

# Stability and Transparency in Mixed Reality Bilateral Human Teleoperation

David Black, Septimiu Salcudean

**Abstract**—Recent work introduced the concept of human teleoperation (HT), where the remote robot typically considered in conventional bilateral teleoperation is replaced by a novice person wearing a mixed reality head mounted display and tracking the motion of a virtual tool controlled by an expert. HT has advantages in cost, complexity, and patient acceptance for telemedicine in low-resource communities or remote locations. However, the stability, transparency, and performance of bilateral HT are unexplored. In this paper, we therefore develop a mathematical model and simulation of the HT system using test data. We then analyze various control architectures with this model and implement them with the HT system to find the achievable performance, investigate stability, and determine the most promising teleoperation scheme in the presence of time delays. We show that instability in HT, while not destructive or dangerous, makes the system impossible to use. However, stable and transparent teleoperation are possible with small time delays ( $< 200$  ms) through 3-channel teleoperation, or with large time delays through model-mediated teleoperation with local pose and force feedback for the novice.

**Index Terms**—Teleoperation, Augmented Reality, Human Computer Interaction, Stability, Transparency, Force Feedback, Haptics

## I. INTRODUCTION

Many remote and underresourced communities experience severe challenges in accessing qualified medical care. For example, ultrasound imaging is important, widely used, and much lower cost than other modalities such as CT or MR. However, capturing and interpreting ultrasound images requires a high degree of expertise that is not commonly present in many small communities. As a result, a sonographer or radiologist must be transported to the town on a regular basis, or patients must be sent to a major medical center. Either case leads to long wait times and difficulty handling urgent cases. In communities across Canada, patients are flown hundreds of kilometers for standard ultrasound exams. This takes up to three days and exerts a high social and financial cost on the community.

Therefore, tele-ultrasound is an important and growing field. However, current commercially available technologies are often impractical. Video teleguidance is simple, low-cost, and accessible to anyone but is highly inefficient and imprecise if the person being guided does not already have ultrasound experience [1]. On the other hand, robotic teleultrasound gives the physician complete and precise control but is expensive and complex to set up and maintain. We thus recently introduced a novel teleguidance method called human teleoperation to address the shortcomings of both existing approaches [1], [2]. This method is also applicable to many other remote guidance applications.

In human teleoperation, a local novice, the “follower”, performs an ultrasound exam on a patient while being guided by a remote operator, the sonographer or radiologist. The follower wears a mixed reality (MR) head-mounted display (HMD) which projects a virtual ultrasound probe into their field of view. The pose of this virtual probe is controlled in real time by the operator, who manipulates a haptic device with an ultrasound probe-shaped end effector. As the operator moves the virtual probe, the follower aligns their real probe to the virtual one and follows it as it moves, thus achieving the operator’s desired motion. The ultrasound image and video from the follower’s HMD are streamed to the operator so they can carry out the procedure. The two sides are also in verbal communication. The tracking of the MR virtual probe is very intuitive, and no prior ultrasound experience is required for the follower. Indeed, the ability of the follower to track the MR input was previously characterized [3], [4], showing tracking lags of 200-300 ms and steady state error of less than 3 mm. Such tight coupling of the operator and follower enable this teleguidance system to be designed and analyzed like a robotic teleoperation system, leading to the name human teleoperation.

While the position and orientation of the ultrasound probe are key during procedures, the applied force is also an important factor. Applying the correct level of force avoids excessive deformation of structures like blood vessels, displaces gas which otherwise obstructs the image, and allows imaging below the ribs, for example. Therefore, the ultrasound probe in human teleoperation is instrumented with force sensing, as previously described [5], [6]. Additionally, the probe pose is tracked visually by the HMD [7] to transform the force to known coordinates. Furthermore, the haptic device of the operator has three actuated joints, allowing it to apply forces to the operator’s hand. In this way, it is possible to render a realistic haptic sense of the patient to the operator.

Haptic feedback is important in teleoperation. In robot-assisted minimally invasive surgery [8] and surgical training [9], haptic feedback has been shown to increase task performance and decrease tissue damage [10], [11]. Similarly, in teleoperation of mobile robots [12], unmanned aerial vehicles [13], and robots for micro-manipulation and assembly [14], haptic feedback has shown utility. In human teleoperation for teleultrasound, there are three primary reasons to use haptic feedback. First, sonographers are accustomed to resting their hand on the patient and exerting relatively large forces up to approximately 20 N [15]. Without this support, the experience feels unrealistic and unintuitive. More importantly, it is very difficult to hold the probe at exactly the height of the patient surface or to make precise motions if the operator’s hand is

floating in space instead of resting on the surface. Finally, the operator uses their sense of touch to perform the exam, for example by palpating tissue or by feeling the ribs so as to image between them. Therefore, haptics is essential in human teleoperation.

From the perspective of teleoperation literature, haptic feedback increases the transparency of the teleoperation. Transparency describes how well the follower and operator sides match each other in terms of force and position (velocity) [16]. In a perfectly transparent system, the velocity and force of the operator and follower (commonly referred to in prior literature as master and slave, respectively) would be exactly equal at all times. Thus, the follower would perfectly track the operator's motion and the operator would feel exactly as if they were touching the remote environment directly. This is a primary goal in teleoperation that has been addressed using many approaches including direct force feedback [17], position feedback [18], combined position and force feedback [19], [20], impedance feedback [21], rate control [22], adaptive control [23], local force feedback at the follower [24], and more.

The other essential objective in teleoperation is stability. Especially in medical applications in which the follower robot interacts with patients, instability in the system can constitute a health hazard. Unfortunately, even a small amount of time delay can destabilize nominally stable bilateral teleoperation systems [25]. Since any remote teleoperation system inevitably includes communication at a distance, leading to delays, this is an important problem that has been studied extensively. Due to the complex, nonlinear nature of teleoperators, system passivity is commonly used to show stability [26]. For example, it has been shown that transmitting wave variables rather than the values themselves guarantees the passivity of the system under arbitrary time delays [27], [28]. However, this can degrade tracking accuracy and transparency. Thus, others have used predictive methods such as Smith predictors [15], [29], [30] and model predictive control [31] to eliminate the time delay by predicting the feedback before it arrives. Similarly, model-mediated control allows instantaneous feedback by keeping a local model of the remote environment on the operator side [32]. As these methods are affected by modeling accuracy, however, others have studied transmitting adaptive combinations of velocity and force [33], and robust control approaches [34]. A particularly successful method called time-domain passivity was proposed by Hannaford and Ryu [35], [36], in which an energy-like quantity of the system is observed and dissipative elements are added to eliminate only the net energy output, thus maintaining passivity with less degradation of performance.

Despite the extensive investigation of robotic teleoperation, no research exists that examines the stability and transparency of human teleoperation. In human teleoperation, the slave robot is replaced by a human, leading to fundamentally different system behavior. In particular, the human follower may be considered passive at all times [37], but their reaction time and imperfect tracking lead to potentially larger delays and inaccurate motions and forces. The follower can intrinsically respond to changing environment conditions and does not require

trajectory planning around joint limits and singularities, nor initial calibration, safe deceleration, or compliant controllers. However, they must be guided effectively according to the capabilities of their vision and perception, without providing too much input so as to cause cognitive overload or rapid fatigue. Thus, human teleoperation is influenced by other factors than robotic teleoperation, but stability and transparency remain paramount. To achieve efficient, performant teleoperation, it is thus important to study the stability and transparency of the human teleoperation system.

To this end, this paper presents an initial application of the concepts of stability and transparency from telerobotics to bilateral human teleoperation. In particular, the following contributions are included:

- A physical model of the human-in-the-loop system is derived and used to create a hybrid matrix formulation of the teleoperation with time delays in Section II-B.
- This model is used with several candidate controllers to investigate their respective transparency and expected performance for human teleoperation in Section II-C.
- In Section II-E, the stability of bilateral human teleoperation is investigated.
- In Section III-D, the robustness of the derived models to parametric and dynamic uncertainty is explored using the structured singular value.
- Finally, limited tests with the real system and volunteer followers were performed to validate the modeling. The setup is described in Sections II-F and II-G, and the results are presented in Section III.

While this paper does not constitute a comprehensive analysis or test of all possible human teleoperation approaches, it lays the foundations both for future practical tests with human volunteers and for further algorithmic development, for example in predictive or robust control schemes.

## II. METHODS

### A. Human Teleoperation

Human teleoperation enables precise, tightly-coupled remote guidance of a person's hand motion by an operator. A block diagram is shown in Fig. 1. The system consists of three primary components: the operator side, follower side, and communication. On the follower side, the follower wears an MR headset, the HoloLens 2 (Microsoft, Redmond, WA), and holds a manual tool that is instrumented with force sensing. Most work so far has focused on remote ultrasound, so the tool is an ultrasound probe with a force sensing shell [5], [6]. The probe also has infrared markers that are tracked by the HoloLens, giving access to the probe's pose in real time [7]. The operator side consists of a monitor on which the ultrasound image and video stream of the follower are displayed, as well as a haptic device (TouchX, 3DSYSTEMS, Rock Hill, SC) that the operator uses to input their desired motion. This motion is executed contemporaneously by a virtual ultrasound probe displayed in the HoloLens 2. The follower aligns their ultrasound probe with the virtual one to match the desired pose of the operator. The measured force

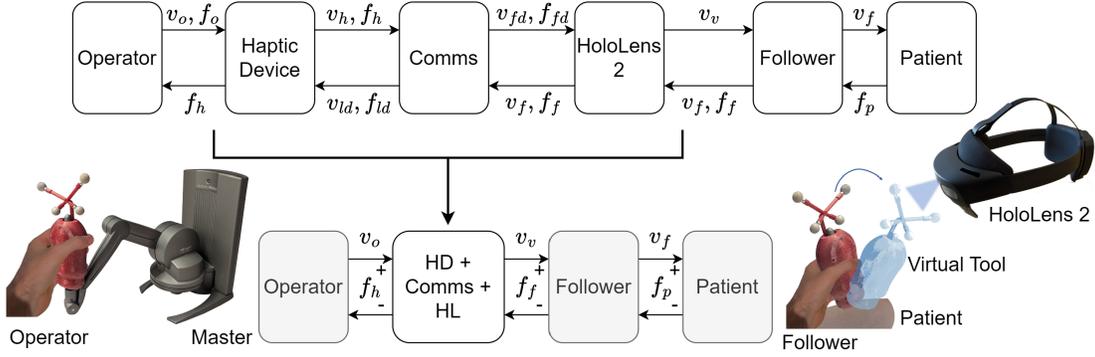


Fig. 1. Conceptual overview of the teleoperation system. The gray boxes of operator, follower, and patient have to be modeled while the haptic device, communications, and HoloLens 2 represent the part of the system we can directly control.

and pose are fed back to the operator, where they can be used to generate forces in the haptic device.

The communication is performed over the Internet through WebRTC, a fast peer-to-peer framework, as described previously [4]. Prior work has also examined the ability of people to track mixed reality inputs, analyzing the tracking error, steady-state error, and lag of their motions and forces compared to the operator's input [3]. Step and frequency response tests were also carried out [4]. The results of these tests are used to inform the mathematical modeling in the following section.

## B. Modeling

A general model of human teleoperation is shown in Fig. 2. We consider a one-dimensional, linear time-invariant (LTI) system which can be extended to three dimensions accordingly. Other than the follower side, the structure is similar to previous work in teleoperation. Indeed, human teleoperation differs from robotic teleoperation primarily in the response of the follower compared to the slave robot. While both follow the commanded motion and/or force and consequently interact with the patient, at a low level the robot is controlled through joint torque commands and must explicitly account for its own dynamics. Conversely, the follower is controlled through position commands by following the virtual tool and implicitly accounts for his/her own arm dynamics. Furthermore, the robot, depending on the controller, can become unstable. On the other hand, the human follower is always stable and can be treated as a passive system [37]. Thus, the stability analysis is simplified. However, the follower's response is slower than most robots' and less accurate, with stochastically varying performance base on focus and other factors, so achieving transparency is more difficult.

Beyond the virtual tool pose, it is possible to give additional input to the follower such as changing the tool color or displaying an error bar, arrow, or visual renderings in the MR headset. These options could be added in the blue  $C_{HL}$  box in Fig. 2. How to model the human response to color or error bars mathematically is, however, an open question. Furthermore, we found previously [3] that further rendering is distracting and leads to cognitive overload for the follower. Hence, this paper examines only the virtual tool pose and other renderings are left for future work.

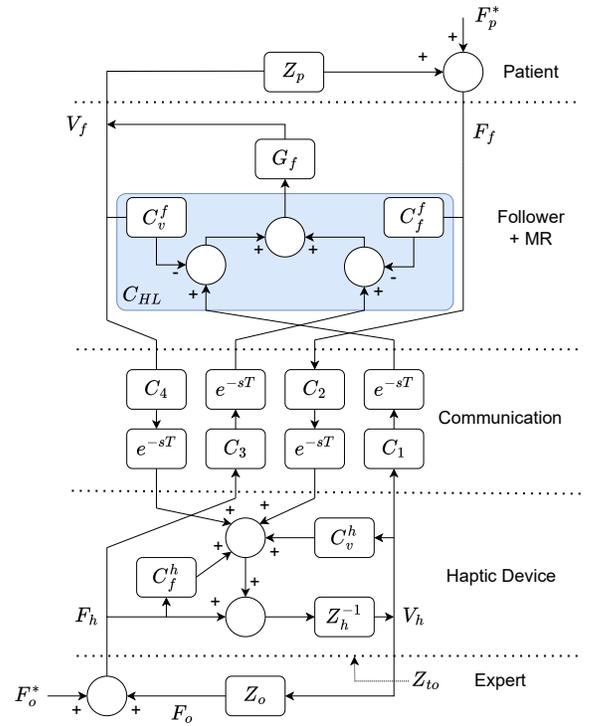


Fig. 2. A general model of human teleoperation. The operator (subscript  $o$ ) interacts with a haptic device (subscript  $h$ ) while the follower ( $f$ ) interacts with the patient ( $p$ ). The communication channel induces time delays of  $T$  on the force and velocity which are transmitted bilaterally. The MR headset creates a visual control output potentially using all four channels of force and velocity, denoted  $C_{HL}$ . The follower responds to the MR input according to the follower transfer function,  $G_f$ . The controllers  $C_j^i$  are on the follower or haptic device respectively with  $i = f$  or  $h$ , and involve force or velocity respectively with  $j = f$  or  $v$ .

Given these assumptions about the human teleoperation system, we develop models of the follower and operator in contact with the patient and haptic device respectively, and study how these can be interconnected in a bilateral teleoperation system.

1) *Follower*: The follower model was derived from measured response data of 11 volunteer followers from previous

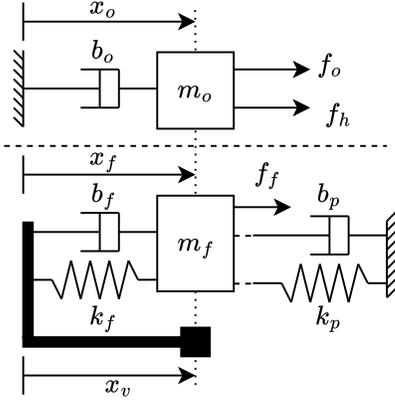


Fig. 3. Models of the operator and haptic device (top), and the follower potentially in contact with the patient (bottom). The damping of the operator arm and haptic device together is represented by  $b_o$ . The black fixture attached to the follower mass is rigid and massless. The square end at position  $x_v$  represents the virtual tool. The follower force  $f_f = -k_p x_f - b_p \dot{x}_f$  cancels the patient dynamics at all times. Thus, when the follower motion has converged (i.e.  $b_f$  and  $k_f$  are at equilibrium), the follower position  $x_f$  should equal  $x_v$ .

studies [3], [4]. It was found that a second-order system with two poles and one zero matched the measured behavior well. This is equivalent to attaching the follower's hand and tool, which have a certain mass, to the virtual tool by a spring and damper, as shown in Fig. 3. The follower either moves freely in space or is in contact with the patient, represented by an impedance  $Z_p = b_p s + k_p$ , which is taken to be constant.

We assume the follower applies whatever force is necessary to match the virtual tool, whether or not they are in contact with the patient, so there is no steady-state error. In other words,  $f_f = -k_p x_f - b_p \dot{x}_f$  at all times, so it perfectly cancels the patient contact dynamics. In this way, intermittent contact with the patient has no effect on the follower's stability or tracking, which is as expected from experience; the human hand does not become unstable, even in high-frequency contact with a stiff surface. The follower can thus be represented by the passive (strictly passive if  $b_f > 0$ ) transfer function

$$x_f = \left( \frac{b_f s + k_f}{m_f s^2 + b_f s + k_f} \right) x_v = G_f x_v \quad (1)$$

$$f_f = -(b_p s + k_p) x_f \quad (2)$$

where  $k_p$ ,  $b_p$  are zero when not in contact with the patient. For this to be LTI, we assume the patient is unmoving and their impedance is constant. This is approximately true when scanning only one region, for example the abdomen. An advantage of this model is that we do not have to consider switching contact, which is handled implicitly by the follower.

2) *Operator and Haptic Device*: The operator hand holding the haptic device is modeled as a mass and damper system which is subject to the applied force of the operator and the haptic device, as shown in Fig. 3.

$$m_o \ddot{x}_o + b_o \dot{x}_o = f_h + f_o \quad (3)$$

The operator changes their applied force depending on the haptic device force, to track a desired trajectory. In particular,

to choose a representative response, we use a proportional-integral-derivative (PID)-inspired controller, choosing

$$f_o = -k_{o1} x_o - k_{o2} \dot{x}_o + k_a \int (x_o^* - x_o) dt \quad (4)$$

where  $x_o^*$  is the operator's desired motion. These gains can be tuned to obtain a response that resembles the recorded operator data. We ignore the operator's desired velocity for simplicity as it introduces a zero eigenvalue and the operator is primarily interested in position. We thus find the following representation of the operator

$$x_o = \frac{f_h + \frac{k_a}{s} x_o^*}{m_o s^2 + (b_o + k_{o2})s + k_{o1} + \frac{k_a}{s}} \quad (5)$$

$$f_o = \frac{k_a}{s} x_o^* - \left( k_{o2} s + k_{o1} + \frac{k_a}{s} \right) x_o \quad (6)$$

The state space formulations of the operator and follower are derived in the Appendix, where we show that the expert controller can be framed as servo control and tuned using pole placement.

3) *Teleoperation System*: The full teleoperation system is shown in Fig. 4, given the models derived above. Let  $y_j = [x_j \ \dot{x}_j \ f_j]^T$   $j = \{o, f\}$  be the output of the operator and follower, respectively. We assume the follower sees only the virtual tool pose and the operator receives feedback only by forces applied to the haptic device. This does not take into account the operator's visual feedback from video or other sensor streams such as ultrasound images, nor the verbal communication between operator and follower. In this case, we control the virtual tool pose,  $x_v$ , for the follower, and the haptic device force,  $f_h$  for the operator. Each can be a function of the position, velocity, and force of the operator and follower as shown in the diagram. Some gains are never used, so they are set to zero. For example, the haptic device is not equipped with a force sensor, so its actual force cannot be fed back. Thus, the general inputs to the operator and follower are

$$f_h = k_{fo} f_f + k_{po}(x_f - x_o) + k_{do}(\dot{x}_f - \dot{x}_o) \quad (7)$$

$$\dot{x}_v = \dot{x}_o + k_{pf}(x_o - x_f) + k_{ff}(f_h - f_f) \quad (8)$$

Let  $K_{oo} = [k_{po} \ k_{do} \ 0]$ ,  $K_{of} = [k_{pf} \ 1 \ k_{ff}]$ ,  $K_{ff} = [k_{pf} \ 0 \ k_{ff}]$ , and  $K_{fo} = [k_{po} \ k_{do} \ k_{fo}]$ . Then in terms of the individual model outputs, these expressions become

$$f_h = K_{fo} y_f - K_{oo} y_o \quad (9)$$

$$\dot{x}_v = K_{ff} y_f - K_{of} y_o \quad (10)$$

For non-zero time delay, the equations are:

$$f_h = e^{-sT} K_{fo} y_f - K_{oo} y_o \quad (11)$$

$$\dot{x}_v = K_{ff} y_f - e^{-sT} K_{of} y_o \quad (12)$$

The full state space notation of the closed-loop teleoperation system is shown in the Appendix.

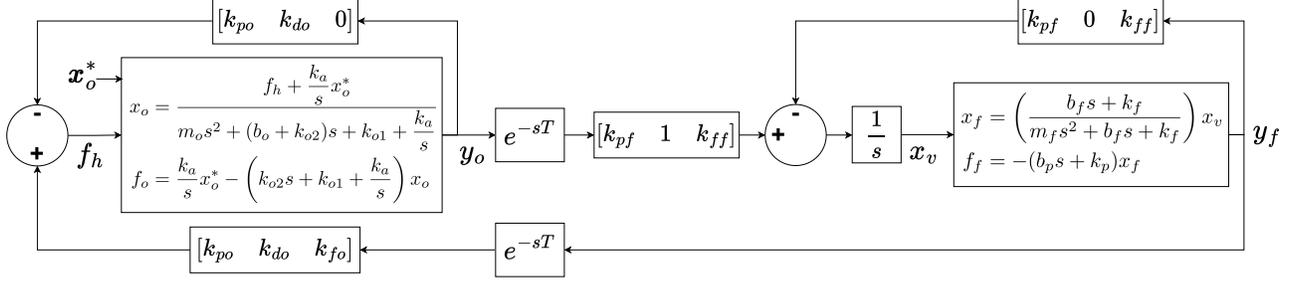


Fig. 4. Interconnection of the derived models in the closed-loop teleoperation system. Different behavior is achieved by changing the feedback and feed-forward gains of the positions, velocities, and forces. The operator-side model is given in Equation 5 while the follower side is in Equation 1.

4) *Simulation:* A Simulink model was developed to simulate the system and test the controllers. This directly implemented the state-space models described in the previous subsections with some practical considerations. The haptic device force was set to saturate at 7 N, approximately the maximum force of the haptic device, and the environment was made approximately as stiff as a human abdomen during ultrasound [38]. Next, the operator and follower parameters were independently estimated to obtain responses similar to previous measurements [3], [4]. Finally, the model and feedback gains were tuned with all components of the teleoperation connected together, to again match the expected performance.

The chosen model parameter values were obtained using the MATLAB linear system identification toolbox and some manual tuning. Specifically, linear grey-box estimation was utilized with Adaptive subspace Gauss-Newton search to fit parameters to the derived ordinary differential equations, with non-negativity constraints and regularization on all parameters, and a stability constraint on the system. The final system fit the measured dataset with a mean-squared error (MSE) of 0.67 mm in the follower motion. The fit accuracy is relatively insensitive to changes in the parameters, which shows robustness in the model. The fit MSE has a partial derivative of 0.125 mm/kg for follower mass with all other parameters fixed at their chosen values. All other parameters have slopes of less than 0.0625 mm/unit in magnitude. The measured and fitted data are shown in Fig. 5.

In this way, approximate numerical values could be assigned to the various patient, follower, and operator parameters. These values and simulated models are useful for testing controllers before implementing them on physical hardware, and for numerically assessing the stability of a controller. An effective set of parameters is shown in Table I.

TABLE I  
APPROXIMATE MODEL PARAMETER VALUES DETERMINED THROUGH SIMULATION AND SYSTEM IDENTIFICATION BASED ON PREVIOUS DATA [3]. ALL UNITS ARE N, M, S, KG.

Patient		Follower		Operator	
$k_p$	10	$k_f$	1	$k_a$	100
$b_p$	1	$b_f$	0.275	$b_o$	0.1
$m_p$	0.02	$m_f$	0.02	$m_o$	0.1
				$k_{o1}$	0
				$k_{o2}$	80

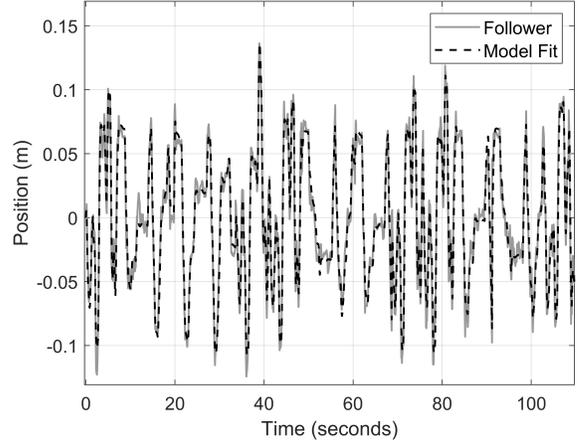


Fig. 5. The derived model with parameters fitted to the measured data from previous tests [3], [4]. The MSE between the two is 0.67 mm.

### C. Transparency

Bilateral teleoperation systems are commonly expressed using a hybrid representation:

$$\begin{bmatrix} f_h \\ -\dot{x}_f \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} \dot{x}_o \\ f_f \end{bmatrix} = H(s) \begin{bmatrix} \dot{x}_o \\ f_f \end{bmatrix} \quad (13)$$

For perfect transparency,  $\dot{x}_o = \dot{x}_f$  and  $f_h = f_f$ , so the ideal hybrid matrix consists of elements  $h_{11} = h_{22} = 0$ ,  $h_{12} = 1$ , and  $h_{21} = -1$ . Equations 11 and 12 together with the operator and follower models can be expressed as a hybrid matrix with the following elements:

$$h_{11} = \frac{e^{-s2T}(k_{po} + k_{do}s)(1 - G_f^{-1})}{G_f^{-1}s + k_{pf} - e^{-s2T}k_{ff}(k_{po} + k_{do}s)} \quad (14)$$

$$h_{12} = \frac{(G_f^{-1}s + k_{pf})k_{fo} - k_{ff}(k_{po} + k_{do}s)}{G_f^{-1}s + k_{pf} - e^{-s2T}k_{ff}(k_{po} + k_{do}s)} e^{-sT} \quad (15)$$

$$h_{21} = -\frac{s + k_{pf} - e^{-sT}k_{ff}(k_{po} + k_{do}s)}{G_f^{-1}s + k_{pf} - e^{-s2T}k_{ff}(k_{po} + k_{do}s)} \quad (16)$$

$$h_{22} = \frac{k_{ff}(1 - e^{-s2T}k_{fo})s}{G_f^{-1}s + k_{pf} - e^{-s2T}k_{ff}(k_{po} + k_{do}s)} \quad (17)$$

Notice that all elements of the  $H$  matrix share the same denominator. Consider the case where time delay approaches

zero ( $T \rightarrow 0$ ) and we use unity-gain force feedback ( $k_{fo} = 1$ ). Then  $h_{22} = 0$  and  $h_{12} = 1$ . The terms  $h_{11}$  and  $h_{21}$  achieve their ideal values of 0 and  $-1$  respectively if and only if  $G_f = 1$ , which would occur if the follower were infinitely fast and perfectly accurate. Of course this is not the case and the transient response will never be exactly unity. Thus, with no time delay and  $k_{fo} = 1$ , the hybrid matrix becomes

$$H(s) = \begin{bmatrix} \frac{(k_{po} + k_{do}s)(1 - G_f^{-1})}{G_f^{-1}s + k_{pf} - k_{ff}(k_{po} + k_{do}s)} & 1 \\ -\frac{s(1 - G_f^{-1})}{G_f^{-1}s + k_{pf} - k_{ff}(k_{po} + k_{do}s)} - 1 & 0 \end{bmatrix} \quad (18)$$

However, at steady-state if there is no remaining error,  $G_f$  tends to 1, so  $h_{11} = 0$  and  $h_{21} = -1$ . Similarly, if the operator velocity is very small, for example when in contact with a patient during an ultrasound,  $v_o \rightarrow 0$ , so transparency is achieved at steady state.

The follower's response,  $G_f$ , cannot be altered, but  $k_{po}$ ,  $k_{do}$ ,  $k_{pf}$ , and  $k_{ff}$  can be tuned to bring  $h_{11}$  closer to 0. Indeed, taking either follower gain,  $k_{pf}$  or  $k_{ff}$ , to infinity drives  $h_{11}$  and  $h_{21}$  to their ideal values. With very large operator gains,  $k_{po}$  and  $k_{do}$ ,  $h_{21} \rightarrow 0$  and  $h_{11} \rightarrow \frac{G_f - 1}{k_{ff}G_f}$ . However, such large gains are prevented by practical considerations such as haptic device joint torque limits and stability. Adding time delay further degrades the transparency.

#### D. Control Architectures

In this subsection, various possible teleoperation architectures based on this model are described. Their stability is analyzed in Section II-E.

1) *2-Channel Teleoperation*: Three 2-channel teleoperation approaches were implemented, differing in which variable is reflected from follower to operator and whether the follower receives local position feedback. In each case, the operator motion is sent to the follower.

*Position-Position* ( $C_2 = C_3 = 0$ ):

In this architecture, the follower position is fed back to the operator, and a force proportional to the position error and its time derivative is applied to the haptic device. This is one of the first bilateral teleoperation schemes to have been developed and has the advantage of not requiring a force sensor at the follower. In this case, however,  $k_{fo} = k_{pf} = k_{ff} = 0$ , which leads to problems with the transparency. In Fig. 2, this is equivalent to setting  $C_2 = C_3 = C_v^f = C_f^f = C_f^h = 0$ . The hybrid matrix becomes

$$H(s) = \begin{bmatrix} e^{-s2T} \left( \frac{k_{po}}{s} + k_{do} \right) (G_f - 1) & 0 \\ -G_f & 0 \end{bmatrix} \quad (19)$$

The obvious problem is that the haptic device force has no relation to the follower force. In practice, the bigger problem is that the follower lags the operator and the returned force is further delayed, so the operator always feels forces resisting their motion which makes operation difficult and frustrating. This is evident in  $h_{11}$ , which becomes further from 0 as the follower response  $G_f$  departs from unity and as the round-trip

time of the communication ( $2T$ ) increases. For this reason, the position-position architecture was implemented and tested once but discarded for very poor performance.

*Force-Position* ( $C_3 = C_4 = 0$ ):

This teleoperation scheme is the naive method of providing force feedback. Quantitatively, the force-position teleoperation sets  $k_{po} = k_{do} = k_{pf} = k_{ff} = 0$ . In Fig. 2, this is equivalent to setting  $C_3 = C_4 = C_v^f = C_f^f = C_v^h = 0$ . The hybrid matrix then becomes

$$H(s) = \begin{bmatrix} 0 & e^{-sT}k_{fo} \\ -G_f & 0 \end{bmatrix} \quad (20)$$

The time-delayed force feedback is seen in  $h_{12}$ , and motion tracking is affected by the follower's response. If the tracking is not fast and accurate, it can lead to oscillation due to the delay. Often in practice  $k_{fo} < 1$  is used to improve stability. Intuitively, this reduces the jerk on the operator when the novice side changes suddenly, e.g., when first touching the patient. Similarly, lowpass filtering the fed-back force injects damping into the system, greatly reducing the oscillatory behaviour. Both, however, decrease transparency.

Additionally, in human teleoperation the follower can easily misjudge the alignment of the virtual and real tools due to imperfect depth perception as well as partial occlusion of the real tool by the virtual one. With this controller there is no possibility to correct the follower's pose to account for such steady-state errors.

*Force-Position with Local Position Feedback* ( $C_3 = C_4 = 0$ ): To correct for these steady state errors, the pose of the follower can be measured, and the tracking error fed back locally. In particular, the control law for the follower is then

$$\dot{x}_v = \dot{x}_o + k_{pf}(x_o - x_f) \quad (21)$$

In Fig. 2, this is equivalent to setting  $C_3 = C_4 = C_f^f = C_v^h = 0$ . To render this on the HoloLens 2, the desired velocity is numerically integrated using the time step between successive frames, and the resulting pose is applied to the virtual probe. To avoid losing track of the virtual probe in case of poor tracking,  $x_v$  is limited to stay within a distance  $d_{max}$  of  $x_o$ . By adding the position error to the virtual probe signal, the tracking error is effectively amplified with an integrator. If the follower is not well aligned in one axis, the virtual probe starts moving further away in that axis, causing the follower to realize their error and better align the probe. A deadband of an acceptably small distance error,  $d_{err}$ , is applied to the integrator since the error is never exactly zero.

This control scheme creates the following hybrid matrix, with  $k_{po} = k_{do} = k_{ff} = 0$ .

$$H(s) = \begin{bmatrix} 0 & e^{-sT}k_{fo} \\ -\frac{s + k_{pf}}{G_f^{-1}s + k_{pf}} & 0 \end{bmatrix} \quad (22)$$

As expected, this is very similar to the position-force architecture, but  $k_{pf}$  provides an opportunity to modify the  $h_{21}$  element, giving kinematic correspondence.

2) *3-Channel Teleoperation* ( $C_4 = 0$ ): As in the 2-channel force-position teleoperation, only the follower's force is reflected to the operator. However, both pose and force are sent from operator to follower, enabling the follower control law to include a local force feedback term as well. This is useful during contact phases of the teleoperation, in which the follower is to apply a certain desired force. With a relatively stiff environment, a minute change in position constitutes a large change in force, so employing only position and velocity in the control law leads to relatively poor force tracking. Therefore, the follower control law is given by

$$\dot{x}_v = \dot{x}_o + k_{pf}(x_o - x_f) + k_{ff}(f_h - f_f) \quad (23)$$

In Fig. 2, this is equivalent to setting  $C_4 = C_v^h = 0$ . The hybrid matrix with this controller becomes

$$H(s) = \begin{bmatrix} 0 & e^{-sT} k_{fo} \\ -\frac{s + k_{pf}}{G_f^{-1}s + k_{pf}} & \frac{k_{ff}(1 - e^{-s2T})s}{G_f^{-1}s + k_{pf}} \end{bmatrix} \quad (24)$$

From this it is apparent that the follower force now has a bearing on follower velocity, as desired. Numerically, this decreases transparency. However, in practice the follower is relatively insensitive to small changes in virtual probe pose and the human hand has limited motion resolution. These effects are not well modeled and make the added local force feedback term desirable.

3) *Model-Mediated Teleoperation*: One problem with the above methods is that the operator's sensation depends directly on the follower's actions. While this gives a true representation of the system on the follower side, it can lead to an unsteady experience for the operator who feels every small jolt and inadvertent movement of the follower. This effect can be reduced using low-pass filtering, but only at the cost of response speed. Furthermore, the performance decreases with increasing time delays.

One method to potentially overcome both problems is to render a local virtual haptic environment for the operator that is a replica of the follower's real environment. To do so, a point cloud or mesh of the follower environment is captured and sent to the operator. This mesh is used as a virtual fixture or keep-out volume by the haptic device, which applies an outward force when moved into the volume by the operator. This is known as model-mediated teleoperation [39]. Additionally, the impedance of the follower environment can be estimated and applied to the virtual mesh to replicate the feel of the real environment.

In methods presented and tested in this paper we assume an accurate mesh of the environment is available. This is feasible through RGB-D cameras, LiDAR, time-of-flight (ToF) depth cameras, stereo reconstruction, or even monocular 3D reconstruction from camera motion. Almost all robotic or human teleoperation systems are equipped with at least one of the aforementioned sensing modalities. For example, the HoloLens 2 has stereo cameras and a ToF depth sensor, and can easily be augmented with a higher-performance external RGB-D camera. The local model of the environment is sent to the operator and can be updated if the environment changes.

In addition, the impedance of the environment is assumed to be estimated accurately in real time by the force and pose tracking of the follower tool. Since the environment impedance is generally relatively constant, temporary loss in communication is less disruptive in this method than those described in previous subsections because the system can continue to render the last known impedance. However, there may also be sudden changes in impedance as the follower moves, so the impedance estimation must converge quickly.

The haptic device and follower forces are given simply by the patient model with the estimated impedance:

$$f_h = -\hat{b}_p \dot{x}_o - \hat{k}_p x_o \quad (25)$$

$$f_f = -b_p \dot{x}_f - k_p x_f \quad (26)$$

Plugging in the follower model gives

$$f_h = -\left(\hat{b}_p + \frac{\hat{k}_p}{s}\right) v_o \quad (27)$$

$$f_f = -\left(b_p + \frac{k_p}{s}\right) G_f v_v \quad (28)$$

For the virtual probe pose, we can again use Equation 8, and substitute in  $v_f = G_f v_v$ .

$$\begin{aligned} v_o &= e^{-sT} v_o + \frac{k_{pf}}{s} (e^{-sT} v_o - G_f v_v) + k_{ff} (e^{-sT} f_h - f_f) \\ &= \frac{1 + \frac{k_{pf}}{s} - k_{ff} \left(\hat{b}_p + \frac{\hat{k}_p}{s}\right)}{1 + \left[\frac{k_{pf}}{s} - k_{ff} \left(b_p + \frac{k_p}{s}\right)\right] G_f} e^{-sT} v_o \end{aligned} \quad (29)$$

Thus we obtain an expression for the follower velocity:

$$v_f = \frac{1 + \frac{k_{pf}}{s} - k_{ff} \left(\hat{b}_p + \frac{\hat{k}_p}{s}\right)}{\frac{1}{G_f} + \frac{k_{pf}}{s} - k_{ff} \left(b_p + \frac{k_p}{s}\right)} e^{-sT} v_o \quad (30)$$

Similarly, from Equations 25 and 26, the follower force is

$$f_f = \left(\frac{b_p + \frac{k_p}{s}}{\hat{b}_p + \frac{\hat{k}_p}{s}}\right) \frac{1 + \frac{k_{pf}}{s} - k_{ff} \left(\hat{b}_p + \frac{\hat{k}_p}{s}\right)}{\frac{1}{G_f} + \frac{k_{pf}}{s} - k_{ff} \left(b_p + \frac{k_p}{s}\right)} e^{-sT} f_h \quad (31)$$

Here it is clear that if  $G_f = 1$  and the impedance estimate is perfect, the follower motion and force will correspond precisely to the delayed operator motion and force, so the hybrid matrix will be

$$H = \begin{bmatrix} 0 & e^{sT} \\ -e^{-sT} & 0 \end{bmatrix} \quad (32)$$

Notice how the time delay in the force is negative. Depending on the communication delay, the operator feels the force feedback from the local model before the follower has applied it. In this way, the model-mediated method gives predictive haptic feedback.

In general, for imperfect impedance estimation, the damping is very small, and at steady-state the velocity should be zero,

so damping is negligible. Assuming  $G_f$  is constant, the steady state gain is

$$f_f = \left( \frac{k_p}{\hat{k}_p} \right) \frac{k_{pf} - k_{ff}\hat{k}_p}{k_{pf} - k_{ff}k_p} f_h \quad (33)$$

If  $k_{pf} - k_{ff}k_p > 0$  and the tissue impedance is underestimated, i.e.  $k_p > \hat{k}_p$ , the operator moves too far into the patient for a given force, so the follower must press harder to match the virtual tool and  $f_f > f_h$ . Conversely, overestimation leads to  $f_f < f_h$ .

If the impedance model is inaccurate, a small  $k_{ff}$  reduces the pose tracking error due to the mismatch, but the force experienced by the operator remains proportionally inaccurate. If the follower response,  $G_f$  is poor, large  $k_{ff}$  can suppress the error. Thus, every effort should be made to have an accurate and fast impedance estimation and patient model while using a reasonably large  $k_{ff}$  to compensate for the follower response. Making  $k_{ff}$  too large, however, becomes unintuitive for the follower and amplifies modeling errors. Additionally, taking  $k_{pf} = 0$  ensures a steady-state gain of 1 but may lead to poor position tracking. Thus, the best strategy may be to have large  $k_{ff}$  and zero  $k_{pf}$  in the normal direction to the tissue (the direction of the force), and non-zero  $k_{pf}$  in the tangent directions.

Other approaches are also possible wherein virtual tool color, error bars, arrows, or other rendering techniques are used to give visual force or pose error feedback to the follower. However, these tend to overwhelm the follower with too much information and should be avoided if possible [3]. These are also difficult to model and are not considered in this paper.

### E. Stability

Since the follower is a person, not a robot, instability does not lead to violent motions that cause harm or damage. However, it does render the teleoperation completely unusable, so some analysis of stability can give insight into how well the system will perform. In this section we use a state space formulation of the model that is derived in the Appendix.

For the zero time delay case, the system stability is given by the eigenvalues of the  $A$  matrix in Equation 49. These depend on the choice of the feedback and feedforward gains from Equations 42 and 8. These can lead to stable or unstable behavior and must be chosen carefully.

For force-position control with no local feedback loops, only  $k_{fo}$  is non-zero. In this case,  $A$  is stable. However, it has one eigenvalue close to zero. Increasing  $k_{fo}$  makes the follower slower and more oscillatory and in turn pushes this eigenvalue closer to zero. At large enough  $k_{fo}$ , the system becomes unstable. Thus, we expect that with poor tracking from the follower, the force-position feedback will be oscillatory.

Conversely, if we add local position and force feedback for the follower, the hybrid matrix is given in Equation 24. By the Routh-Hurwitz criterion for a third degree polynomial,  $h_{21}$  is stable if

$$b_f(k_f + k_{pf}b_f) - m_f k_{pf} k_f > 0 \quad (34)$$

This is clearly true if  $k_{pf} = 0$ , meaning no local position feedback. With the feedback, however, using the parameter

values that gave realistic results in the follower model simulations, the system is stable for any  $k_{pf} \geq 0$ . Depending on the follower's response, however, this may change. Notice also that with no time delay the stability is independent of the local force feedback,  $k_{ff}$ .

For the general time delayed case, we express the state space model from Equation 52 in the Appendix with time delays as follows

$$\dot{\mathbf{x}}(t) = A_0 \mathbf{x}(t) + \sum_{n=1}^2 A_n \mathbf{x}(t - nT) + B \mathbf{x}_o^*(t) \quad (35)$$

The  $A_0$ ,  $A_1$ , and  $A_2$  matrices are derived in the Appendix.

According to Mori et al. and Cheres et al. [40], [41], the system is asymptotically stable independent of time delays if

$$\mu(A_0) + \sum_{n=1}^2 \|A_n\|_2 < 0 \quad (36)$$

Where  $\mu(A) = \frac{1}{2} \lambda_{max}(A^\top + A)$  and  $\|A\|_2 = \sqrt{\lambda_{max}(A^\top A)}$ , where  $\lambda_{max}$  is the maximum eigenvalue. This is difficult to show in general, but can be used to test specific known gains and parameters.

We can also consider the controllers described above using the hybrid matrices. Anderson and Spong (1989) showed that a system is passive if the norm of its scattering operator is less than or equal to 1:

$$S(s) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} (H(s) - I)(H(s) + I)^{-1} \quad (37)$$

$$\|S(s)\|_2 \leq 1 \implies \text{Stable}$$

For example, the hybrid matrix in Equation 32 leads to a scattering operator norm of 1, so stability is maintained irrespective of delay. However, this is more difficult to show when the follower response or parameter estimation are non-ideal. In reality, the impedance estimation is time-varying, delayed, and not perfectly accurate. Furthermore, passivity is a very conservative method of guaranteeing stability, so a system may be stable even if  $\|S(s)\|_2 > 1$ .

In the following experiments, we will show empirically that stability is much less of a concern in human teleoperation than in robotic teleoperation, and is in this case secondary to performance and transparency.

### F. Experimental Setup

To test the human teleoperation architectures in a controlled and close-to-ideal environment, we developed the following system. The follower and operator sides both consisted of a host PC running Windows 11 and a haptic device (Touch X, 3D Systems, Rock Hill, SC). The haptic devices were controlled via C++ programs using OpenHaptics, with a graphical user interface written in Qt. The two clients communicated over a fast WebRTC connection, exchanging position, velocity, and force data, as well as occasional synchronization messages to measure the communication time delay [4]. The time delay was controlled by handling all sending and receiving of WebRTC messages on the operator side on a separate thread, which delayed every message by a set time before

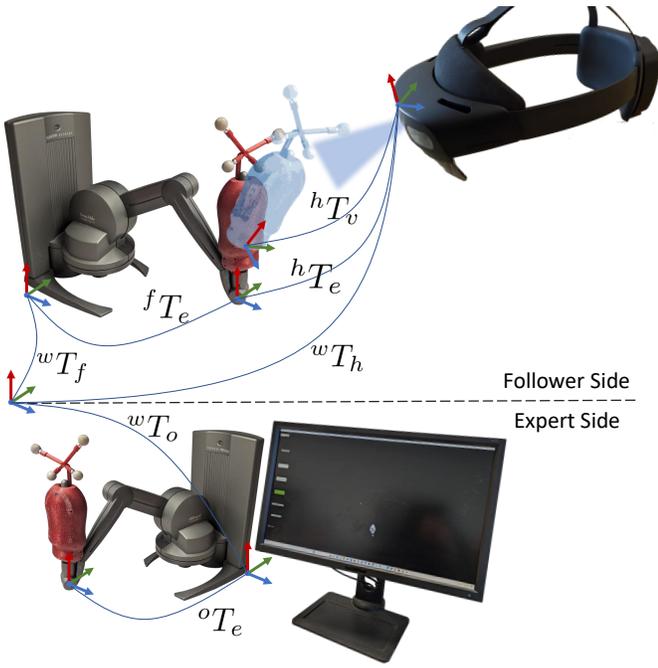


Fig. 6. Experimental Setup showing the operator and follower sides which shared one host computer but were separated by a screen. The virtual probe rendered by the HoloLens 2 is shown in semitransparent blue. The red objects on the haptic devices are the ultrasound probe-shaped end effectors with retro-reflective IR markers.

sending it or forwarding it to the main graphics and haptics threads. The follower side also had a HoloLens 2 MR HMD which communicated with the follower side PC application over a local WebSocket, receiving messages forwarded from the WebRTC connection.

The virtual tool used in the tests was an ultrasound probe viewed through the HoloLens 2. An identical ultrasound probe shape was 3D printed and attached to the follower’s Touch X stylus. Four infrared (IR) reflective spheres were also attached to the dummy probe and were tracked by the HoloLens 2, allowing the HoloLens to compute the haptic device end effector pose in the HoloLens frame [7]. This was used to perform a registration before every test. Suppose the operator’s haptic device has base frame represented by the homogeneous transform  ${}^wT_o$  relative to the world while the follower’s has base frame  ${}^wT_f$  and the HoloLens is at  ${}^wT_h$ , as shown in Fig. 6. When the operator moves their haptic device end effector to a pose  ${}^wT_o{}^oT_e$ , the virtual probe is rendered in the HoloLens at pose  ${}^wT_h{}^hT_v$ , and the follower moves their handle to pose  ${}^wT_f{}^fT_e$ . The virtual probe should be rendered such that when the follower aligns perfectly to the virtual probe (i.e.  ${}^wT_f{}^fT_e = {}^wT_h{}^hT_v$ ), their handle is at  ${}^fT_e = {}^oT_e$ . In this way, if the follower matches the operator, the poses output by the haptic device control software, OpenHaptics, are equal. Note that the follower pose is  ${}^fT_e = ({}^wT_f)^{-1}{}^wT_h{}^hT_v$ , so to achieve  ${}^fT_e = {}^oT_e$ , the virtual probe pose must be set to  ${}^hT_v = ({}^wT_h)^{-1}{}^wT_f{}^oT_e$ . For this, we must determine the follower haptic device base frame relative to the HoloLens:  ${}^hT_f = ({}^wT_h)^{-1}{}^wT_f$ .

Thus, at the start of each test, the follower haptic device

end effector was moved to several different positions and orientations while the HoloLens recorded the measured pose ( ${}^hT_e$ ) and received the actual pose from OpenHaptics ( ${}^fT_e$ ) over the WebSocket. These are related by  ${}^wT_h{}^hT_e = {}^wT_f{}^fT_e$ , so  ${}^hT_f = ({}^wT_h)^{-1}{}^wT_f = {}^hT_e({}^fT_e)^{-1}$ . After recording several hundred samples of ( ${}^hT_e$ ,  ${}^fT_e$ ), the least squares-optimal value of  ${}^hT_f$  was computed using the Kabsch-Umeyma algorithm, which ensures a valid homogeneous transform using singular value decomposition (SVD) [42].

During tests, the operator performed arbitrary motions with their haptic device. The motion and force were sent to the follower PC over WebRTC, and forwarded to the HoloLens over the WebSocket. Here they were transformed into HoloLens coordinates and rendered to the follower through the virtual probe pose. The follower grasped the dummy ultrasound probe on their haptic device end effector and aligned it with the virtual probe to follow the desired motion. A virtual environment in the form of a flat surface of specified stiffness and damping was rendered haptically to the follower, leading to forces generated due to the motions. The applied forces and follower motions were sent to the HoloLens for rendering local feedback, as described above, and to the operator for haptic feedback. Depending on the architecture being tested, the operator’s haptic device either applied forces based on this feedback and/or simply rendered an identical virtual environment to the follower one to test the model-mediated feedback.

The operator and follower PC clocks were synchronized by sending timestamps along with every data message, and all forces, positions, orientations, and velocities were in the same coordinate frame since both sides used identical haptic devices. All data received on the operator side was recorded to a file for later analysis in MATLAB (Mathworks Inc., Natick, MA).

### G. Tests

Using this setup, we performed preliminary tests of the various teleoperation architectures. These tests provide a validation of our mathematical modeling under controlled conditions. The architectures we considered are listed below and described in the previous sections.

- Position-Position (PP):  
 $\dot{x}_v = \dot{x}_o$  and  $f_h = k_{po}(x_f - x_o) - k_{do}\dot{x}_o$
- Force-Position (FP):  
 $\dot{x}_v = \dot{x}_o$  and  $f_h = k_{fo}f_f - k_{do}\dot{x}_o$
- Force-Position with local position feedback (FPP):  
 $\dot{x}_v = \dot{x}_o + k_{pf}(x_o - x_f)$  and  $f_h = k_{fo}f_f - k_{do}\dot{x}_o$
- Force-Position with local pose and force feedback (FPPF):  
 $\dot{x}_v = \dot{x}_o + k_{pf}(x_o - x_f) + k_{ff}(f_h - f_f)$  and  
 $f_h = k_{fo}f_f - k_{do}\dot{x}_o$
- Model-mediated (M):  
 $\dot{x}_v = \dot{x}_o$  and  
 $f_h = \begin{cases} k_p(x_o - x_o) - b_p\dot{x}_o & \text{if } x_o \text{ inside mesh} \\ -b_p\dot{x}_o & \text{otherwise} \end{cases}$   
where  $x_o$  is the position of the virtual mesh surface.
- Model-mediated with local position feedback (MP):  
Same as M but with  
 $\dot{x}_v = \dot{x}_o + k_{pf}(x_o - x_f)$

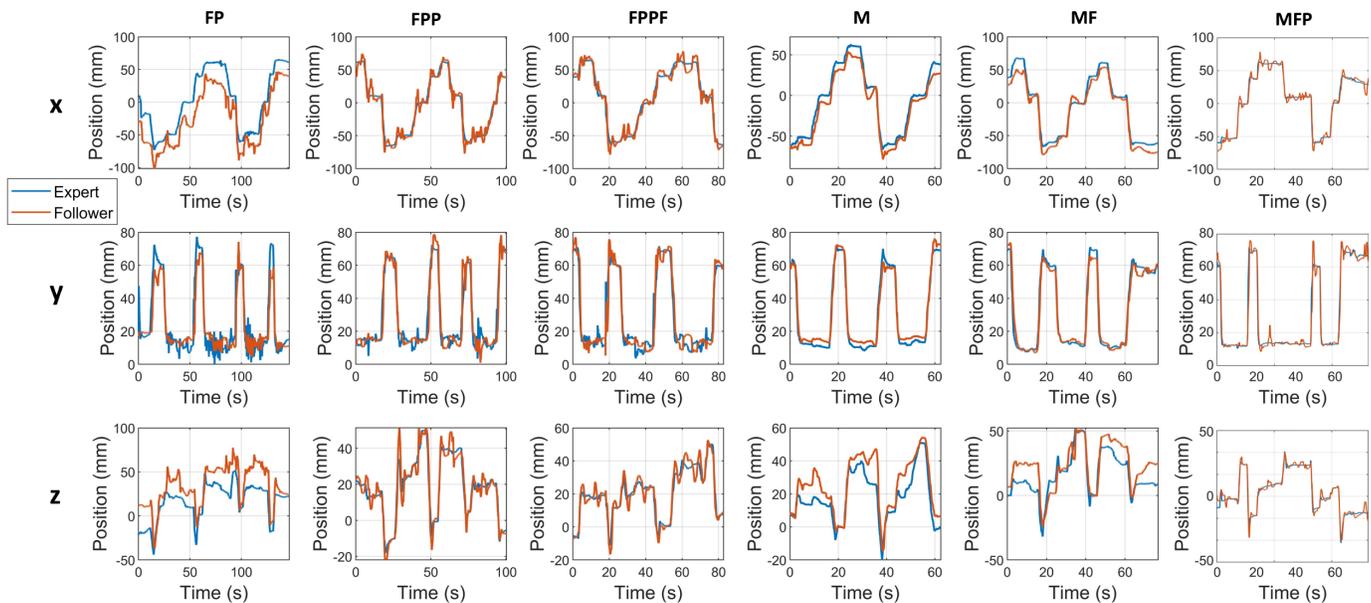


Fig. 7. Position tracking with no time delay for the six teleoperation architectures. FP shows significant oscillation during contact while FPP, M, and MF have relatively poor position tracking. Adding explicit local position feedback for the follower greatly improves the tracking accuracy and speed.

- Model-mediated with local position and force feedback (MFP):

Same as M and MP but with

$$\dot{x}_v = \dot{x}_o + k_{pf}(x_o - x_f) + k_{ff}(f_h - f_f)$$

Note, in MFP the force feedback was given normal to the surface and the position feedback was tangent to the surface. The surface in these tests was a flat, horizontal virtual surface in the haptic device with set stiffness and damping.

For each teleoperation architecture, the operator performed arbitrary smooth motions for 1 to 4 minutes with the follower tracking. The position and force root-mean-squared (RMS) tracking errors were analyzed, as well as the steady state error for step-like motions. Next, the two most promising methods of each type (force feedback and model-based feedback) were selected and tested again five times, each time with increasing communication time delay (0, 50, 250, 500, and 1000 ms). Then the same methods were tested with three different environment stiffness values (0.2, 0.4, and 0.8 N/mm) and no time delay, with and without low-pass filtering.

In the tests, the operator was one of the authors, and the follower was a volunteer. Five followers were tested, including three males and two females with age ranging from 22 to 57, with various backgrounds. None were experienced in ultrasound or MR. Ethics approval for the tests was obtained from the University of British Columbia Behavioural Research Ethics Board (BREB), approval number H22-01195. Informed consent was given by all participants.

### III. RESULTS

#### A. Tracking Tests with no Delay

The average results of the first test are shown in Table II, and one set of tests is plotted in Figs. 7 and 8. For these nominally zero time delay tests, the communication round trip time was measured to be  $3.69 \pm 2.67$  ms (average  $\pm$  standard deviation)

TABLE II  
TELEOPERATION PERFORMANCE FOR DIFFERENT CONTROL SCHEMES AT ZERO TIME DELAY. THE RMS POSITION TRACKING ERROR,  $e_p$ , STEADY-STATE POSITION ERROR,  $e_{pss}$ , RMS FORCE TRACKING ERROR,  $e_f$ , STEADY STATE FORCE ERROR,  $e_{fss}$ , AND APPROXIMATE TIME DELAY,  $\tau$ , BETWEEN FOLLOWER AND OPERATOR ARE SHOWN.

	$e_p$ (mm)	$e_{pss}$ (mm)	$e_f$ (N)	$e_{fss}$ (N)	$\tau$ (ms)
FP	7.92	6.57	<b>0.18</b>	0.12	440
FPP	5.91	3.83	0.16	0.11	38
FPPF	6.18	<b>1.52</b>	0.33	<b>0.07</b>	48
M	6.75	5.12	2.66	2.45	254
MF	6.91	6.23	0.94	0.9	284
MFP	<b>3.86</b>	3.38	0.58	0.47	<b>36</b>

using the method described in [4]. For later non-zero delay tests, the nominal time delay is also reported but may vary similarly.

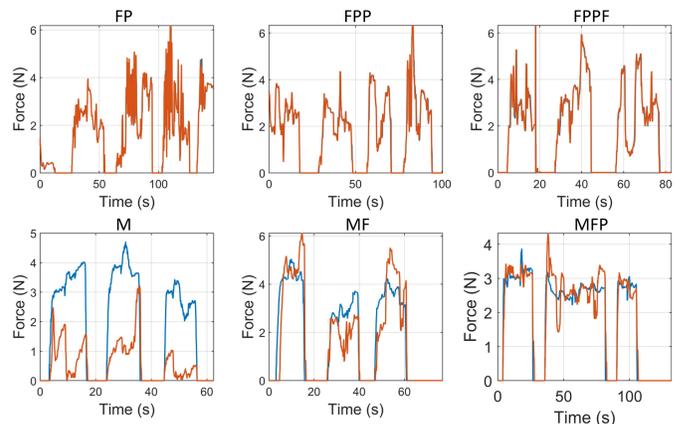


Fig. 8. Force tracking with no time delay using various teleoperation schemes.

Even with no delay, it was immediately apparent that position-position teleoperation was impractical. The lag of

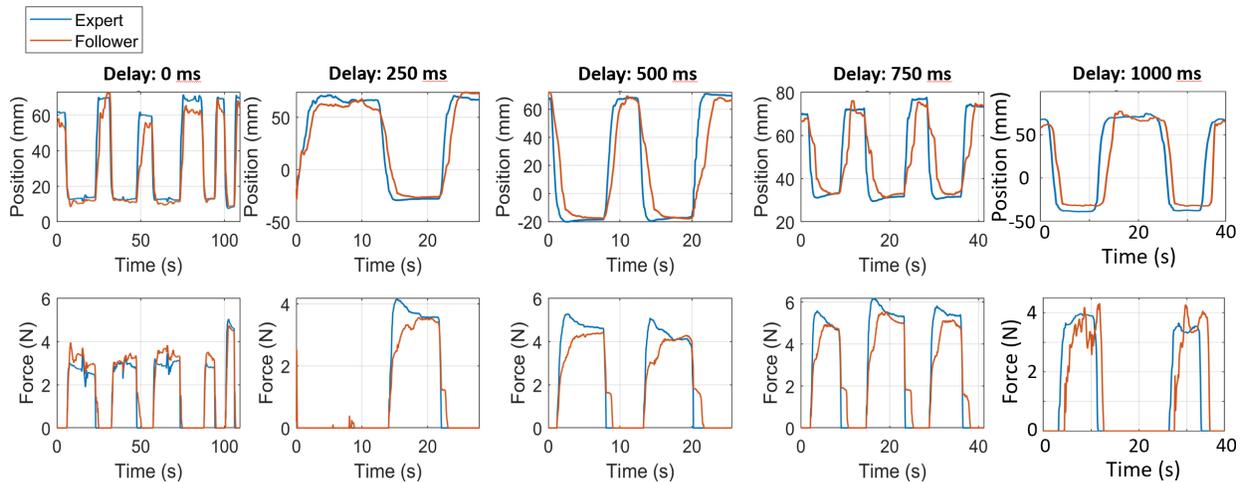


Fig. 9. Position and force tracking (both normal to the virtual surface) in the presence of time delays with the MFP controller (mesh local model with local position and force feedback). Time delay has little effect on the force tracking and only delays the position tracking.

the follower in tracking the input signal made it difficult for the operator to move at all if the stiffness was high, while low stiffness gave the operator very poor feeling of where the surface was. Additionally, the follower, being human, does not track very smoothly, so the operator continually received unwanted and jarring feedback. Low-pass filtering helped but increased the delay, which again made it harder for the operator to move. Thus, while the position tracking was by definition good, the teleoperation experience was intolerable and the method was removed from the tests.

Direct force feedback performed much better, though the lack of explicit position coordination on the operator side was apparent, with relatively large tracking errors. Additionally, contact was oscillatory: due to the tracking lag of the follower, the operator initially moved well below the surface before receiving force feedback which was stronger than anticipated due to the follower's attempt to reach the virtual tool inside the surface. This jerked the operator's hand upward which the follower tried to track, thus decreasing the force, so the operator's hand moved down again. This leads to periodic and uncomfortable motion when in intermittent contact. This is similar to the chattering effect in robots, but at a lower rate and larger amplitude. This back-and-forth reflection of waves is very apparent in Fig. 8. Both behaviors match what was expected from Equation 20.

Adding local position feedback for the follower helped decrease position error and substantially reduced tracking lag. By effectively magnifying any tracking error, the local feedback has a predictive effect, moving the probe further than the operator has yet moved and encouraging the follower to react immediately. This reduced the oscillation during contact. The feedback also led to much more accurate position tracking, as shown in Fig. 7. With local position and force feedback, position tracking suffered slightly, since the virtual probe no longer exclusively represents the desired pose, but aims to control the force as well. In Equation 24, this is reflected by the non-zero  $h_{22}$  element, while  $k_{pf}$  enables the more responsive position tracking. Interestingly, RMS force tracking

also became slightly worse since the increased feedback to the follower sometimes caused initial overshoot. However, in steady-state the force quickly converged to be more accurate than without local feedback. The addition of force additionally had a damping effect during contact, discouraging sudden changes in force and thus leading to much less oscillation.

The teleoperation architectures based on a local model or mesh showed different behavior. Contact oscillation was completely eliminated due to the stable and unchanging mesh. This led to a much more comfortable and intuitive experience. With local position feedback, the position tracking was the best of any of the methods. On the other hand, force tracking was completely lacking unless explicitly enforced with local feedback, as shown in Fig. 8. Even so, the force experienced by the operator was not as accurate as when the measured force was fed back directly, and is affected by modeling errors in the surface shape and impedance. Again, adding local position feedback greatly reduced the tracking lag.

### B. Tracking Tests with Delay

The effect of communication delays on the force and position tracking is shown in Figs. 9 and 10 for the FPPF and MFP controllers. Due to the delay, discussing RMS error is not meaningful. However, several important results are apparent in the plots. With the 3-channel architecture, the character of the response underwent two distinct phases. Initially, up to between 250-500 ms delay, the response became increasingly oscillatory as the wave reflections were exacerbated by the larger delays. At 500 ms, however, the operator's approach changed and they started moving more slowly and carefully, anticipating the unexpected application of force feedback. As a result, the oscillations ceased. However, as the delay continued increasing, the operator was deeper and deeper within the surface before the follower's force was applied, measured, and fed back. This led to excessive offsets between the virtual probe and the follower, making tracking impossible.

Conversely, with model-mediated teleoperation, the operator always received immediate haptic feedback with no delay.

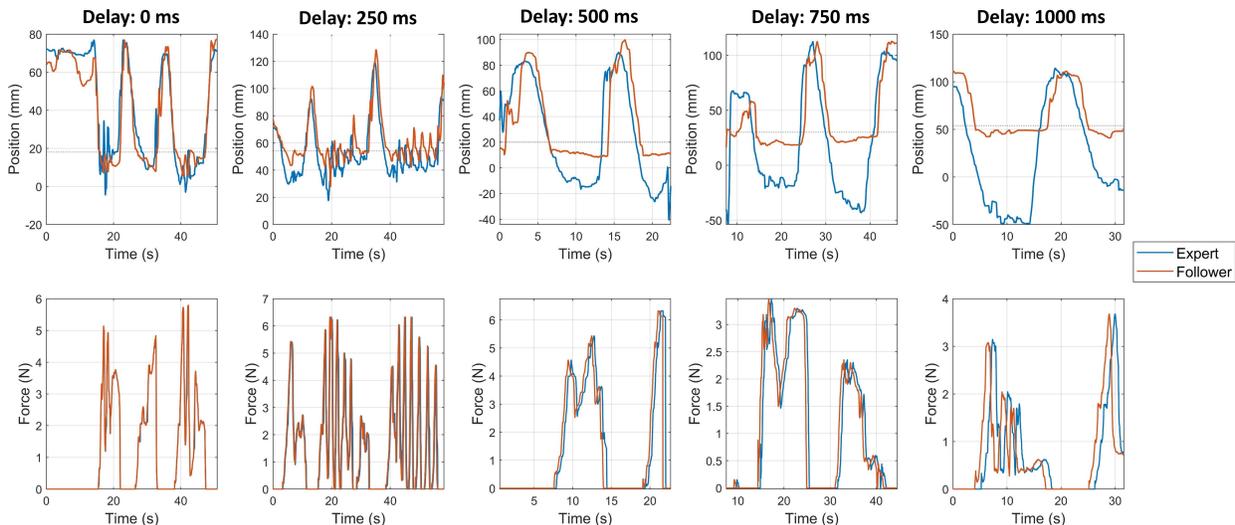


Fig. 10. Position and force tracking (both normal to the virtual surface) in the presence of time delays with the FPPF controller (force-position teleoperation with local position and force feedback for the follower). With delays over approximately 500 ms, tracking becomes infeasible.

Consequently, their motion was unaffected and the follower was able to track without problems, simply delayed by the given amount. Thus, while the delay made teleoperation impractical with direct force feedback, it had little effect when operating with a local model. Due to the local position and force feedback on the follower side, the tracking was good for both. This reflects what was expected in Equation 32 with perfect modeling, but will degrade with an imperfect local model.

### C. Environment Stiffness and Low-pass Filtering

Finally, the effect of environment stiffness on the teleoperation performance was tested. At zero delay, the stiffness was varied from 0.4 N/mm (the value used for all other tests) down to 0.2 N/mm to simulate a very soft environment and up to 0.8 N/mm for a stiff environment. The damping was kept constant and was not noticeable. These tests were performed with the FPPF architecture since the stiffness has little effect on the local mesh feedback as long as it is approximately equal for the operator and follower.

TABLE III  
TELEOPERATION PERFORMANCE (RMS TRACKING ERROR IN MM) FOR STIFF AND SOFT ENVIRONMENTS, WITH AND WITHOUT LOW-PASS FILTERING (LP).

Stiffness (N/mm)	0.2	0.4	0.8
FPPF	6.48	6.18	7.53
FPPF + LP	5.04	4.78	6.31

It was already noted in previous tests that the direct haptic feedback approaches had oscillatory behaviour in contact due to the delay of the follower. Therefore, low-pass filtering of the fed-back force was also tested, using a single-pole infinite impulse response filter with a time constant of approximately 0.1 seconds. This injects damping into the system and is expected to decrease the oscillation.

The results of these tests are shown in Table III and Fig. 11. The oscillatory behaviour increases with stiffness, as expected, but strongly decreases in the presence of low-pass filtering, without greatly increasing the phase lag. This leads to better RMS error, but more importantly substantially improved the feel for the operator. The feedback was much more stable and less disruptive than without the filtering. The lower environment stiffness slightly increased tracking error compared to normal. This is because on the softer virtual surface, the operator and follower could move more for small changes in force, making it harder to maintain a precise position.

### D. Robustness to Parameter Variation

Throughout this paper, we have made a number of simplifying assumptions about the system. In particular, the follower and operator models are taken to be constant. However, they may be affected by external, time-varying influences such as distractions or fatigue and, as described in Section II-D1, the follower may have imperfect depth perception. Additionally, the follower is often hesitant or unable to apply large forces on a patient, so may fail to follow a virtual tool placed too far within the patient. In this case,  $f_f \neq -k_p x_f - b_p \dot{x}_f$ , as assumed in Fig. 3, so additional dynamics are introduced. The linear time-invariant assumption in the model is thus not always accurate, and there is both parametric and dynamic uncertainty in the follower and operator models. Thus, the mathematical modeling can be improved through stochastic approaches as both users' actions are not deterministic and their imperfect sensory perception and motor skills lead to random fluctuations in the motions. These imperfections could be treated as disturbances in the existing model in future work. However, as shown in Fig. 5 and the related fit error, the models with constant parameters were able to capture complex recorded motions for an extended period of time without any drift due to parameters changing. Indeed, with

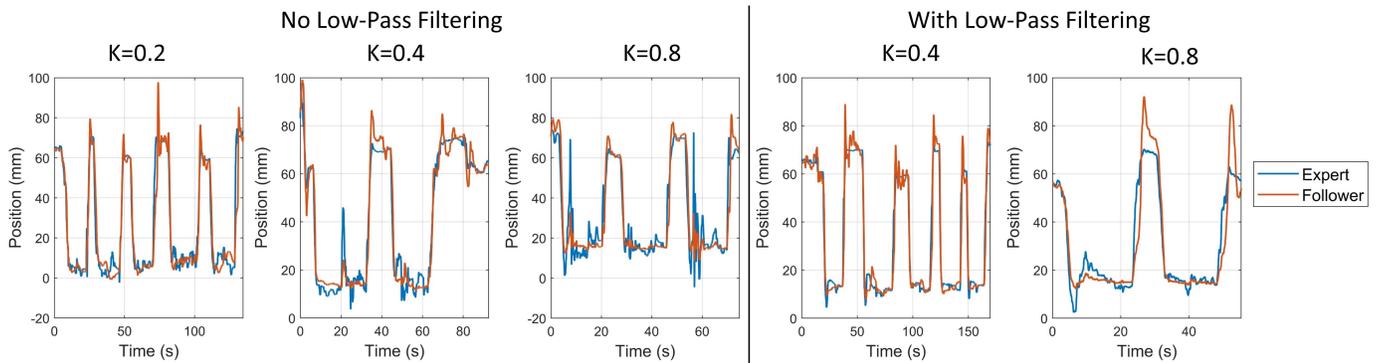


Fig. 11. Position tracking at different environment stiffnesses with and without low-pass filtering, with the FPPF teleoperator. In contact (position close to 0 mm), the oscillations increase with stiffness but decrease with low-pass filtering.

controller gains  $K_{fo}$ ,  $K_{of}$ ,  $K_{oo}$ , and  $K_{ff}$  chosen to give realistic performance at the parameter values identified in Table I, the acceptable ranges of the model parameters are shown in Table IV. To compute these values, the structured singular value,  $\mu$ , of the closed-loop system was computed in MATLAB with uncertainty on one parameter at a time. On average, the parameters can individually vary approximately 75% without causing instability. Thus, the LTI assumption gives a reasonable representation of the system.

TABLE IV

UPPER AND LOWER BOUNDS OF PARAMETER VALUES TO MAINTAIN STABILITY. THE AVERAGE ACCEPTABLE PERCENT VARIATION FROM THE MEAN PARAMETER VALUE IS 74.7% FOR THE SYSTEM TO REMAIN STABLE.

Parameter	Minimum	Maximum	Percent Variation
$k_p$	4.14	15.9	58.6%
$b_p$	0.093	1.91	90.7%
$b_f$	0.207	0.571	46.8%
$m_f$	0.020	0.056	47.4%
$m_o$	0.000	0.2	100%+
$b_o$	0.000	36.5	100%+
$k_a$	0.002	200	100%
$k_{o1}$	0.000	1.02	100% +
$k_{o2}$	63.7	96.3	20.4%

Additionally, though it is assumed that the follower can handle intermittent patient contact implicitly without affecting their tracking, switching contact affects the force fed back to the operator and may thus destabilize the closed-loop system. This is shown in Fig. 10, where the tracking is unstable at 250 ms delay during intermittent contact, leading to very ineffective teleoperation. This shows that instability in human teleoperation does not constitute a health or equipment hazard but makes guidance impossible. However, at higher delays this instability disappears due to an apparent change in strategy. The instability is also eliminated when using model-mediated teleoperation.

#### IV. DISCUSSION

This paper has introduced the concept of bilateral human teleoperation as a control problem analogous to robotic bilateral teleoperation but with some unique characteristics. The modeling and simulation of the system were described, and various control architectures were explored in terms of

their implementation, theoretical transparency and stability, simulation, and practical performance on a test system. The results show that high performance, stable, transparent bilateral human teleoperation is possible but that time delays pose a challenge. With less than approximately 200 ms communication latency, a three channel teleoperation system performed best. In this scheme, the follower's set point velocity is a function of the operator's velocity and the follower's position and force tracking error, and the follower's force is fed back to the operator with significant low-pass filtering. With time delays, however, direct feedback becomes impractical and alternate approaches are required. One such approach of model-mediated teleoperation, in which the operator interacts with a local virtual model of the follower's environment, had excellent performance irrespective of time delay, though it is dependent on an accurate model.

The paper represents an initial exploration of a novel topic and as a result has some limitations. Firstly, the tests were performed with few subjects and in idealized conditions. They will have to be augmented by practical tests with numerous subjects. For example, to achieve the performance of the ideal model-mediated teleoperation in this paper, future work will have to integrate a depth camera and fast, accurate environment impedance estimation into the system. Both exist, so the assumptions in this paper are not inaccurate. However, very fine details such as the patient's ribs during an ultrasound procedure may have a substantial effect on the measured force and will likely not be captured well in the mesh.

Similarly, the controllers presented here use pose and force, which have to be measured in real scenarios where the follower is not moving a haptic device. For human teleoperation, the force sensing must be low-profile and have sufficient load capability [5], [6], while pose tracking presents another set of challenges. Optical trackers are accurate but susceptible to occlusion, while electromagnetic trackers are affected by external disturbances. Extensive research has been performed in this direction [7], but the solutions are imperfect and future work will have to test practical bilateral teleoperation with all of these challenges.

Other future work includes investigation of further control architectures, for example predictive control using the mathematical models to extrapolate trajectories, or robust control by

defining uncertainties on the parameters in the derived state space model. This paper also did not consider architectures with non-zero  $k_{po}$ , i.e. four-channel approaches where the follower position is fed back to the operator. It has been shown that using all four channels of velocity and force is important for optimal transparency [43]. However, with the follower's unpredictable motions this may not be true for human teleoperation. Traditional passivity-based methods such as wave variables [27] for achieving time delay-robust stability were not explored in this paper because they trade off performance to maintain stability for the sake of safety. As instability in human teleoperation degrades performance rather than safety, this trade-off is not helpful. Time domain passivity control [36] is less conservative than wave variables and may lead to better performance but is left for future work.

Finally, orientation tracking was not explicitly evaluated in this paper but can be taken to be analogous to position tracking - i.e. consisting of similar lag and accuracy. This was the case in previous tests of human teleoperation [3]. Moreover, in [3], different rendering schemes were tested. These tests could be extended to find the best rendering method for improved depth perception. During the step-like motions in this paper, the distribution of the follower position was approximately Gaussian about the desired location. However, in the depth direction from the follower's point of view (approximately  $z$  in Fig. 7), the position variance was significantly greater than in the other directions, pointing to possible depth perception limitations.

In model-mediated teleoperation, the patient impedance model affects the stability. We assumed in the tests that the patient model is fixed and known, with no delay and no error in the parameter estimation. Of course, in reality the patient breathes and moves, and the impedance varies when examining different anatomies. For example, when scanning the liver and kidneys, the probe moves from ribs to soft tissue and back. Practical impedance estimation schemes based on pose and force measurement require several measurement samples to converge to an estimate. Thus, the impedance model is delayed, varying, and not perfectly accurate. However, as shown in Equation 33, the follower in the model-mediated teleoperation can converge to the correct normal force and tangent position despite modeling errors by having local force feedback as well as local position feedback in directions orthogonal to the force. In this case only the operator feels a slightly incorrect force that is only transient while the impedance estimation converges. In practice, the ultrasound probe is moved slowly and contact is maintained, so the impedance is approximately constant. Thus, the test results are realistic and show the effectiveness of model-mediated teleoperation in this system.

## V. CONCLUSION

This paper is a first exploration of bilateral human teleoperation from a controls perspective. The results show that three channel teleoperation is effective at small time delays while model-based teleoperation with local force and pose feedback achieves good performance even with large latency

and imperfect impedance estimation. With the developed system model, simulation, and practical setup, future work can build off the presented methods and tests to achieve high-performance, transparent, stable human teleoperation despite time delays. This will in turn enable effective remote guidance and execution of important tasks such as ultrasound exams with the relatively simple, low-cost mixed reality human teleoperation system.

## REFERENCES

- [1] D. Black, Y. Oloumi Yazdi, A. H. Hadi Hosseinabadi, and S. Salcudean, "Human teleoperation-a haptically enabled mixed reality system for teleultrasound," *Human-Computer Interaction*, pp. 1–24, 2023.
- [2] D. Black, M. Nogami, and S. Salcudean, "Mixed reality human teleoperation with device-agnostic remote ultrasound: Communication and user interaction," *Computers & Graphics*, vol. 118, pp. 184–193, 2024.
- [3] D. Black and S. Salcudean, "Human-as-a-robot performance in mixed reality teleultrasound," *International Journal of Computer Assisted Radiology and Surgery*, pp. 1–8, 2023.
- [4] D. G. Black, D. Andjelic, and S. E. Salcudean, "Evaluation of communication and human response latency for (human) teleoperation," *IEEE Transactions on Medical Robotics and Bionics*, 2024.
- [5] D. Black, A. H. H. Hosseinabadi, N. R. Pradnyawira, M. Pol, M. Nogami, and T. Salcudean, "Towards differential magnetic force sensing for ultrasound teleoperation," in *2023 IEEE World Haptics Conference (WHC)*, pp. 333–339, IEEE, 2023.
- [6] D. Black, N. R. Pradnyawira, M. Nogami, A. H. H. Hosseinabadi, and S. Salcudean, "Low-profile 6-axis differential magnetic force/torque sensing," *IEEE Transactions on Medical Robotics and Bionics*, 2024.
- [7] D. Black and S. Salcudean, "Robust object pose tracking for augmented reality guidance and teleoperation," *IEEE Transactions on Instrumentation and Measurement*, 2024.
- [8] A. Saracino, A. Deguet, F. Staderini, M. N. Boushaki, F. Cianchi, A. Menciacchi, and E. Sinibaldi, "Haptic feedback in the da vinci research kit (dvrk): A user study based on grasping, palpation, and incision tasks," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 15, no. 4, p. e1999, 2019.
- [9] O. A. Van der Meijden and M. P. Schijven, "The value of haptic feedback in conventional and robot-assisted minimal invasive surgery and virtual reality training: a current review," *Surgical endoscopy*, vol. 23, pp. 1180–1190, 2009.
- [10] C. C. Alleblas, M. P. Vleugels, S. F. Coppus, and T. E. Nieboer, "The effects of laparoscopic graspers with enhanced haptic feedback on applied forces: a randomized comparison with conventional graspers," *Surgical endoscopy*, vol. 31, pp. 5411–5417, 2017.
- [11] G. Tholey, J. P. Desai, and A. E. Castellanos, "Force feedback plays a significant role in minimally invasive surgery: results and analysis," *Annals of surgery*, vol. 241, no. 1, pp. 102–109, 2005.
- [12] N. Diolaiti and C. Melchiorri, "Teleoperation of a mobile robot through haptic feedback," in *IEEE International Workshop HAVE Haptic Virtual Environments and Their*, pp. 67–72, IEEE, 2002.
- [13] T. M. Lam, H. W. Boschloo, M. Mulder, and M. M. Van Paassen, "Artificial force field for haptic feedback in uav teleoperation," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 39, no. 6, pp. 1316–1330, 2009.
- [14] A. Bolopion and S. Régner, "A review of haptic feedback teleoperation systems for micromanipulation and microassembly," *IEEE Transactions on automation science and engineering*, vol. 10, no. 3, pp. 496–502, 2013.
- [15] N. Smith-Guerin, L. Al Bassit, G. Poisson, C. Delgorge, P. Arbeille, and P. Vieyres, "Clinical validation of a mobile patient-expert telechography system using isdn lines," in *4th International IEEE EMBS Special Topic Conference on Information Technology Applications in Biomedicine, 2003.*, pp. 23–26, IEEE, 2003.
- [16] D. A. Lawrence, "Stability and transparency in bilateral teleoperation," *IEEE transactions on robotics and automation*, vol. 9, no. 5, pp. 624–637, 1993.
- [17] G. Burdea and J. Zhuang, "Dextrous telerobotics with force feedback—an overview part 2: control and implementation," *Robotica*, vol. 9, no. 3, pp. 291–298, 1991.
- [18] J. Kim, P. H. Chang, and H.-S. Park, "Transparent teleoperation using two-channel control architectures," in *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1953–1960, IEEE, 2005.

- [19] W.-H. Zhu, S. E. Salcudean, and M. Zhu, "Experiments with transparent teleoperation under position and rate control," in *Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No. 99CH36288C)*, vol. 3, pp. 1870–1875, IEEE, 1999.
- [20] K. Hashtrudi-Zaad and S. E. Salcudean, "Analysis and evaluation of stability and performance robustness for teleoperation control architectures," in *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No. 00CH37065)*, vol. 4, pp. 3107–3113, IEEE, 2000.
- [21] S. E. Salcudean, K. Hashtrudi-Zaad, S. Tafazoli, S. P. DiMaio, and C. Reboulet, "Bilateral matched impedance teleoperation with application to excavator control," *IEEE Control Systems Magazine*, vol. 19, no. 6, pp. 29–37, 1999.
- [22] F. Mobasser and K. Hashtrudi-Zaad, "Transparent rate mode bilateral teleoperation control," *The International Journal of Robotics Research*, vol. 27, no. 1, pp. 57–72, 2008.
- [23] H.-K. Lee and M. J. Chung, "Adaptive controller of a master–slave system for transparent teleoperation," *Journal of Robotic systems*, vol. 15, no. 8, pp. 465–475, 1998.
- [24] K. Hashtrudi-Zaad and S. E. Salcudean, "On the use of local force feedback for transparent teleoperation," in *Proceedings 1999 IEEE international conference on robotics and automation (cat. no. 99CH36288C)*, vol. 3, pp. 1863–1869, IEEE, 1999.
- [25] T. B. Sheridan, "Space teleoperation through time delay: Review and prognosis," *IEEE Transactions on robotics and Automation*, vol. 9, no. 5, pp. 592–606, 1993.
- [26] E. Nuño, L. Basañez, and R. Ortega, "Passivity-based control for bilateral teleoperation: A tutorial," *Automatica*, vol. 47, no. 3, pp. 485–495, 2011.
- [27] G. Niemeyer and J.-J. Slotine, "Stable adaptive teleoperation," *IEEE Journal of oceanic engineering*, vol. 16, no. 1, pp. 152–162, 1991.
- [28] A. Aziminejad, M. Tavakoli, R. V. Patel, and M. Moallem, "Transparent time-delayed bilateral teleoperation using wave variables," *IEEE Transactions on control systems technology*, vol. 16, no. 3, pp. 548–555, 2008.
- [29] A. C. Smith and K. Hashtrudi-Zaad, "Adaptive teleoperation using neural network-based predictive control," in *Proceedings of 2005 IEEE Conference on Control Applications, 2005. CCA 2005.*, pp. 1269–1274, IEEE, 2005.
- [30] A. C. Smith and K. Hashtrudi-Zaad, "Smith predictor type control architectures for time delayed teleoperation," *The International Journal of Robotics Research*, vol. 25, no. 8, pp. 797–818, 2006.
- [31] S. Sirouspour and A. Shahdi, "Model predictive control for transparent teleoperation under communication time delay," *IEEE Transactions on Robotics*, vol. 22, no. 6, pp. 1131–1145, 2006.
- [32] X. Xu, B. Cizmeci, C. Schuwerk, and E. Steinbach, "Model-mediated teleoperation: Toward stable and transparent teleoperation systems," *IEEE Access*, vol. 4, pp. 425–449, 2016.
- [33] W.-H. Zhu and S. E. Salcudean, "Stability guaranteed teleoperation: an adaptive motion/force control approach," *IEEE transactions on automatic control*, vol. 45, no. 11, pp. 1951–1969, 2000.
- [34] A. Haddadi and K. Hashtrudi-Zaad, "Robust stability of teleoperation systems with time delay: A new approach," *IEEE transactions on haptics*, vol. 6, no. 2, pp. 229–241, 2012.
- [35] B. Hannaford and J.-H. Ryu, "Time-domain passivity control of haptic interfaces," *IEEE transactions on Robotics and Automation*, vol. 18, no. 1, pp. 1–10, 2002.
- [36] J.-H. Ryu, C. Preusche, B. Hannaford, and G. Hirzinger, "Time domain passivity control with reference energy following," *IEEE Transactions on Control Systems Technology*, vol. 13, no. 5, pp. 737–742, 2005.
- [37] N. Hogan, "Stable execution of contact tasks using impedance control," in *Proceedings. 1987 IEEE International Conference on Robotics and Automation*, vol. 4, pp. 1047–1054, IEEE, 1987.
- [38] C. F. Guimarães, L. Gasperini, A. P. Marques, and R. L. Reis, "The stiffness of living tissues and its implications for tissue engineering," *Nature Reviews Materials*, vol. 5, no. 5, pp. 351–370, 2020.
- [39] P. Mitra and G. Niemeyer, "Model-mediated telemanipulation," *The International Journal of Robotics Research*, vol. 27, no. 2, pp. 253–262, 2008.
- [40] T. Mori and M. Kuwahara, "A way to stabilize linear systems with delay," *IFAC Proceedings Volumes*, vol. 14, no. 2, pp. 165–170, 1981.
- [41] E. Cheres, Z. Palmor, and S. Gutman, "Quantitative measures of robustness for systems including delayed perturbations," *IEEE Transactions on Automatic Control*, vol. 34, no. 11, pp. 1203–1204, 1989.
- [42] J. Lawrence, J. Bernal, and C. Witzgall, "A purely algebraic justification of the kabsch-umeyama algorithm," *Journal of research of the National Institute of Standards and Technology*, vol. 124, p. 1, 2019.
- [43] K. Hashtrudi-Zaad and S. E. Salcudean, "Analysis of control architectures for teleoperation systems with impedance/admittance master and slave manipulators," *The International Journal of Robotics Research*, vol. 20, no. 6, pp. 419–445, 2001.

## APPENDIX

In this section, we derive a state space model of the full teleoperation system. This formulation is used in the paper to assess stability and robustness.

From Section II-B, the follower is given by

$$\begin{aligned} \dot{\mathbf{x}}_f &= \begin{bmatrix} \dot{x}_f \\ \dot{\ddot{x}}_f \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{k_f}{m_f} & -\frac{b_f}{m_f} \end{bmatrix} \begin{bmatrix} x_f \\ \dot{x}_f \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \frac{k_f}{m_f} & \frac{b_f}{m_f} \end{bmatrix} \begin{bmatrix} x_v \\ \dot{x}_v \end{bmatrix} \\ \mathbf{y}_f &= \begin{bmatrix} x_f \\ \dot{x}_f \\ f_f \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -k_p & -b_p \end{bmatrix} \begin{bmatrix} x_f \\ \dot{x}_f \end{bmatrix} \end{aligned} \quad (38)$$

where  $k_p$ ,  $b_p$  are zero when not in contact with the patient. For this to be LTI, we assume the patient is unmoving and their impedance is constant. This is approximately true when scanning only one region, for example the abdomen. An advantage of this model is that we do not have to consider switching contact, which is handled implicitly by the follower.

The operator model is given by

$$m_o \ddot{x}_o + b_o \dot{x}_o = f_h + f_o \quad (39)$$

The operator changes their applied force depending on the haptic device force, to track a desired trajectory. In particular, to choose a representative response, we can represent the system in state space and use servo control with pole placement, choosing

$$f_o = -K_o \mathbf{x}_o + k_a \int (x_o^* - x_o) dt \quad (40)$$

where  $\mathbf{x}_o = [x_o, \dot{x}_o]^\top$ , and  $x_o^*$  is the operator's desired motion. The state feedback gains are  $K_o = [k_{o1} \ k_{o2}]$ . These can be tuned to obtain a response that resembles the recorded operator data. We ignore the operator's desired velocity for simplicity as it introduces a zero eigenvalue and the operator is primarily interested in position. Setting  $x_a = \int (x_o^* - x_o) dt$ , we find the following model

$$\begin{aligned} \dot{\mathbf{x}}_o &= \begin{bmatrix} \dot{x}_o \\ \dot{x}_a \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ -\frac{k_{o1}}{m_o} & -\frac{b_o - k_{o2}}{m_o} & \frac{k_a}{m_o} \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_o \\ \dot{x}_o \\ x_a \end{bmatrix} \\ &\quad + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} x_o^* + \begin{bmatrix} 0 \\ \frac{1}{m_o} \\ 0 \end{bmatrix} f_h \\ \mathbf{y}_o &= \begin{bmatrix} x_o \\ \dot{x}_o \\ f_o \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -k_{o1} & -k_{o2} & k_a \end{bmatrix} \begin{bmatrix} x_o \\ \dot{x}_o \\ x_a \end{bmatrix} \end{aligned} \quad (41)$$

where the gain on  $f_h$  is denoted  $E_o$ .

1) *Teleoperation System*: The full teleoperation system is shown in Fig. 12, given the state space matrices derived above. We assume the follower sees only the virtual tool pose and the operator receives feedback only by forces applied to the haptic device. This does not model the operator's visual feedback from video or other sensor streams such as ultrasound images, nor the verbal communication between

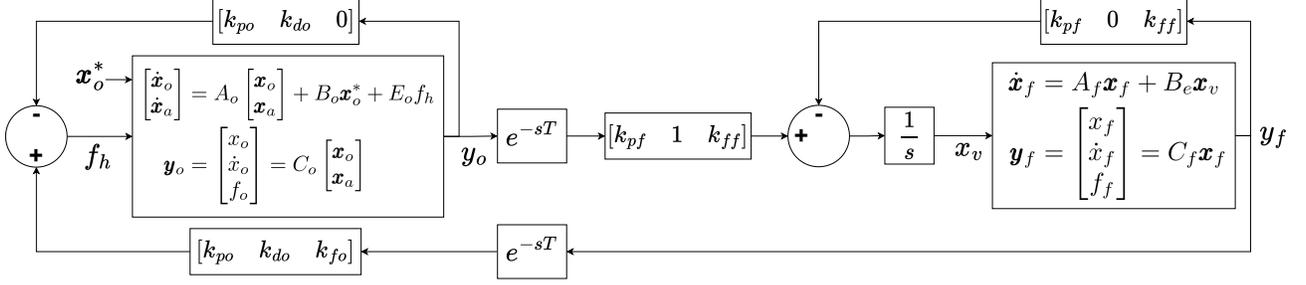


Fig. 12. State space representation of teleoperation system. Different behavior is achieved by changing the feedback and feed-forward gains of the positions, velocities, and forces. The operator-side model is given in Equation 41 while the follower side is in Equation 38.

operator and follower. Instead, the operator's decisions based on the ultrasound images determine their desired trajectory,  $x_o^*$ , which is the input to the model. In this case, we control the virtual tool pose,  $x_v$ , for the follower, and the haptic device force,  $f_h$  for the operator. Each can be a function of the position, velocity, and force of the operator and follower as shown in the diagram. Some gains are never used, so they are set to zero. For example, the haptic device is not equipped with a force sensor, so its actual force cannot be fed back. Thus, the general inputs to the operator and follower are

$$f_h = k_{fo}f_f + k_{po}(x_f - x_o) + k_{do}(\dot{x}_f - \dot{x}_o) \quad (42)$$

$$\dot{x}_v = \dot{x}_o + k_{pf}(x_o - x_f) + k_{ff}(f_h - f_f) \quad (43)$$

Let  $K_{oo} = [k_{po} \ k_{do} \ 0]$ ,  $K_{of} = [k_{pf} \ 1 \ k_{ff}]$ ,  $K_{ff} = [k_{pf} \ 0 \ k_{ff}]$ , and  $K_{fo} = [k_{po} \ k_{do} \ k_{fo}]$ . Then in terms of the state space outputs, these expressions become

$$f_h = K_{fo}\mathbf{y}_f - K_{oo}\mathbf{y}_o \quad (44)$$

$$\dot{x}_v = K_{of}\mathbf{y}_o - K_{ff}\mathbf{y}_f \quad (45)$$

as in Equation 9. Let  $\mathbf{x}_v = [x_v \ \dot{x}_v]^\top$  be the state of the virtual tool. The state equations of the operator, follower, and virtual tool are then

$$\begin{aligned} \dot{\mathbf{x}}_o &= A_o\mathbf{x}_o + B_o\mathbf{x}_o^* + E_o(K_{fo}C_f\mathbf{x}_f - K_{oo}C_o\mathbf{x}_o) \\ &= (A_o - K_{oo}C_o)\mathbf{x}_o + E_oK_{fo}C_f\mathbf{x}_f + B_o\mathbf{x}_o^* \end{aligned} \quad (46)$$

$$\dot{\mathbf{x}}_f = A_f\mathbf{x}_f + B_{f1}x_v + B_{f2}(K_{of}C_o\mathbf{x}_o - K_{ff}C_f\mathbf{x}_f) \quad (47)$$

$$\dot{x}_v = K_{of}C_o\mathbf{x}_o - K_{ff}C_f\mathbf{x}_f \quad (48)$$

Where  $B_f = [B_{f1} \ B_{f2}]$ . We now define a combined state  $\mathbf{x} = [\mathbf{x}_o^\top \ \mathbf{x}_f^\top \ x_v^\top]^\top$ . This allows us to state the system as one equation:

$$\begin{aligned} \dot{\mathbf{x}} &= \begin{bmatrix} A_o - E_oK_{oo}C_o & E_oK_{fo}C_f & 0_{3 \times 1} \\ B_{f2}K_{of}C_o & A_f - B_{f2}K_{ff}C_f & B_{f1} \\ K_{of}C_o & -K_{ff}C_f & 0 \end{bmatrix} \mathbf{x} \\ &+ \begin{bmatrix} 1 \\ 0_{5 \times 1} \end{bmatrix} x_o^* \\ \mathbf{y} &= \begin{bmatrix} C_o & 0_{3 \times 2} & 0_{3 \times 1} \\ 0_{3 \times 3} & C_f & 0_{3 \times 1} \end{bmatrix} \mathbf{x} \end{aligned} \quad (49)$$

The dimensions of the matrices are  $A \in \mathbb{R}^{6 \times 6}$ ,  $B \in \mathbb{R}^{6 \times 1}$ , and  $C \in \mathbb{R}^{6 \times 6}$ .

This analysis has only considered the case of time delay  $T = 0$ . For non-zero delay, we must rewrite Equations 44 and 45:

$$f_h = e^{-sT}K_{fo}\mathbf{y}_f - K_{oo}\mathbf{y}_o \quad (50)$$

$$\dot{x}_v = e^{-sT}K_{of}\mathbf{y}_o - K_{ff}\mathbf{y}_f \quad (51)$$

The system model with time delays is thus

$$\dot{\mathbf{x}}_o = A_o\mathbf{x}_o + B_o\mathbf{x}_o^* + E_oe^{-sT}f_h$$

$$\dot{\mathbf{x}}_f = A_f\mathbf{x}_f + B_{f1}x_v + B_{f2}\dot{x}_v$$

$$x_v = K_{of}C_oe^{-sT}\mathbf{x}_o - K_{ff}C_f\mathbf{x}_f$$

Combining this with equations 50 and 51, and with a slight abuse of notation to include time delays in the matrices, we obtain

$$\begin{aligned} \dot{\mathbf{x}} &= \begin{bmatrix} A_o - E_oK_{oo}C_oe^{-sT} & E_oK_{fo}C_fe^{-s2T} & 0_{3 \times 1} \\ B_{f2}K_{of}C_oe^{-sT} & A_f - B_{f2}K_{ff}C_f & B_{f1} \\ K_{of}C_oe^{-sT} & -K_{ff}C_f & 0 \end{bmatrix} \mathbf{x} \\ &+ \begin{bmatrix} 1 \\ 0_{5 \times 1} \end{bmatrix} x_o^* \\ \mathbf{y} &= \begin{bmatrix} C_o & 0_{3 \times 2} & 0_{3 \times 1} \\ 0_{3 \times 3} & C_f & 0_{3 \times 1} \end{bmatrix} \mathbf{x} \end{aligned} \quad (52)$$

Thus, we have the following  $A$  matrices:

$$\begin{aligned} A_0 &= \begin{bmatrix} A_o & 0_{3 \times 2} & 0_{3 \times 1} \\ 0_{2 \times 3} & A_f - B_{f2}K_{ff}C_f & B_{f1} \\ 0_{1 \times 3} & -K_{ff}C_f & 0 \end{bmatrix} \\ A_1 &= \begin{bmatrix} -E_oK_{oo}C_o & 0_{3 \times 2} & 0_{3 \times 1} \\ B_{f2}K_{of}C_o & 0_{2 \times 2} & 0_{2 \times 1} \\ K_{of}C_o & 0_{1 \times 2} & 0 \end{bmatrix} \\ A_2 &= \begin{bmatrix} 0_{3 \times 3} & E_oK_{fo}C_f & 0_{3 \times 1} \\ 0_{2 \times 3} & 0_{2 \times 2} & 0_{2 \times 1} \\ 0_{1 \times 3} & 0_{1 \times 2} & 0 \end{bmatrix} \end{aligned} \quad (53)$$

These can be used with the Mori and Cheres identity in Equation 36 to determine the stability independent of time delays.