Juggling the Effects of Latency: Software Approaches to Minimizing Latency in Dynamic Projector-Camera Systems

Jarrod Knibbe[1,2], **Hrvoje Benko**[1], **Andrew D. Wilson**[1]

[1] Microsoft Research, Redmond, USA
{benko,awilson}@microsoft.com

[2] University of Bristol, UK
Jarrod.Knibbe@bristol.ac.uk

Figure 1. The Juggling Display; projection alignment on fast-moving objects through software-based latency reduction.

ABSTRACT

Projector-camera (pro-cam) systems afford a wide range of interactive possibilities, combining both natural and mixed-reality 3D interaction. However, the latency inherent within these systems can cause the projection to 'slip' from any moving target, so pro-cam systems have typically shied away from truly dynamic scenarios. We explore software-only techniques to reduce latency; considering the best achievable results with widely adopted commodity devices (e.g. 30Hz depth cameras and 60Hz projectors). We achieve 50% projection alignment on objects in free flight (a 34% improvement) and 69% alignment on dynamic human movement (a 40% improvement).

Author Keywords

Projector-camera systems; latency; prediction

INTRODUCTION

Latency in projector-camera (pro-cam) systems can easily result in projection 'slipping' from a moving target. As a result of this, the use of pro-cams in truly active scenarios has been avoided. Previous work exploring latency reduction has primarily focused on advanced hardware-based approaches (e.g. [5]). This approach is at odds with the lightweight, commodity-hardware based development style favored by both the research and enthusiast communities (i.e. 30Hz depth-cameras with 60Hz projectors, as in [1, 6] for example). We explore software only approaches to latency reduction that can be applied to any hardware configuration, in order to facilitate pro-cam usage in dynamic settings. We report on a preliminary investigation of projection on both objects under free flight and on-the-body of fast moving users.

RELATED WORK

Pro-cam systems have long been a popular tool for the HCI community (e.g. [1, 6]). However, latency is a problem for the experience; constraining movement speeds [6] or resulting in image misalignment, as can be seen in the video for [1]. As a result of latency effects, research has considered hardware-based solutions, such as high-speed cameras [5] and multi-camera setups [4]. For example, Lumospheres projects on balls in free flight [4]. Using 6 cameras operating at 120Hz, Lumospheres tracks and predicts the balls' future position. With this technique, Lumospheres achieves on-object projection in all frames, with 80% projection accuracy. We employ similar principles using only one commodity 30Hz depth-camera to explore achievable projection accuracy.

Other works have considered software-based motion modelling across a range of domains, such as projectile motion in sports tracking [3]. We draw upon this as one further motivation for our work.

MOTION PREDICTION FOR LATENCY REDUCTION

In order to reduce the effects of latency and increase projection alignment we use motion prediction to model and calculate the future states of objects (similarly to [4].)

In our work we consider two different scenarios for latency reduction. In the first, we explore *projection on objects under free flight*. In the second, we explore *projection on bodies in dynamic motion*. Our system uses a Kinect camera, a commodity 60Hz projector and a Windows 7 laptop. We measured the latency of 9 different projectors when paired with a 30Hz Kinect and calculated an average latency of 102.5ms (+/- 6ms std. dev). Alongside latency effects, other sources of projection slip also exist in pro-cam systems, such as camera calibration errors and rolling shutter effects (see [2] for a discussion) and pro-cam synchronization errors.

Objects in Free Flight – Predictable Motion

We developed a *Juggling Display*. In standard 3 ball juggling the balls can easily exceed 4m/s, resulting in 40cm of projection slip given ~100ms of system latency. At the zenith of the ball's flight, as speeds decrease, fleeting alignment occurs, resulting in only ~14% of the balls' flight being illuminated without latency correction.

UIST '15 Adjunct, November 08-11, 2015, Charlotte, NC, USA
ACM 978-1-4503-3780-9/15/11.
http://dx.doi.org/10.1145/2815585.2815735

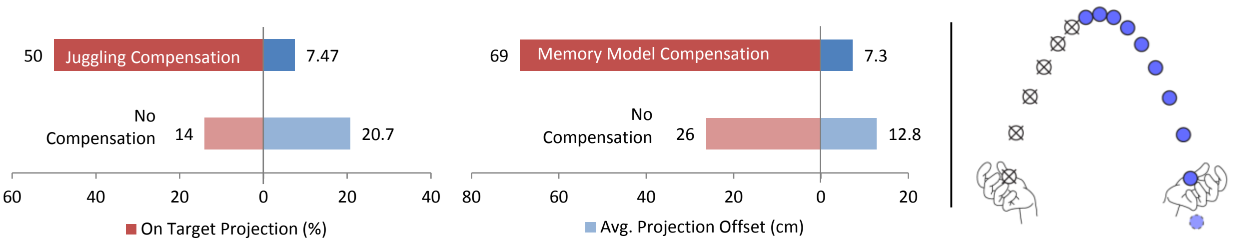

Figure 2. Left: Objects in free flight evaluation results. Mid: On-body projection evaluation results. Right: Blue balls indicate frames available for projection alignment (after Kalman filter initialization, 3 frames, and 100ms prediction step, 3 frames).

As the balls' location at any time during flight can be calculated based on a set of physical laws, we term this *Predictable Motion*. We fit a projectile motion model to a Kalman filter to smooth our tracking values (from the depth camera) and predict the balls location into the future (to account for latency). We specify low values for process noise, but high uncertainty for observation (to account for the rolling shutter and calibration inaccuracy in our camera).

Preliminary Evaluation

Three jugglers performed 'standard' 3-ball cascade juggling for 2 minutes under no latency correction (condition 1) and full latency correction (condition 2). The juggler's used softballs (9.7cm diameter) and stood 2.5m from the pro-cam unit. We provide 2 accuracy measurements; average projection offset captured automatically using image processing, and a binary projection 'aligned vs. missed' measure based on all frames with an estimated >10% projection alignment.

With no latency correction, we found that 14% of balls in flight are projected upon and that the average projection offset is 20.7cm (Figure 2). With latency correction, we achieve 50% projection accuracy and an average offset of 7.47cm.

On-body Projection – Semi-Predictable Motion

Human motion includes a wide range of patterns and repetition, for example in walking, dancing and athletic performance (e.g. the basic performance of different shots in tennis) [7], thus we term this *Semi-Predictable Motion*.

In order to predict bodily motion, we train a motion lookup model based on a person's previous movements. After a brief training window (20secs), our lookup model can take into account the intricacies of personal performance, such as acceleration patterns, maximal reach and personal style. The performer's ongoing movement details continue to further train the model for ongoing prediction.

Preliminary Evaluation

Three participants performed up-down and circular hand movements at 1.5m/s, as guided by the system. Each participant performed each movement for 15s. A 10cm graphic was projected onto the hand's location, as tracked through the Kinect Skeletal tracker.

With no latency correction, only 26.3% of projected frames fall on target (with an average offset of 12.8cm.) With our motion model approach, 69% of frames fell on target, with an average offset of 7.3cm.

DISCUSSION AND CONCLUSION

At 50% alignment for objects in free-flight and 69% alignment for on-body projection, our results demonstrate an improvement in projection alignment of 34% and 40% respectively. These results are constrained by a couple of key challenges. Firstly, a Kalman filter requires a few frames of data to initialize and smooth sensor values, and combined with a prediction step of 100ms (per system latency), this results in approximately 6 or 7 frames of missed projection in a 15 frame ball flight (observed average) – suggesting an achievable maximum of 60% with our approach. We are keen to explore further machine learning techniques to reduce the initial missed frames and increase the achievable maximum projection. Secondly, at 4m/s a juggling ball travels 13.2cm between (non-linearly sampled) camera frames and measures only 22 pixels across in camera space (at 2.5m) – resulting in position measurement noise that detracts from projection alignment. Similarly in the on-body scenario, at speeds greater than 1m/s the Kinect skeletal tracker becomes increasingly inaccurate, effecting our on-body projection alignment.

When combined with the effects of visual motion blur, we believe our results are compelling and we encourage the reader to view our associated video. Although these results still fall short of the projection accuracy achievable through hardware augmentation (such as an estimated average offset of 2.91cm, derived from [4]), we believe they form a good launch point for further exploration in this area and hope this will motivate further work on software-only approaches to latency reduction.

REFERENCES

1. Harrison, C. et al. OmniTouch: Wearable Multitouch Interaction Everywhere. *UIST '11*, ACM (2011).
2. Khoshelham, K. Accuracy analysis of kinect depth data, in archive of ISPRS, (2011), 133–138.
3. Kitani, K. et al. BallCam!: Dynamic View Synthesis from Spinning Cameras. *UIST Adjunct '12*, ACM (2012), 87–88.
4. Koike, H. et al. LumoSpheres: Real-time Tracking of Flying Objects and Image Projection for a Volumetric Display. *AH '15*, ACM (2015), 93–96.
5. Okumura, K. et al. Lumipen: Projection-Based Mixed Reality for Dynamic Objects. (2012), 699–704.
6. Sodhi, R. et al. LightGuide: Projected Visualizations for Hand Movement Guidance. *CHI '12*, ACM (2012).
7. Coupled Oscillators and Biological Synchronization. http://www.scientificamerican.com/article/coupled-oscillators-and-biological/.