Can gamers detect cloud delay?

Kjetil Raaen
Westerdals — Oslo School of Arts, Communication and Technology
Faculty of Technology
Simula/University of Oslo
Email: raakje@westerdals.no

Ragnhild Eg
Simula Research Laboratory
1325 Lysaker, Norway
Email: rage@simula.no

Carsten Griwodz
Simula Research Laboratory
1325 Lysaker, Norway
University of Oslo
Email: griff@simula.no

***Abstract*—In many games, a win or a loss is not only contingent on the speedy reaction of the players, but also on how fast the game can react to them. From our ongoing project, we aim to establish perceptual thresholds for visual delays that follow user actions. In this first user study, we eliminated the complexities of a real game and asked participants to adjust the delay between the push of a button and a simple visual presentation. At the most sensitive, our findings reveal that some perceive delays below 40 ms. However, the median threshold suggests that motor-visual delays are more likely than not to go undetected below 51-90 ms. These results will in future investigations be compared to thresholds for more complex visual stimuli, and to thresholds established from different experimental approaches.**

I. Introduction

Games resemble real life in many ways, and players expect them to behave like the physical world. There, most actions we perform lead to instant reactions. Unfortunately, because of technical restrictions, immediate results are impossible in computer games. To study how we perceive these delays between input actions and visual results, we apply work on sensory interactions to current gaming shortcomings imposed by network limitations.

Our sensory systems process many external stimuli at the same time, but somewhere along the way they converge and align to create a unified experience. For instance, in many human-computer interactions, a button push and a visual event have to coincide in order to ensure fluent operations. Fortunately, they need not be in perfect synchrony. Humans learn from experience what to expect following a familiar action, moreover, the perceptual system can compensate for, and even adapt to, short time displacements [1]. Causality is an important factors in motor-visual interaction, if too much time passes after an action, the delayed consequence may be attributed to another event [2]. Online games introduce concerns related to causality and anticipation, considering that gamers expect immediate reactions from their actions. Unfortunately, network limitations slow down the reaction time, creating temporal delays between the motor and the visual signals. Eventually, the delays become too long for the perceptual system to compensate, and they become detectable.

Multiplayer games communicating over a network exhibit two types of delay. *Interface delay* is the most critical, but shortest, occurring between a user's input and the resulting visual presentation. Further, when games are played across a network, information exchange between client and server, in addition to processing on the server, introduces delays. This type of *network delay* can extend to tens or even hundreds of milliseconds, and developers strive to hide it using various techniques [3]. Recently, a new way of delivering games has gained popularity: *Cloud gaming*, the concept of running the entire game remotely and using the local computer as a dumb terminal. In this scenario, network delays appear between input and output, and add to the interface delay. Latency hiding techniques have become much more difficult to implement, which highlights the importance of understanding users' basic latency tolerance and ability to detect it. Jarschel and colleagues [4] ran a study on player sensitivity to latency in cloud gaming, but their manipulations included no latencies shorter than 80 ms. Because subjective quality of experience was noticeably reduced at this value, their results emphasise mainly how sensitive players are to fairly short delays.

Psychologists typically tackle their research questions using controlled experimental designs that heed statistical power. On the other hand, game developers tend to use rules-of-thumb and estimates based on experience when working to reduce lag in networks and computational processes. In the study of real game scenarios, stringent experimental methods with repeated presentations could require participants to spend hours on a single experiment. To circumvent this, earlier studies on in-game delays have applied the experimental method to real games, and have instead restricted the number of participants [5]. With a smaller pool of participants to average across, generalisations come with a note of caution. Our motivation is to empirically establish thresholds for detectable motor-visual delays, and in this first step we explore their temporal interaction in isolation.

II. Experiment design and method

We designed and conducted a behavioural experiment that uses a button device and a simple visual stimulus, and we ran it on standard computers[1]. The button devices are Griffin click+spin USB controllers, or jog shuttles, which are designed solely for button-pushes and left/right rotations[2]. The visual stimuli consisted of a black disc on a white background that would flash on or off in response to the button pushes. We wanted to explore object size as an additional factor that could shape the perceptual process. Thus we included both a small and a large disc, with diameters of 20 and 200 pixels.

[1]HP Z200 (Intel Xeon X3430 CPU 4 cores@2533MHz, 8GB RAM) running Windows 7 and connected to Acer AL1916W monitors with1440x900 screen resolution and 60 Hz refresh rate

[2]http://store.griffintechnology.com/powermate

978-1-4799-6882-4/14/$31.00

Fig. 1. Illustration of the experimental set-up with the USB controller placed conveniently in front of the participant and the visual stimulus presented on the monitor. The visual stimulus is here represented by a 20 pixel black disc.

At a viewing distance of 60 cm, these disc sizes correspond to visual angles of 5.6° and 0.56°, respectively. During the experiment, a visual guideline in the form of a simplified clock face helped participants keep track of rotations made with the USB controller. An example of the experimental set-up is visualised in Figure 1. The initial delay at which stimuli were introduced varied between 200 ms, 300 ms, and 400 ms. Each stimulus condition was repeated four times, so that participants completed a total number of 24 trials, in addition to two initial practice trials. The experiment took approximately 10-15 minutes to complete, including an initial questionnaire that assessed possible background variables, such as gaming experience.

We recruited 13 female and 28 male volunteers, aged between 19 and 43 years (mean = 24), and conducted the experiment in a computer lab at the Norwegian School of IT. Participants were instructed to click the knob of the USB controller in order to make the disc presented on the monitor flash on or off, and to rotate the controller when they wanted to adjust the temporal delay between the push and the disc-flash. The delay would change proportionally to the adjustment angle. Because the controller provided no reference points, participants were always unaware of the value of the delay and the direction of their adjustments. Values for delay could be adjusted from 0 to 500 ms; if adjusted past these extremes, the values would gradually decrease or increase away from the extreme. We emphasised that they could spend as much time and make as many pushes and adjustments as they wanted. When they were satisfied that they could no longer perceive the temporal delay, they proceeded to the next trial by pressing the spacebar.

An experimental set-up that depends on a computer system is bound to be influenced by physical and computational limitations. We highlight these limitations in order to explain the precautions taken. First of all, the screen refresh rate of 60 Hz could introduce up to 16.7 ms visual lag. Assuming a random distribution of clicks between screen updates, this corresponds to an average of 8.3 ms delay. Clicks collected by the USB controllers also have limited temporal resolutions, the experiment machines polled these at a rate of 125 Hz. This equals a maximum delay of 8 ms and an average of 4 ms. In total, these uncertainties add up to 12.3 ms average and 24.7 ms maximum technical delay. An additional uncertainty relates to the experimental task. A standard approach would be to adjust delays continuously downwards, meaning that the exact detection threshold could be surpassed. Accepted values may therefore correspond to a point below the threshold. In the worst case, our collected data may be uniformly distributed over a region of imperceptible delays. Finally, the controller rotations adjusted the delays at 25 ms increments, which leads to a clustering of the collected data. These concerns are all addressed in the interpretation and presentation of our findings.

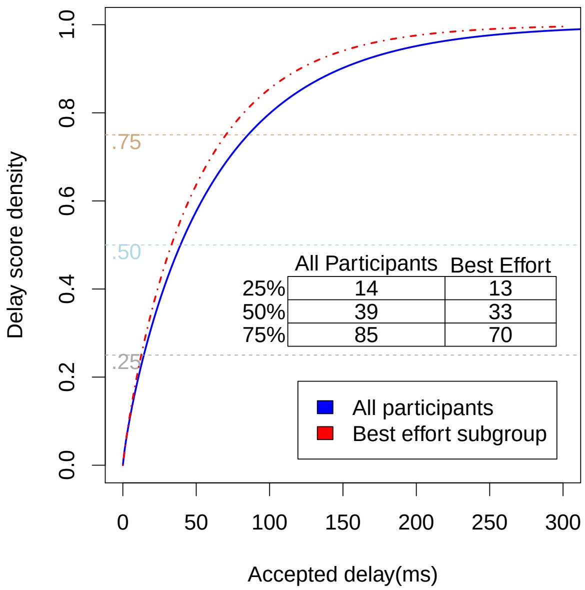
III. Results

In our study of the perception of motor-visual delays, we designed and conducted an experiment where participants adjusted the delay between a button-push and a visual disc flash. We also explored the relationship between accepted delay scores and the size of the presented disc, as well as participants' gaming experience.

Any initial delay value that participants accepted without adjustments was categorised as an accidental accept; accidental accepts were thus labelled as missing values and treated like such for the main analyses. We also judged that participants who made two or more of these accidents (corresponding to approximately 5% of all trials) did not adhere to the experimental procedures and we therefore excluded the data from two participants from the analyses.

We explored potential effects with Wilcoxon signed rank-sum tests, but found no significant differences in accepted delays due to disc size ($W = 660, n = 38, Z = 0.06, p > .5$) or to gaming experience ($W = 89189, n = 893, Z = -0.46, p => .3$). However, we observed great differences in task effort between participants, defined by the number times they pushed the button. Using the number of button-pushes as an estimate for the number of adjustments made, we ran a linear regression with a logarithmic fit. This revealed a significant, but negative, relationship between task effort and accepted delays ($accepted\ delay = -19.96 * log(clicks) + 118.14, R^2 = .04, p < .001$). In other words, frequent adjustments led to lower delay values. Furthermore, we found that experienced gamers made significantly more adjustments than those with less experience ($W = 66254, n = 893, Z = -5.31, p < .001$), meaning that experienced gamers make more attempts than less experienced gamers before finding an acceptable motor-visual delay. Accordingly, we surmised that the accepted delays depended in part on the number of adjustments made, contributing to greater temporal sensitivity among those most dedicated to the task. In light of this finding, we decided to run separate analyses for the full participant group and for the best-effort subgroup. The latter group was defined by their average number of adjustments, and it includes all participants who made more than the median number of 18 button-pushes during a trial.

To establish detection thresholds for motoric-visual temporal delays, we plotted an empirical cumulative density distribution with all delay scores and then established the best-fitting gamma distribution. From this distribution, we derived the 25th, 50th, and 75th percentiles. We did the same for the best-effort sub group. The distributions and percentiles are portrayed in Figure 2. When interpreting these results, keep in mind the limitations outlined in section II. Notably, even when taking these cautionary measures, the 25th and 50th thresholds still fall below 100 ms.



	All Participants	Best Effort
25%	14	13
50%	39	33
75%	85	70

Fig. 2. Cumulative density of participants' accepted delay scores plotted with the best-fitting gamma distributions. The thresholds listed in the table correspond to the 25th, 50th and 75th percentiles of the distributions.

IV. Discussion

This work presents the first step in our studies of the interplay between human actions and digital events, where we aim to establish thresholds for detectable motor-visual delays. The applicability of these perceptual thresholds is particularly prominent in current challenges facing the gaming industry.

In our first venture, we explored the detectability of these delays under the most ideal conditions, using simple, isolated stimuli. We also considered individual differences, although we found no correlation between gaming experience and accepted delays. Similarly, we found no effect related to the size of the visual object. Instead, our results showed that some participants made more adjustments during a trial than others, and this effort was reflected in decreased values for accepted delays. Moreover, experienced gamers had a greater tendency to make more adjustments than non-experienced gamers.

Because the experiment task allowed participants to adjust delays below the actual point of detection, we prefer to err on the side of caution. Our cautionary measures take into account that the collected data could be uniformly distributed across the individuals' detection ranges, as well as the average 12 ms delay added by our system.

When incorporating these limitations into our motor-visual delay thresholds, we find that a small share of participants cannot perceive visual lags shorter than 97-182 ms, while those whose scores fall below the median are able to detect delays around 51-90 ms. The most sensitive of our participants could even perceive visual lags as short as 26-40 ms. Because of the strong correlation between task-effort and accepted delay, we included a separate analysis for participants who made more adjustments than the median of 18. For this half of participants, the median accepted delay lies between 45-78 ms. Although these thresholds allow room for uncertainty, they serve as guidelines to the sensitivity of the human perceptual system when encountering motor-visual delays. Importantly, they suggest that a large proportion of our participants can easily perceive delays shorter than 100 ms. Moreover, for one out of every four trials, our participants could even detect delays below 40 ms.

In cloud gaming scenarios, all network latency appears as interface latency. The presented experiment tackles this temporal challenge using a button-device and a simple visual presentation, and herein lies the most pronounced difference between real-life cloud games and our experimental scenario. Most games involve reactions to moving stimuli, and this points the direction for our next step in the study of motor-visual delay perception. For the time being, the isolation of the motor-visual interaction allowed us to investigate how much delay people can detect when they are at the most sensitive. By presenting game developers with a lower bound for acceptable motor-visual delays, we hope to introduce a level of confidence in the development stage. Below the shortest of our established thresholds, gamers are very unlikely to consciously experience any visual lags. On the other hand, gaming performance may still be negatively affected by imperceptible delays, which is another issue we aim to address in future works.

V. Conclusions and Future Directions

If our participants are representative of a larger population of regular computer users, we can assume that half of these individuals will have a difficult time tolerating services that operate with up to 100 ms delay. At the most sensitive, some of these people are also able to detect motor-visual delays as short as 26-40 ms. Consequently, providers of cloud gaming services should bear in mind that some of their players could be very sensitive to the visual consequences of network latency.

With this foundation, we now have a scale of noticeable delays to build our work on. In the next planned step, we will compare experimental methodologies to ensure that the established thresholds do not merely reflect the assigned task. We also plan to apply more dynamic and more complex stimuli, to explore whether this could alter the detection of motor-visual delays. Our end objective is to conclude the project with a fully operational game; this will build the foundation for an investigation into the ability of game players' to compensate for lags.

Acknowledgements

We acknowledge the valuable feedback provided by Professors Mark Claypool, Pål Halvorsen, and Dawn Behne.

References

[1] C. Stetson, X. Cui, P. R. Montague, and D. M. Eagleman, "Motor-sensory recalibration leads to an illusory reversal of action and sensation." *Neuron*, vol. 51, no. 5, pp. 651–659, Sep. 2006.

[2] J. Heron, J. V. M. Hanson, and D. Whitaker, "Effect before cause: Supramodal recalibration of sensorimotor timing," *PloS One*, vol. 4, no. 11, p. e7681, Jan. 2009.

[3] Y. Bernier, "Latency compensating methods in client/server in-game protocol design and optimization," *Game Developers Conference*, vol. 98033, no. 425, Mar. 2001.

[4] M. Jarschel, D. Schlosser, S. Scheuring, and T. Hoßfeld, "An Evaluation of QoE in Cloud Gaming Based on Subjective Tests," *Innovative Mobile and Internet Services in Ubiquitous Computing*, pp. 330–335, Jun. 2011.

[5] M. Claypool and K. Claypool, "Latency and Player Actions in Online Games," *Communications of the ACM*, vol. 49, no. 11, pp. 40–45, Nov. 2005.