# A Predictive Approach for an End-to-End Touch-Latency Measurement

**Elie Cattan**
University of Grenoble
LIG, UJF
Grenoble, France
elie.cattan@imag.fr

**Amélie Rochet-Capellan**
University of Grenoble
Gipsa-lab, CNRS
Grenoble, France
amelie.rochet-capellan@gipsa-lab.grenoble-inp.fr

**François Bérard**
University of Grenoble
LIG, Grenoble-INP
Grenoble, France
francois.berard@imag.fr

## ABSTRACT
With direct-touch interaction, users are sensitive to very low levels of latency, in the order of a few milliseconds [3, 6]. Assessing the end-to-end latency of a system is thus becoming an important part of touch-devices evaluation, and this must be precise and accurate. However, current latency estimation techniques are either imprecise, or they require complex setups involving external devices such as high-speed cameras.

In this paper, we introduce and evaluate a novel method that does not require any external equipment and can be implemented with minimal efforts. The method is based on short-term prediction of the finger movement. The latency estimation is obtained on the basis of user calibration of the prediction to fully compensate the lag. In a user study, we show that the technique is more *precise* than a similar "low overhead" approach that was recently presented [1].

## ACM Classification Keywords
H.5.2 User Interfaces: Input devices and strategies (e.g., mouse, touchscreen)

## Author Keywords
latency; measure; direct-touch; prediction

## INTRODUCTION
Recent expansion of direct touch systems has prompted several research efforts on the effect and measurement of systems' end-to-end latency (or lag). The lag is defined as the time between a user's action and the display of the system's corresponding feedback. Lag was shown to reduce users' performances and the sense of presence in direct touch interaction. Users are able to perceive latencies down to 6ms [6] and their pointing performances are still affected at 10ms [3]. Most of current commercial touch surfaces exhibit levels of latency higher than 50ms [3], i.e. levels that are far from optimal. Researchers and designers of direct-touch systems should thus control the latency of their systems, which, as a first step, requires measuring it.

Measuring the end-to-end latency of touch system, however, is not an easy task. Several measuring instruments have been proposed in the literature. These instruments often require important efforts to set-up and operate, and they often require additional hardware such as cameras, pendulums or oscilloscopes. Berard and Blanch recently introduced a novel approach that requires the assistance of a human operator, but it is easy to implement and does not require any external hardware [1]. The approach is based on users' ability to accurately trace a circular path at constant speed. This is actually a difficult task that requires significant training to obtain good results.

In this paper, we introduce a new method that follows a similar approach of using the help of a human operator. The method requires the operators to slide a finger in straight lines at a constant speed of their choice, which is arguably an easier task that tracing circles at a constant rotational speed imposed by the system. This approach is based on the finger's motion prediction and the operator ability to perceive a mismatch between the finger and the lagging feedback. The technique only requires minimal efforts to implement and does not require any external device, or specific skills. We compared our predictive method to the low overhead (LO) method of Berard and Blanch [1] in a user study and showed that our new method is more *precise* than the LO method and generally preferred by participants.

## RELATED WORK
Estimating latency requires that, at a particular instant, the positions of the finger and the feedback are known. The feedback position is computed by the system and is thus easily reported. By contrast, it is more difficult to measure the physical position of the finger at the moment the feedback was displayed. A common approach to solve this problem is to use an external video camera that takes pictures of both the finger and the feedback [6]. However, this approach requires the time consuming hand labeling of many images to account for the variability of the latency. In addition, the precision of the approach is limited by the thickness of the finger, which makes difficult the estimation of the position of contact and the speed of the finger.


ACM 978-1-4503-3899-8/15/11 ...$15.00.
DOI: http://dx.doi.org/10.1145/2817721.2817747

Berard and Blanch recently introduced an automated method of touch latency measurement [1]. It is based on the extraction of the finger position as well as the display frame number by computer vision processing of the images captured by the camera. The approach provides an accurate estimation of the physical position of the finger. In addition, it allows the automated computation of latency on many samples, which results in an accurate estimation of the average latency, but also provides the spread of the latency. However, the approach requires an external camera, several calibration steps, and complex software, which make it unsuitable for quick latency estimations.

Rather than using an external camera to measure the physical position of the finger, the system may require an operator to follow a predefined trajectory with the finger. Hence, the system does not have to actually measure the position of the finger, it simply assumes that the finger is where it should be. The idea was introduced and implemented by Berard and Blanch [1] in their low overhead approach (or "LO approach"). This approach saves the requirement for external equipment and complex calibrations, and it can be implemented with minimal efforts. However, Berard and Blanch used the tracing of a circle at constant speed as the predefined trajectory. They report that the task was difficult to perform by the operators, which resulted in a range of latency estimations across participants having a width of 8ms. They also report that the average estimation was offset by 2ms compared to the ground truth, which gives an indication of its accuracy. Our approach builds on the idea of benefitting from the help of operators to save the complexity of external equipment, but we chose a task that is simpler to perform in order to improve the precision of the measure.

A predictive approach to measure the latency of the system was introduced by Knibbe et al. [4]. Using a projector-camera system, a trajectory prediction improved the accuracy of the real-time projection on juggling balls. With no assumption on its own latency, the system is able to calibrate itself: starting from 0ms, the prediction length is incremented (by 33ms and then by 5ms for refinement) and the projection error is measured at each step until a local minimum is found. The prediction length when the error is minimal is then considered as the end-to-end latency of the system. We follow a similar approach, but we adapt it to touch surfaces, and we ask human operators to match the object and the feedback, which saves the requirement for external cameras and calibration.

## A NEW MEASUREMENT OF TOUCH LATENCY

Predicting the finger trajectory can reduce the perception of latency and improve user performances [2]. With prediction, the feedback is not displayed at its last known position but at a predicted position so as to compensate the system lag. A perfect prediction can theoretically compensate the latency totally. Reasoning backwards, if the prediction length is adjusted until the feedback exactly matches the physical finger, then the prediction length is equal to the system's latency. We use this as a means to measure the latency of the system.

Our approach consists in showing a simple vertical line to users that follows the index finger on the touch surface. Users can tune the prediction length until they observe the best match between the position of their finger and the position of the feedback. The prediction length is used as the latency estimation. Compared to the "LO approach" [1], users' motion is less constrained by the system. Users must perform straight line trajectories at the (constant) speed of their choice, which is arguably easier than tracing circles at a constant speed imposed by the system.

Ng et al. have showed that users are more likely to notice difference of latency when dragging small objects compared to bigger ones [5]. Therefore, we minimize the size of the dragged object, to improve users' perception accuracy: we display a 1-pixel line (0.28mm). To ease the matching between the line and the finger, a line is also drawn (with a pen) on users' index finger, and the positions are matched on a static finger.

This predictive approach provides accurate latency estimation only if the prediction model is correct. However, predicting the future is a difficult task prone to errors. We use a simple constant speed linear model for the prediction [2]. At each display cycle $i$, a predicted position $\mathbf{\hat{x}_i}$ is computed as

$$\mathbf{\hat{x}_i} = \mathbf{x_i} + L * \mathbf{\hat{s}_i} \qquad \mathbf{\hat{s}_i} = \frac{\mathbf{x_i} - \mathbf{x_{i-1}}}{\Delta t} \tag{1}$$

where the $\mathbf{x}$ are the observed finger positions, $\mathbf{\hat{s}_i}$ is the speed estimation and $\Delta t$ is the display cycle time. The prediction length L is tunable and is used to measure the latency.

By design, this model perfectly predicts the finger position only when the finger is moving at constant speed in straight lines. But actual user trajectories are neither perfectly straight nor performed at perfectly constant speed. We remove the *straight line* problem by working in 1D: we ask users to make mostly horizontal translations on the surface, and display a vertical line at the x-position of the prediction. The *constant speed* constraint is more problematic: every user trajectory has an acceleration and a deceleration phase, which respectively induce undershoots and overshoots of the prediction model. We thus ask users to focus on the finger-feedback alignment in the middle of their motion, and ask them to stabilize their finger's speed as much as possible. Even so, the speed is never perfectly constant. However, a user choses the best finger-feedback alignment after several translation gestures in both directions. We assume that the prediction error averages across these gestures and allows for an accurate latency estimation. The accuracy is assessed in the user experiment presented in the next section.

In addition, large system latencies require large prediction lengths, which result in larger prediction instabilities. We thus envision that the precision of our approach will decrease with increasing latency of the system. In the user experiment, we evaluate two different levels of latency to study this phenomenon.

## VALIDATION: USER EXPERIMENT

We designed a user experiment to assess the precision and accuracy of our approach. We also compared our approach to

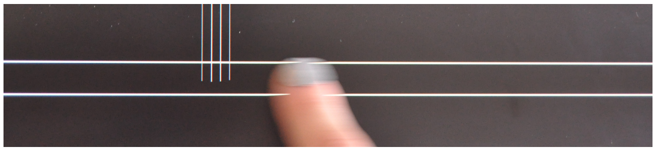

**Figure 1. A user executing the predictive approach measurement: the finger, equipped with an optical marker, moves between the two guidelines. Users have to tune the prediction length to make the feedback match their finger when moving. Several vertical lines appear because of the exposure time of the camera.**

the “LO approach”, considered as the state of the art in quick latency measurements that require no external equipment.

8 participants took part in the experiment, with mean age = 26.75 [21-33], and including 3 women. All the participants came from our lab and had previous knowledge in computer science but no experience with neither the LO approach nor the predictive approach.

Our experiment was performed on a custom device that allowed us to control its latency with precision, and to test a low level of latency (30ms). However, it should be clear that the approach is not bound to this specific hardware and can be easily implemented on any commercial tablet for example. We used the same high performance touch system as in [2]. The system uses 120Hz optical tracking and requires that an optical marker be taped to the user’s fingernail. The precision of the sensing was measured at less than 0.1mm (subpixel). We used the “high accuracy” approach [1] to measure the ground truth latency of our system at the various studied levels.

### Task

In the LO task, we used a wheel similar to the one used by Berard and Blanch [1]: we displayed a circle of radius 8.8cm with a spinning radius at constant speed $s = 3 rad s^{-1}$. Participants were requested to follow the intersection of the radius and the circle. Pressing the space bar of a keyboard began and ended the recording. We asked participants to perform at least one and a half of a circle before ending the measurement.

In the predictive task, a white 1 pixel vertical line was displayed at the x-position of the finger. While the finger was static on the surface, we asked participants to draw a line on their fingernail that matched the feedback line on the screen. As illustrated in Figure 1, two white horizontal lines were displayed, and we asked participants to try to keep their finger within these lines when moving. Participants could tune the prediction length with the keyboard: two keys were used for ±1ms increments, two others for ±0.1ms increments. Participants were asked to tune the prediction length until the line on the finger and the feedback line matched when executing horizontal movements on the screen at a speed as constant as possible.

### Protocol

The experimental design includes two factors: METHOD was either *LO* (following the LO approach) or *PRED* (following our prediction approach); LATENCY was either 30ms or 80ms.

Participants were asked to execute 5 measurements in each of the 4 conditions (2 METHOD x 2 LATENCY). We thus recorded 160 latency measurements (8 participants x 4 conditions x 5 repetitions). Presentation of conditions was ordered by METHOD first, then LATENCY. The order of presentation of METHOD and LATENCY was balanced across participants as a means to equilibrate a potential order effect. Post study ANOVA tests confirm no significant effect of ordering. At the beginning of the experiment, we explained both methods to participants and gave them a demonstration.

When executing the LO method, participants were first given 1min to train to following the wheel. Then, they executed five measurements in the first latency condition and five measurements in the second one. To encourage participants to improve their performances, standard error was displayed at the end of each measurement.

When executing the predictive method, participants were asked to move at the most constant speed as possible from one side of the screen to the other and to focus on the middle of their motion when the speed is the most constant. The prediction length, hidden to participants, was set to 0ms at the beginning of the run. They were given as much time as they wanted to tune the prediction length. The task was more easily performed using the *key repeat* feature of the keyboard, which is hardly reproducible, and ensured that participants did not simply repeat the same keyboard actions at each measurements. When they were satisfied with the feedback matching their finger, the final prediction length was recorded and reset to start another evaluation.

### Results and discussion

Finger’s speed during prediction measurements was variable amongst participants. We observed a trend of increasing speed during a single measurement. When the finger moves faster, it is easier to perceive a gap due to uncompensated lag; yet, speed also increases prediction error (as it is difficult to maintain a constant high speed) and creates a blurry effect on the feedback that makes difficult to match the lines. Globally, speeds peaked around 1.1m/s, where participants found a tradeoff between precision and blur.

On the 160 latency measurements, we removed 10 within-subject outliers that were 1.5 interquartile either above the upper quartile, or below the lower quartile. Measurements of participant 2 at 80ms with the LO method were around 35ms and very variable. This whole condition was considered as an outlier compared to the other results and was discarded from the study, given 145 measurements left for analysis.

The within subject variability of the estimation was greater with the LO method compared to the prediction method, as illustrated on Figure 2. The standard deviation of the five measurement repetitions, averaged across participants and the two latencies was significantly smaller with the predictive method (1.6ms) than with the LO method (8.4ms) ($F_{1,7} = 21.9, p < .005$).

Estimating the latency with the prediction method takes more time (between 30s to 1min depending on participants) compared to a *single run* of the LO method (less than 10s). This

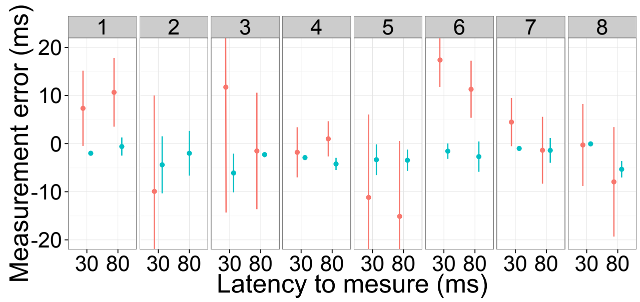


**Figure 2. Averaged measurement error for each participant at 30ms and 80ms latency with the LO method (red, left) or the predictive method (blue, right). 95% confidence intervals indicate the within subject variability.**

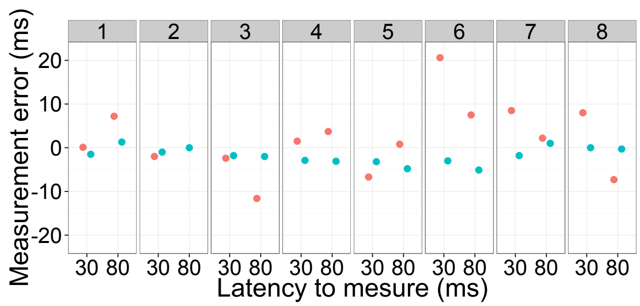


**Figure 3. Error of the most stable LO run (red, left) compared to the error of the first predictive measurement (blue, right) for each participant at 30ms and 80ms latency.**

is due to the multiple adjust and observe steps needed in the predictive approach. However, Berard et Blanch recommend that multiple runs of the LO method are performed by an operator in order to select the best run, defined as the one with lowest standard error. We thus compare a single run of the prediction method (the first run) to the best of five runs of the LO method, so that the operator time was similar. The LO method was found to be surprisingly accurate when averaged across participants: the best measurement average was 33.5ms at 30ms and 80.4ms at 80ms. The predictive approach gives a similar accuracy of 28.1ms at 30ms and 78.4ms at 80ms. However, the confidence interval for the mean of the measurement error (taking both 30ms and 80ms values) is narrower with the predictive approach (2.0ms) than with the LO approach (8.7ms). This indicates that a quick 2min latency measurement will be as *accurate* and more *precise* with the predictive method compared to the LO method.

With the LO method, the results uncover a high variability within and between participants. This could be explained by the limited training that the participants performed. The task certainly required more training to be mastered. The LO method provides good accuracy when averaged over several users. However, our results indicate that using the estimation of a single user could lead to large errors. On the contrary, the predictive method provides better accuracy and more precise results when considering the measurement of a single user.

The predictive measurements appear to suffer form a negative offset, compared to the ground truth. It can be due to the prediction length zeroed at the beginning of the measurement. A convergence of the prediction length from lower value could have lead to under-estimation. In further studies, we will set different prediction values at the beginning of each evaluation.

Contrary to our hypothesis, the predictive method did not yield less precision at 80ms than at 30ms. One explanation is that the human eye seems to be good at averaging the blurred feedback caused by the prediction instabilities.

When asked, "Which approach was the easier to perform?", 6 participants answered that they preferred the predictive approach. They argued that the wheel was hard to follow as they were constrained by the system. By contrast, the predictive approach allowed them to adopt their own rhythm.

Further evaluation will be needed to validate our approach at lower sensing rate. Even if 120Hz sensing is already available in commercial devices (for example, on the iPad Air 2), most current commercial devices have lower sensing frequencies in the range [60Hz-85Hz]. This could lead to more unstable predictions. However, the similar results of our experiment at 30ms and 80ms of latency indicate that users are good at coping with blurry feedback due to unstable prediction.

In conclusion, we introduced a new latency measurement approach using prediction. It is easy to implement and does not need any external device. A user study indicates that our approach is as *accurate*, more *precise*, and easier performed by users compared to previous work.

## ACKNOWLEDGEMENT

This work has been partially supported by the LabEx PERSYVAL-Lab (ANR-11-LABX-0025-01).

## REFERENCES

1. François Bérard and Renaud Blanch. 2013. Two Touch System Latency Estimators: High Accuracy and Low Overhead. In *ACM ITS*. ACM, 241–250.
2. Élie Cattan and François Bérard. 2015. Reducing Latency with a Continuous Prediction: Effects on Users' Performance in Direct-Touch Target Acquisitions. In *ACM ITS*. ACM.
3. Ricardo Jota, Albert Ng, Paul Dietz, and Daniel Wigdor. 2013. How Fast is Fast Enough? A Study of the Effects of Latency in Direct-Touch Pointing Tasks. In *ACM CHI*. ACM.
4. Jarrod Knibbe, Hrvoje Benko, and Andrew D Wilson. 2015. Juggling the Effects of Latency: Motion Prediction Approaches to Reducing Latency in Dynamic Projector-Camera Systems. *Microsoft Research Technical Report MSR-TR-2015-35* (2015).
5. Albert Ng, Michelle Annett, Paul Dietz, Anoop Gupta, and Walter F. Bischof. 2014. In the Blink of an Eye: Investigating Latency Perception During Stylus Interaction. In *ACM CHI*. ACM, 1103–1112.
6. Albert Ng, Julian Lepinski, Daniel Wigdor, Steven Sanders, and Paul Dietz. 2012. Designing for low-latency direct-touch input. In *Proceedings of ACM UIST*. ACM, 453–464.