An Input Predictor for Time-Delayed Teleoperation

Christian Smith
Centre for Autonomous Systems, Royal Institute of Technology
Stockholm, Sweden. Email: ccs@kth.se

Abstract— **We describe a method for bridging internet time delays in a teleoperation scenario where the the time delays are large compared to the task execution time. The method proposed uses minimum jerk motion models to predict the input from the user a time into the future that is equivalent to the communication delay. We present results from a teleoperated ball catching experiment, where we show that the proposed method makes a significant improvement over traditional methods for teleoperation over intercontinental distances.**

I. Introduction

Teleoperation in robotics has been studied for more than half a century. One typical teleoperation scenario consists of an operator controlling a robot manipulator via haptic input, while viewing video feedback from the remote site. This would for example be tele-surgery, hazardous material handling, underwater robotics, or rescue robotics. In many of these cases, the perception and reasoning abilities as well as the task related skills of the human operator are of paramount importance for successfull task completion. It is therefore necessary to enable the operator to interact with the remote environment as transparently as possible.

The ultimate challenge for tele-operation control is for systems with significant dynamics, and with a need for closed loop control in reaction to sensory feedback. For such a system control becomes a major challenge when the time delay (τ) is a significant portion of the overall process time (T). Time-delayed teleoperation is a well-studied problem, but there is as yet no clear solution to this problem [1].

To study these kinds of systems we have chosen to study tele-operated catching. The flight time of a small object thrown across a room in an indoor setting is typically 600–900 ms, and if we were to consider teleoperated catching over the internet from one continent to another the roundtrip time-delay could easily be 150 ms which is on the order of 15–25% of the overall flying time. How does one select a control strategy for such systems?

It is well known that humans are visually dominated [2], and there are clear models that suggest that much of human actions are feed-forward driven rather than feedback driven [3]. Under such a control regime one would expect that motion for interaction is largely prospective, in particular for dynamic situations. The fundamental question we address is: given this knowledge, would it be possible to estimate and predict the motion input by the user with sufficient accuracy to allow it be to used to generate a predictive motion control signal for the slave device, thus eliminating the effects of the time delay?

We study the estimation of commands issued by the user under significant time pressure. The questions addressed are:

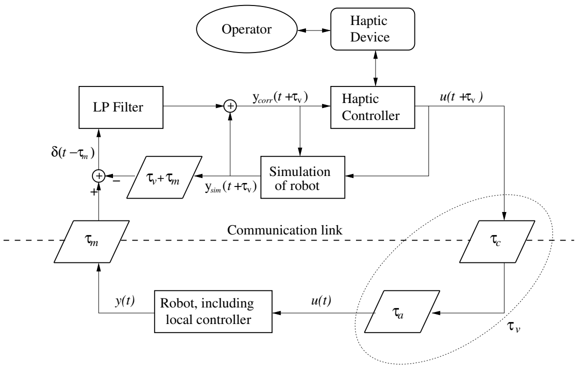
- What are good estimation models for identification of the trajectory specified by the user?
- How can these models be integrated into the control system?
- What kind of performance can be achieved with a real system?

II. Proposed Method

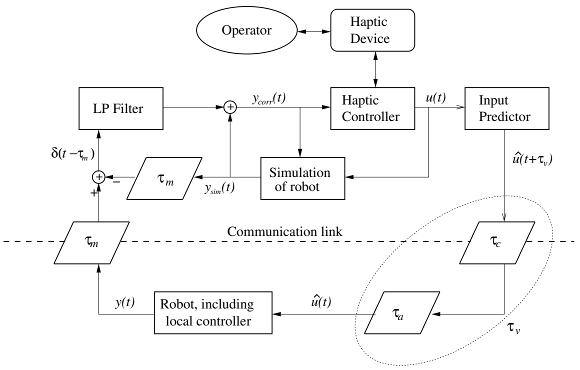
A well-tried approach to teleoperation under time delays is to use different methods to predict the remote site, or put in control terms, one substitutes the unavailable future measurement y with the predictor $\hat{y}$. Examples of methods that fall under this classification are Smith predictors, predictive displays, and different kinds of model predictive control (MPC). With these traditional approaches, the measurements are predicted for the entire roundtrip delay 2τ. An example of this is shown in the schematic of a Smith predictor in Figure 1(a).

We propose a novel control structure. Instead of handling the entire roundtrip delay by predicting the remote state with a simulation, the delay handling is split into an input predictor and a feedback predictor, each predicting the one-way delay τ. The principal structure is shown in Figure 1(b). For nonlinear systems, where errors grow more than linearly with the size of the delay, the total errors are significantly smaller with this approach.

The model we use to predict human motion is the well known *minimum jerk* (MJ) model [4]. Hand trajectories can be well explained by using this model in which the square sum of the third derivative of position, *jerk*, integrated over time is minimized. I.e, given a starting point, an end point and a time to move between the two, the trajectory that minimizes the jerk on this interval is the MJ trajectory. All MJ trajectories share the property that the 6th derivative is zero for the duration of the motion, and that they thus can be described as 5th degree polynomials. If a more complex motion is desired, or if the target of the motion is changed in mid-motion, the trajectory can be described by superpositioning several MJ trajectories. In our implementation, we fit the input data to a MJ model using least squares. The resulting polynomial is then extrapolated to obtain a prediction. The principles of this method have been presented in earlier work [5], [6].



(a) Control structure A, with Smith predictor only (skewed boxes represent delays)



(b) Control structure B, with Smith prediditor and MJ input predictor (skewed boxes represent delays)

Fig. 1. Comparison of control structures

III. Evaluation Experiments

Ballcatching has been chosen as an example task in order to study teleoperation of a process where the time-delay is a large fraction of the process cycle time. The task for the operator is to guide the robot's end effector to catch the ball. The operator has to perform this without any direct observations of the ball or the manipulator, with all information being relayed through a graphical user interface. A detailed description of the manipulator and the catching task can be found in [7].

The robot used in these experiments is a fast lightweight manipulator with six degrees of freedom. The end effector used is a simple cardboard cylinder. The user interface consists of a stereo display realized with a 21-inch CRT monitor and shutter glasses. An Omega haptic unit from Force Dimension served as input channel for user hand motion.

Internet time delays between the Royal Institute of Technology (KTH) in Stockholm and University of Genoa (UGE) in Genoa, Italy and Georgia Institute of Technology (GT) in Atlanta, Georgia, respectively, were used in the experiment. The setup was tried with a traditional Smith predictor working on the feedback signal (system A), and with the proposed MJ input predictor (System B). Also a scenario with no delays and no predictions was included for reference.

TABLE I

Results from ball catching experiment

Setup	Success rate ($\pm 2\sigma$)
No delay	0.288 (± 0.045)
System A, UGE delay (27 ms)	0.256 (± 0.062)
System A, GT delay (63 ms)	0.160 (± 0.052)
System B, UGE delay	0.275 (± 0.063)
System B, GT delay	0.280 (± 0.065)

10 novice subjects were each given 20 tries in each of these 5 different scenarios, totalling 200 measured tries for each scenario.

IV. Results

The performance results are presented in Table I. With the longer GT delay times, the performance with System A drops to approximately half the success rate that was achieved with zero delays, while the success rate of system B drops by less than 0.01 compared to the zero delay baseline. This result is statistically significant at $p < 0.05$.

With the shorter UGE delays, the success rate of System A drops by 0.03, while the success rate of System B drops by 0.01 as compared to the zero delay baseline. However, here the difference in performance is not large enough to be statiscally significant.

V. Conclusions

Experimental results show that the MJ predictor approach can succesfully bridge one-way time delays of more than 60 ms, for a task where a traditional Smith predictor approach encounters problems. In geographic terms, the difference of the two approaches is if a dynamic internet-based teleoperation task is limited to the same continent (within Europe), or if it is feasable to perform it intercontinentally (between Europe and North America).

References

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

[2] A. B. Schwartz, D. W. Moran, and G. A. Reina, "Differential representation of perception and action in the frontal cortex," *Science*, vol. 303, pp. 380–384, Jan 2004.

[3] A. Berthoz, *The Brain's sense of movement*, ser. Perspectives in Cognitive Science. London, UK: Harvard University Press, 2000.

[4] N. Hogan, "An organizing principle for a class of volontary movements," *Journal of Neuroscience*, vol. 4, pp. 2745–2754, 1985.

[5] M. Bratt, C. Smith, and H. I. Christensen, "Minimum jerk based prediction of user actions for a ball catching task," in *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. San Diego, Ca, USA: IEEE/RSJ, Oct 2007, pp. 2710–2716.

[6] C. Smith, M. Bratt, and H. I. Christensen, "Teleoperation for a ball-catching task with significant dynamics," *Neural Networks, Special Issue on Robotics and Neuroscience*, vol. 24, pp. 604–620, May 2008.

[7] C. Smith and H. I. Christensen, "Using COTS to construct a high performance robot arm," in *Proceedings of the 2007 IEEE International Conference on Robotics and Automation*. Rome, IT: IEEE, April 2007, pp. 4056–4063.