and streaming.

A local network based on Ethernet connection allows the PC-pat at the patient site to communicate with a second PC (PC-doc) located at the doctor's site. At this site, a 3 DoF delta-kinematics haptic interface equipped with an embedded controller [9] allows the doctor to drive the haptic interface at the patient site thus being able to place the stethoscope on the patient's skin. At the same time, the doctor is provided with a 2D stream of the video that is being recorded at the patient site and s/he can hear the audio stream from the stethoscope.

The architecture of the whole system along with details of the components are shown in figure 1.

B. *Patient Site*

1) Hardware components: At the patient site, the patient lays supine on a table. The robotic interface is placed behind her/him (see figure 2). The robotic interface is composed of a 3DoFs fully actuated robot that allows to position its end point in any location of the patient's trunk. This interface has a spherical RRP (rotational-rotational-prismatic joints) kinematics. The workspace of the interface is the subtraction of two sphere sectors which and it is parametrized by azimuth angle $q_1 \in [-20, 20]^\circ$, inclination $q_2 \in [-40, 40]^\circ$ and radius $q_3 \in [0.4, 0.8]$m. In the worst-case configuration (i.e. the farthest from the RR center) end-effector position resolution is 0.13 mm whereas the maximum continuous and peak forces are 4N and 10N. The robotic interface is managed by a custom embedded controller which features

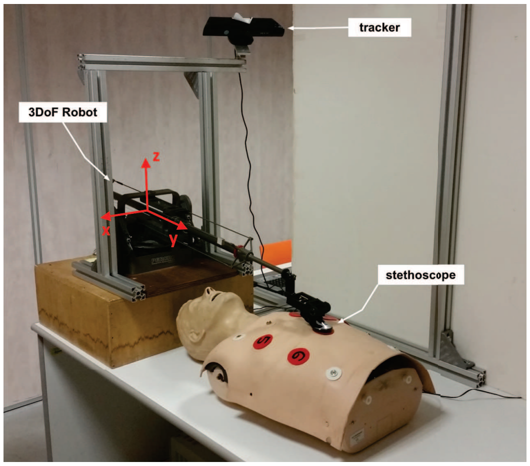

Fig. 2. The patient site. The robotic interface holds a stethoscope on the patient (in this case a mannequin). At the same time a RGB-D sensor tracks the scene.

an STM® Discovery board based on an ARM Cortex-M4 STM32F407VGT6 microcontroller. The board drives via PWM three H-bridge based drivers (Pololu® VNH5019) featuring current sense. The board also features the interface for the three optical encoders attached to each motor shaft and the serial communication with the PC-pat.

The end-effector was designed purposely for the auscultation task and it serves to hold and orient the stethoscope during the examination. The end effector is composed of three parts (see figure 3). The first allows for an offset of the stethoscope with respect to the end point, whereas parts two and three make the stethoscope have two rotational DoFs with respect to the robotic interface using two hinge joints. These joints have low friction thanks to ball bearings and their axis intersect at the point P_S that is 15 mm above the center P_C of the surface of the stethoscope which is in contact with the patient. This mechanical solution makes the force F_c generated during the contact with the patient align the stethoscope to the skin without the need for actuating these DoFs. In facts the torques due to F_c on the two passive rotational joints make the stethoscope rotate until F_c direction includes point P_S. When the force is aligned, the 3 actuated DoFs of the robotic interface allow exerting any force on the patient. This solution works in practice only if the stethoscope motion is kept limited around a neutral pose (see figure 3). Therefore, a spring was added between parts 2 and 3 to limit the motion of the stethoscope during non-contact phases. All parts of the end-effector were realized via 3D printing in ABSplus-P430 (Stratasys®) plastic.

A Littmann® Electronic Stethoscope Model 3200 was selected for auscultation because it features online streaming of the acquired audio via Bluetooth. The chestpiece (diameter 5.1) was separated from the diaphragm and the headset to be mounted on the end effector. A rubber and foam layers were added as interfaces between the end effector and the chestpiece to minimize noise due to the motion of the robotic interface. The chestpiece acquires 16bit mono audio at 4000Hz and sends it via Bluetooth to the PC-pat in packets of 16ms (60Hz). Then the PC-pat streams them over the network via UDP.

The RGB-D sensor is placed over the robotic interface 30 aside from the robotic interface so as to minimize occlusions of the scene due to the robotic arm. Images from the sensor are available at 30Hz frequency and 640x480 resolution. The PC-pat (ASUS PC powered by a quad-core Intel Core i7-3610QM, 8GB of RAM and a GPU Nvidia GeForce GTX 650M running Microsoft Windows 10.) acquires video from the Kinect and streams it over the network using h.264 compression with zero-latency tuning.

2) Software components: The micro of the embedded board runs the firmware of the robotic interface that allows for setting either position, velocity or force control at either joint or end-effector. It also computes and drives joint torques needed for gravity compensation, and it provides the currents measured at each driver.

The PC-pat runs three software components. The first one allows the serial communication at 3Mbit/s over highspeed USB to the embedded controller and runs the teleoperation loop. This software is implemented as a compiled Matlab® Simulink model that runs at 1 kHz frequency. From one hand it receives data from the doctor site via UDP and set the target to the embedded controller. This target is modified according to the experimental condition, i.e. the robotic interface can be given either force or position targets. Given that the robotic interface is gravity compensated and that friction and gravity contributions to the joint torques are negligible during auscultation, available current sense allows us to estimate the force that is applied to the patient. In particular, it allows for detection of contact phases as well as estimation of the force direction. Given the aforementioned properties of the end-effector this is also an estimation of the normal to the surface of the patient skin (i.e. the surface of the chestpiece). This normal is sent over the network along with the robotic interface force and position at the end-effector.

The video acquisition, streaming and playback is managed by the Compact Components (CoCo) framework[1] for Mixed Reality [10], an Open Source C++-based software that has been designed around the concept of data-flow and multicore execution flexibility. While for the current setup an audio-video streaming software would suffice, CoCo has been chosen for the future use of the depth channel from the Kinect camera and alternative viewing modalities such as Head Mounted Displays.

C. Doctor Site

1) Hardware components: The doctor acts on a parallel haptic interface that is composed of a custom Delta-like (Delta.3, Force Dimension, Nyon, CH) haptic device and an end-effector shaped as the chestpiece of the stethoscope (see figure 4). The interface allows for a 3 degrees of freedom (DoFs) translational motion of the end-effector; the workspace is included in a cylinder whose diameter and

[1] https://github.com/cocomr/coco

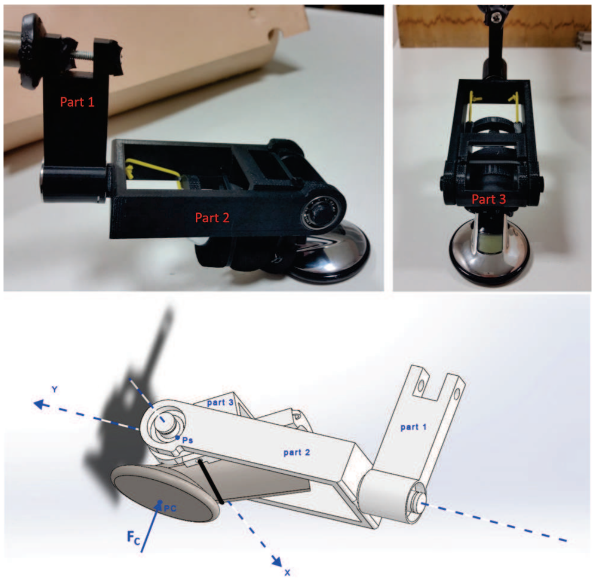

Fig. 3. The attachment of the stethoscope. The three parts are visible. The perspective allow to note the rotation axes intersect over the stethoscope's chestpiece. The yellow rubber recalls the device to the neutral pose as in the figure.

height are 0.26m and 0.12m respectively. The force that the device can display is 40 N in each direction within the aforementioned workspace. A joystick is also available to give further input to the system by means of two buttons. The haptic interface is driven by a copy of the embedded controller which drives the robotic interface at patient site. In front of the haptic interface the doctor can watch the video stream of the scene at the patient site. Currently the video is shown in 2D on a 25" LCD screen, but it is also possible to display the scene at the patient site in 3D by means of a head mounted display. Finally, a speaker is fixed to the diaphragm of the stethoscope (see figure 4) that is used as output device for the doctor to hear the sound at the patient site. The screen and the haptic interface are oriented so as to match the reference frame attached to the robotic interface at the patient site. This frame is depicted in figure 4. The high-level control of the haptic interface, the audio stream output and the rendering of the VE are executed on the PC-doc, a Dell Alienware x51 computer powered by a quad-core Intel Core i7-4790K, 16GB of RAM and a GPU Nvidia GeForce GTX 670 running Microsoft Windows 10.

2) Software components: The PC-doc runs three software components. Similarly to the patient site, the first one manages the serial communication via highspeed USB to the embedded controller and runs the teleoperation loop. Also this component is implemented as a Matlab® Simulink model that runs at 1 kHz frequency. From one hand it receives data from the patient site via UDP and set the target to the embedded controller. Moreover, this Simulink model includes a set of blocks that allow the experimenter to switch on/off the haptic feedback, to label the experimental condition and the trial, so that it is possible to trace the execution of an experimental protocol. A Matlab graphical

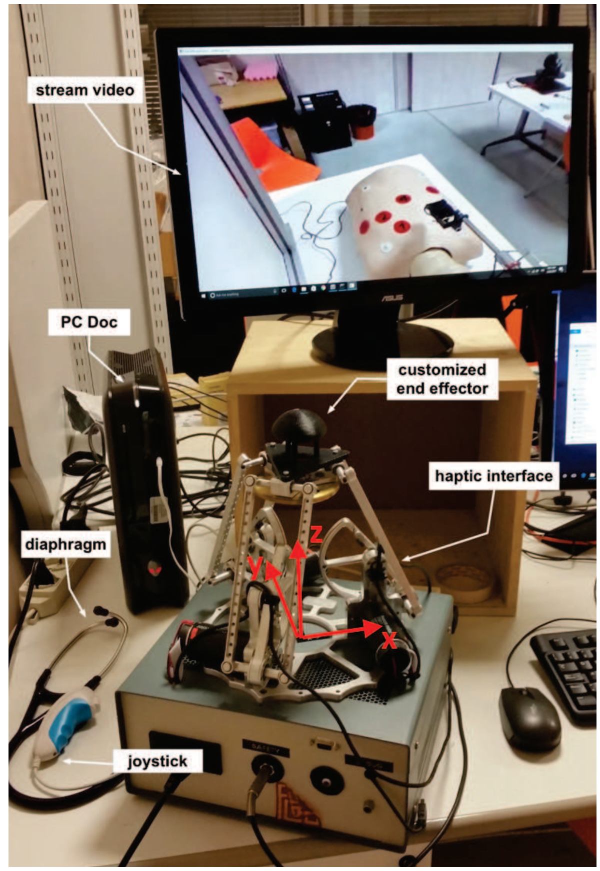

Fig. 4. The doctor site. The doctor holds the customized end-effector to position the stethoscope at the patient site. The haptic interface and the video stream make this task possible. The diaphragm allows the doctor too listen to the audio stream from the patient site. The joystick is sued as an interface for the user to give an input. Its use is described in the following.

user interface allows the experimenter to interact with this model to ease the process of managing the experimental protocol.

To implement the audio stream the PC-doc listens for UDP packets to arrive on a port then packets are sent to speakers.

D. Teleoperation and force rendering for auscultation

A teleoperation loop occurs between the haptic interface at the doctor site and the robotic interface at the patient site. The doctor site device is the master of the teleoperation. Once the experimenter switches on the master device, both the haptic and the robotic interface calibrates, then the slave interface waits for the master to be commanded. The position p_D of the end effector of the haptic interface at the doctor site is sent via UDP to the slave's Simulink model where a position control of the robotic interface is set. A 1.5 scaling factor is applied to the haptic interface position to make the workspace available at such interface allow the stethoscope to span the whole trunk of the patient.

When the stethoscope target position is inside the patient's body a collision occurs and the position control causes a force F_P to be exerted by the robotic interface on the patient. For the safety of the patient, the force F_P is saturated at 3 N. This value is much smaller than the force limit (210 N) reported in the ISO norm 15066 [11] for collaborative robots. However, the saturation is geometrical, i.e. the direction n_P of F_P is not affected by saturation. Given that the contact surface is planar and that F_P is almost normal to this plane (usually the user's approach to the patient makes friction be negligible) n_P is an estimation of the normal to the patient's surface at the contact point. Both n_P and F_P are sent via UDP to the doctor site to manage force feedback.

At the doctor site the doctor perceives a force feedback which is calculated according to the following model. When F_P magnitude is under $1.5\,\mathrm{N}$ no force is displayed at the doctor, whereas when F_P exceeds the threshold the point where it occurred is recorded and it is used as proxy p_0 on the patient skin surface. p_0 is updated every time a new contact of the stethoscope with the patient is identified. Normal vector n_p allows us to define a plane at p_0 with respect to which force rendering is calculated: the patient skin is supposed to be a plane defined by n_p and the body of the patient is locally modeled as a uniform stiffness k_p body. This simple model is suitable until the target position of the stethoscope is near p_0, which usually happens given the force saturation and the stiffness of the rib cage. The force F_D displayed to the doctor is directed along n_p and it is proportional to the indentation, i.e.

$$F_D = k_p||p_D - p_0||n_P \quad (1)$$

where $k_p = 800\,\mathrm{N/m}$. This method is a simple application of the model mediated teleoperation approach that allows us to limit the instability effects due to the communication channel.

IV. Assessment of System Performance

Components were preliminarily assessed to provide the performance they were designed for. In particular, static trials were carried out on both the robotic and the haptic interface to check that positions were correctly estimated and forces correctly rendered. These trials included also test of gravity compensation. As a further step we assessed the latency of the haptic loop, that is we estimated the round trip time of a signal along the path marked with numbers from 1 to 6 in figure 1. The time stamp from the haptic interface embedded controller (i.e. STM Discovery timer) was propagated to the PC-doc high level Simulink model (arrow 1), from here to the PC-pat one (arrow 2) and to the haptic interface embedded controller (arrow 3). From here the signal followed the opposite path to land in the haptic interface embedded controller (arrows 4,5 and 6). Here this old time stamp is subtracted to the current one, this difference being an estimation of the system's latency. From 6 minutes trials we obtained an average latency of $10.64\,\mathrm{ms}$ with standard deviation 6.93,ms and peaks of $75\,\mathrm{ms}$.

A. Preliminary Usability Assessment

A preliminary usability assessment was set up with inexperienced participants to assess whether they were able to carry out basic tasks needed for auscultation. Ten healthy participants (aged 27.1 ± 2.51 years) participated in the experiment after being instructed about the experiment and after providing informed consent to record and use data from the experiment. The material of the experiment is the teleauscultation system where participants play the role of the doctor whereas a mannequin replaces a human patient. They were asked to carry out three tasks.

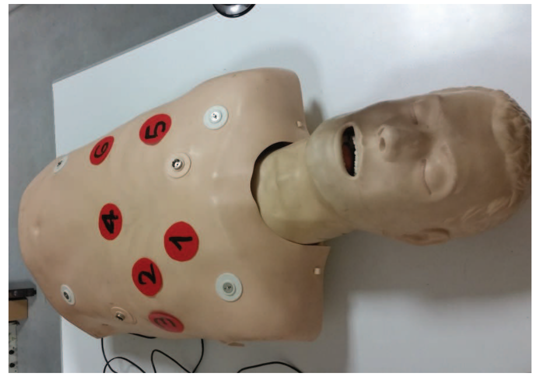

Fig. 5. The six targets attached to the mannequin.

The first task consists of placing the chestpiece of the stethoscope on six targets placed on the patient's trunk as accurately as possible. The targets are as big as the flat surface of the chestpiece so that hiding a target could be a successful strategy to accomplish the task. These targets were selected among anterior and lateral locations on the chest to stress the capabilities of the system (see figure 5 for the location of the 6 targets) The stethoscope is moved by means of the haptic interface at the doctor site while looking at the screen where the video from the patient site is shown. The participant tells the experimenter when s/he believes to have aligned the stethoscope to the target. This task is carried out alternately with or without haptic feedback for a total of six trials. The correct positions of the six targets are recorded at the beginning of the experiment by manually placing the stethoscope on the target. A metric of error e is defined as the Euclidean distance between the correct position and the stethoscope position claimed by the participant.

The second task consists of placing the stethoscope on target 4 and listen to the audio stream from the chestpiece by means of the headset. These tracks which contain heartbeats sound at 27, 55 and 115 beats per minute (bpm) are played by means of a speaker under the mannequin chest. Participants have to estimate the heart frequency they hear. To do that they are provided with a joystick and they can push a button synchronizing with the heartbeat. On a second screen they have a visual cue of the times per minute they are pushing the button. They can use this number as an aid for the estimation of the heard heart frequency. They repeat this task alternately with and without haptic feedback three times for a total of six trials. A metric of error HR_e is defined as the absolute value of the difference between the correct frequency and the one stated by the participant.

The last task consists of listening to ten pairs of tracks by means of the stethoscope's headset. Participants are asked to rate which of the two tracks the judge better in terms of

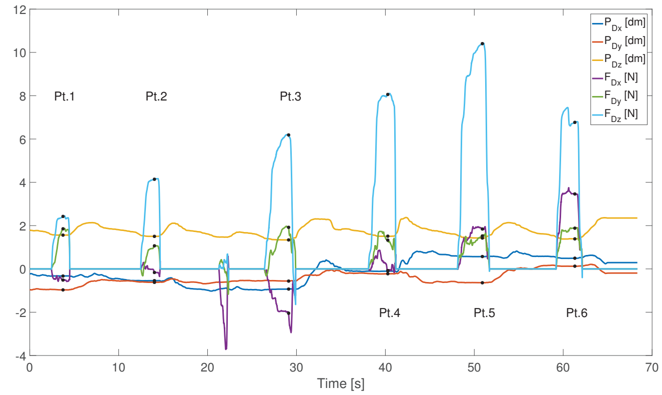

Fig. 6. Haptic interface positions and forces over time during task 1. Positions are in decimeters to have the same numerical scale of the force plot.

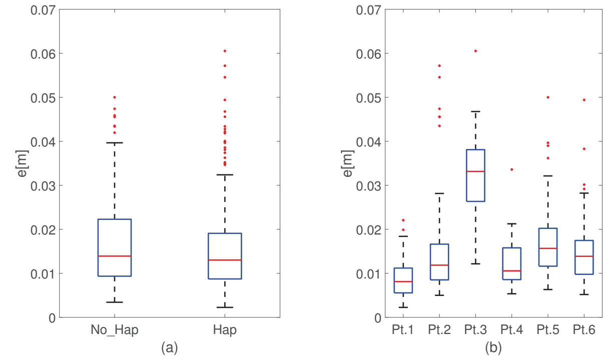

Fig. 7. Error e during task 1 grouped by (a) haptic condition, (b) target point. Boxplots report 25(th) and 75(th) percentiles in blue, median in red. Whiskers are at 1.5 times interquartiles q1 and q3. Crosses are outliers.

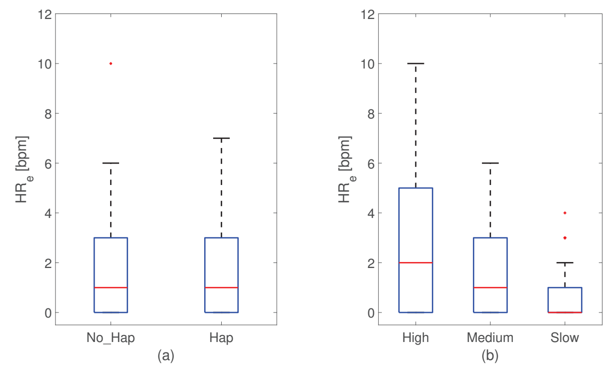

Fig. 8. Error HR_e during task 2 grouped by (a) haptic condition, (b) heart rate. Boxplots report 25(th) and 75(th) percentiles in blue, median in red. Whiskers are at 1.5 times interquartiles q1 and q3. Crosses are outliers.

absence of noise and of overall sound quality. Participants use a 5 points Likert scale to give their preference, one meant they believe track one is totally preferable to track 2, three mean no preference whereas 5 means total preference towards track 2. The first track of the pair was recorded on a human healthy volunteer by manually placing the stethoscope on his chest. The second track contains a second recording of the same person on the same point of the first track. However, this time the stethoscope was place on that point by means of the teleauscultation system. Participants did not know what the origin of the recording was.

V. RESULTS

As a first result we report a plot of the p_D and F_D during task 1 (see figure 6). In the figure the behavior of the participant is apparent: the body of the mannequin is explored while moving the stethoscope in the $x - y$ plane until close to a target. Then a decrease of p_{D_z} shows the approach to the target. This pattern is of course not valid for target point 3, which lies on the side of the trunk. In facts, in this case both p_{D_x} and p_{D_z} vary to approach the target. Forces perceived by the participant are as high as $10\,\mathrm{N}$ and they are generally smooth (no filter is applied), thus showing the stability of the whole system. It is interesting to note how the participant felt more comfortable to push the haptic interface after point three. Asterisks show the frame at which the participant claimed to be in the correct position.

Moving to the experiments, we first report results for positioning accuracy. Figure fig:exp1res shows the results of the computed errors grouped by either haptic feedback condition (figure 7 (a)) or target point (figure 7 (b)). The error is generally as small as $1\,\mathrm{cm}$. From the figure, a smaller error is performed when haptic feedback is active. Moreover, there is a clear difficulty in placing the stethoscope on target 3, the one which lies on the lateral part of the trunk. We tested these hypotheses via statistical analysis. We applied a log-transformation to e (which is defined as a distance) and checked for normality of distributions using the Lilliefors test. Then we ran a 2x6 repeated measures ANOVA to test main effects of factors haptic condition and target point as well as their interaction. Mauchly's test of sphericity was passed by both target point and interaction. ANOVA resulted in no significant effect for haptic condition (F(1,9)=0.929, p=0.36, observed power 0.139) and for interaction (F(5,45)=0.715, p=0.615, observed power 0.139). Instead, target point proved to be significant (F(5,45)=24.373, p<0.001). Post hoc test using Sidak showed that target point 3 was the only one to differ significantly from the other, being the error higher than for other points, whereas the differences among the others were not significant.

Experiment two results are reported in figure 8. The computed errors are grouped by either haptic feedback condition (figure 8 (a)) or target point (figure 8 (b)). In general, the error is small (generally less than two bpm). The figure shows that haptic feedback has a minor effect on error, whereas heart rate influences the error made by participants. A Friedmann was carried out (distributions were not normal) to check for the effect of heart rate. The test showed no significant effect of heart rate on the error ($\chi^2(2) = 3.6, p = 0.202$).

We finally report results from task three in figure 9. Results show a preference towards tracks from manually placed stethoscope (56% of the votes). However, the other group was chosen in 29% of the pairs. In the remaining cases (15%) tracks were judged equally good.

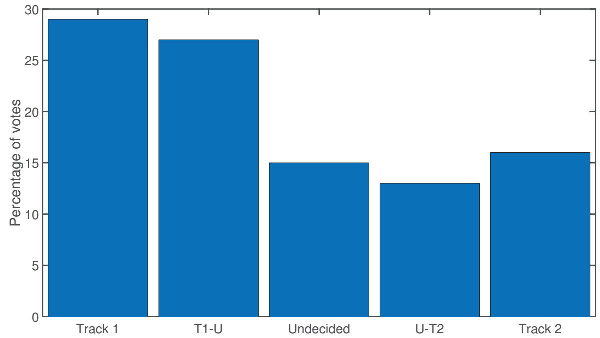

Fig. 9. Haptic interface positions and forces over time during task 1. Positions are in decimeters to have the same numerical scale of the force plot.

We finally report that many participants reported that haptic feedback is useful when added to visual feedback. They also complained about the low resolution that in their opinion could have had an effect on their performance.

A. Discussion

The preliminary assessment as well a visual inspection of position and force plots show that the system allows for a smooth remote manipulation of the stethoscope over and on the patient. Clearly, this results are currently limited to a local network in which latency is sufficiently small to limit stability issues due to the communication channel. However, the proposed approach to teleoperation allows breaking the loop of exchanging forces and positions between the two teleoperation sides so making us confident that this solution could also work with higher latencies.

Although the reported issue with visualization, errors in positioning were rather small. Haptic feedback does not produce a significant effect on error, but the power of the statistical test leads to the need for further experiments to draw conclusions. Target point showed to produce an effect. However, we noted that the only point where the error was significantly larger is located where the robotic interface partially occludes the target. This leads to the need of improving the visualization system not only in terms of resolution bu also as number and location of the sensors.

Heart rate estimation is successfully accomplished regardless the presence of haptic feedback, which is a reasonable consequence of the minor effect of haptic feedback on positioning.

Interestingly, the effect of using the teleauscultation system on the audio quality is not as big as to shift all the preferences towards recordings carried out by manually placing the stethoscope. If the effect of the teleauscultation system on audio quality were strong, then all the preferences should have been directed to track one. However, it has happened only in half of the cases, and in less than one-third of the total votes there was a strong preference. This is promising to move to more challenging tasks such as performing a diagnosis based on lung noises.

As a final remark, we kept the mannequin supine during this preliminary tests because it would have been harder to keep it in a standing pose. However, the robotic interface features enough workspace and payload to allow for carrying out an examination while the patient seats on the table or lies on a side.

VI. Conclusion

We presented a novel system for teleauscultation. The features of the systems were shown and demonstrated through experiments with inexperienced participants. In particular, the system shows to allow positioning the stethoscope and to listen to the noises from the patient site fruitfully. We will exploit the results and the hints that we received to improve the system, especially for what regarding visualization, taking advantage of the Kinect depth information. Further work will be directed to investigate more the role of haptic feedback. Then we will test the system in settings in which latencies are much bigger and unpredictable, requiring a treatment of the synchronization of the multimodal data channels provided by CoCo.

Acknowledgment

This work has been sponsored by the European Commission in the context of the FP7 ReMeDi project (no. 610902). We thank also the project partners.

References

[1] J. C. Bellido, G. De Pietro, and G. Sannino, "A prototype of a real-time solution on mobile devices for heart tele-auscultation," in *Proceedings of the 8th ACM International Conference on PErvasive Technologies Related to Assistive Environments*. ACM, 2015, p. 30.

[2] M. A. Khan, S. F. Jawed, M. O. Khan, and O. Mazhar, "An innovative approach towards e-health in development of tele auscultation system for heart using gsm mobile communication technology," in *Design and Technology in Electronic Packaging (SIITME), 2013 IEEE 19th International Symposium for*. IEEE, 2013, pp. 201–204.

[3] K. Hori, Y. Uchida, T. Kan, M. Minami, C. Naito, T. Kuroda, H. Takahashi, M. Ando, T. Kawamura, N. Kume *et al.*, "Tele-auscultation support system with mixed reality navigation," in *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*. IEEE, 2013, pp. 4646–4649.

[4] P. Arbeille, A. Capri, J. Ayoub *et al.*, "Use of a robotic arm to perform remote abdominal telesonography," *American Journal of Roentgenology*, vol. 188, 2007.

[5] T. R. Coles, N. John, D. A. Gould, and D. G. Caldwell, "Integrating haptics with augmented reality in a femoral palpation and needle insertion training simulation," *Haptics, IEEE Transactions on*, vol. 4, no. 3, 2011.

[6] A. Filippeschi, F. Brizzi, E. Ruffaldi, J. M. Jacinto, and C. A. Avizzano, "Encountered-type haptic interface for virtual interaction with real objects based on implicit surface haptic rendering for remote palpation," in *IROS IEEE*. IEEE, 2015, pp. 5904–5909.

[7] E. Ruffaldi, F. Brizzi, A. Filippeschi, and C. A. Avizzano, "Co-located haptic interaction for virtual usg exploration," in *IEEE EMBC*. IEEE, 2015, pp. 1548–1551.

[8] M. Bergamasco, C. A. Avizzano, A. Frisoli, E. Ruffaldi, and S. Marcheschi, "Design and validation of a complete haptic system for manipulative tasks," *Advanced Robotics*, vol. 20, no. 3, pp. 367–389, 2006.

[9] J. M. J. Villegas, C. A. Avizzano, E. Ruffaldi, and M. Bergamasco, "A low cost open-controller for interactive robotic system," in *IEEE EMS*, 2015, pp. 462–468.

[10] E. Ruffaldi and F. Brizzi, "Coco - a framework for multicore visuo-haptics in mixed reality," in *SALENTO AVR, 3rd International Conference on Augmented Reality, Virtual Reality and Computer Graphics*, L. D. Paolis and A. M. (Eds.), Eds. Springer, 2016, pp. 339–357.

[11] I. ISO, "Ts 15066: 2011: Robots and robotic devices collaborative robots," *Geneva, Switzerland: International Organization for Standardization, draft*, 2011.