19th IEEE International Symposium on Robot and
Human Interactive Communication
Principe di Piemonte - Viareggio, Italy, Sept. 12-15, 2010

Effect of Delay on Dynamic Targets Tracking Performance and Behavior in Virtual Environment

Vittorio Lippi, Carlo Alberto Avizzano, Denis Mottet, Emanuele Ruffaldi

Abstract— **We analyze the impact of the delay between the sensors and the visualization on the performance at a catching task in virtual environment. Testers have been asked to interact with a virtual environment displayed on a 2D screen in front of them, catching, with the virtual hand, virtual balls falling from the top of the screen. Two thresholds have been inferred: a *behavioral threshold delay* beyond which the behavioral organization is sensibly affected and a *performance threshold delay* beyond which task performance begins to worsen. We found a drop in performance when the total delay is over 70-80 ms, yet without significant changes in the organization of successful catches. Although the presence of a threshold on behavior is still to be further investigated, these two thresholds give a practical indication for the design and the validation of virtual reality based training systems. This is especially important in the context of training human skills in virtual environment: in particular to assess the quality of a training platform simulating the three ball cascade juggling pattern.**

I. INTRODUCTION

Virtual environments could be powerful training platform for real tasks. The training simulators not only can simulate the virtual environment of the training situation, but this environment can be enhanced via multimodal renderings, fostering the training process and accelerating the training effect [2]. In a simulated environment is possible to create the optimal conditions and the situations relevant to the task to be learned, to analyze variables not directly accessible in the real world, and to apply learning accelerator methods not possible real world physics, such as slowing down the simulation or giving visual hint [11]. On the other hand the differences between simulated and real world could lead to different behavioral schemes to achieve the same results and this can affect the quality of the training system. In particular, it is impossible to avoid some delay between sensors readings and visualization and this delay might impede performance. It is known that adding delay in visual feedback of one's movement decreases performance in pointing [13] or tracking task [18], even though humans can adapt their control and use anticipation strategies to keep performance at its best [21]. Anticipation allows to adapt to the delay so to minimize errors, but some changes in more subtle aspects of the behavior are often observed [7]. For example, when bouncing a ball on racquet in VR, the delay is not consciously perceived below 90-100ms [16], though behavioral changes in the phasing of movement occur for smaller delays [17]. Such behavioral changes can take the form of complex oscillatory behaviors induced by the time delay [1] or even delay-induced phase transitions leading to qualitatively different perception-action behaviors [20].

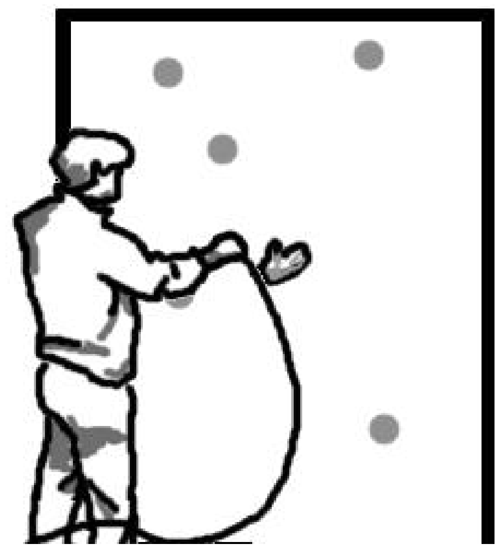

Fig. 1. System usage: the user stands close to a large screen controlling a virtual hand. The task consist in catching balls falling at a constant speed.

Though very brief, this review of previous works suggest that the effect of increasing the delay could be governed by two thresholds:

- a *behavioral threshold*: increasing the delay has no effect up to this first threshold. If the delay is increased, behavior is successfully adapted to sustain performance level (e.g., using simple *Smith predictor* [15] or by integration of the delay in a dynamic synchronization process [19]).
- a *performance threshold*: increasing the delay has no significant effect on *performance* up to this second threshold, but the *behavior* is differently organized [13], [18], [5], [17]. If the delay is increased, the performance starts to drop, because behavioral adaptability limits are attained.

The presence of the *performance threshold* is rather obvious and it became a well documented issue that is critical for the design of VR and teleoperation systems [5]. Conversely, the presence of the *behavioral threshold* is often supposed but

This work was supported by EU SKILLS Integrated Project

Vittorio Lippi, Emanuele Ruffaldi and Carlo Alberto Avizzano are with PERCeptual RObotics Laboratory, Sant'Anna Superior School of Advanced Studies, Pisa. v.lippi@sssup.it e.ruffaldi@sssup.it carlo@sssup.it

Denis Mottet is with Movement to health (EA2991), University Montpellier 1, France denis.mottet@univ-montp1.fr

978-1-4244-7990-0/10/$26.00 ©2010 IEEE

far less documented. For example, tracking performance is found to be linearly dependent on the delay in the range 0-500 ms, but only partial conclusions are derived about behavioral changes [3]. Even when tracking a chaotic target, synchronization performance is proportional to delay in the range 0-1000 ms, but the variability of the behavior suggests a critical reorganization at some point between 400 and 600 ms [19].

II. EXPERIMENT

The goal of the present report is to test the hypothesis of that the usability of VR systems is a function of the delay, with a *behavioral threshold* over which behavioral adaptation allows to sustain performance and a *performance threshold* over which performance drops.

A. Subjects

Ten healthy young subjects volunteered to participate in the study and gave informed, written consent. All had normal or corrected to normal vision. The study was approved by the local ethics committee and complied with the Declaration of Helsinki for human experimentation.

B. Proposed Task

The users interact with a virtual environment displayed on a 2D screen in front of them. Their task consists in catching, with the virtual hand, all the virtual balls falling at constant speed from the top of the screen. A new ball is created each half second, with random horizontal position, and would fall for 3 seconds if not caught. A ball is caught when the hand is near enough, without any other requirement, as in the case of the Light Weight Juggling (see appendix A). When a ball is caught it disappears, hence giving feed back to the user. The users controls the virtual hand position by means of an electromagnetic tracker (Polhemus Liberty©) held in their real preferred hand.

C. Set Up

The experiment has been performed with the user standing in front of a large screen (200 x 150 cm). The displayed environment is in scale 1:1 to the real world in order to reduce as soon as possible the user's effort to map his/her hand movement to the position of virtual hand and focus on the tracking task. For the same reason the user performed the task as close as possible to the screen (touching it would have disturbed the electromagnetic tracker). The screen was retroprojected with a 60 Hz refresh rate. Note that the environment, although displayed in three dimensional graphic is two dimensional. This means that the virtual hand could move on a plane only (vertically and horizontally). The Polhemus Tracker works at 120 Hz tracking the position of users hand and its orientation. Hand orientation does not affect the task. Any movement perpendicular to the surface of the screen is not relevant because the motion of the virtual hand is two dimensional.

D. Protocol

Each user has been requested to perform the task to test the added delays of 0, 10, 20, 30, 40, 50, 60, 100 and 200 ms. First, users get familiarized with the task and the device with 0 additional delay. After a short rest period of time, we recorded their performance at each added delay, presented in random sequence to avoid, as much as possible, a bias due to the training achieved during the trials. The unavoidable delay in this application has been estimated to be 28 ms, a value that does not affect user performance [13], [17]. For the experiment, the delay was added on the communication between the sensors and the dynamics simulations environment. This is not, of course, the only place where delay could be produced in a virtual reality application. The typical sources of delay are [3]:

- in the tracker signal
- in communication between the tracker and the computer system
- due to computations required to process the tracker data
- due to graphical rendering

Anyway, for the present application that does not feature a big overhead due to computation of environment dynamics, it is reasonable to model the delay in this way. Furthermore the delay between the sensors and the simulated environment proved to be more critical for the performance in[22].

Each trial was 120 seconds long. The rest between each trial was 1 minute. 10 series of 9 trials have been performed, for a total of 90 trials and 180 minutes of performance recorded.

E. Data analysis

We measured the overall performance at the catching task and the underlying organization of the movements. Performance is defined as the percentage of balls that were successfully caught. For each performer, the series has been normalized dividing it by the value of the performance with no added delay. The traker gives only the positon of user hand, hence speed, acceleration and jerk have been computed through numerical differentiation. A linear low pass filtering has been added to avoid a magnification of noises through derivation.

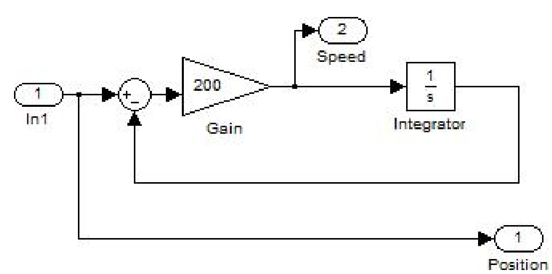

Fig. 2. Filtering and first order derivation of position signal implemented in Matlab/Simulink.

Movement organization is assessed in successful catches only. A first global variable is movement time (MT), defined as the duration of the hand movement for a successful catch.

This variable is useful to assess the cognitive demand for the control of the movement [6] and varies as a function of the delay in pointing movements [13]. We also assessed jerk cost (see appendix B).and the number of corrective sub-movements during a successful catch. These variables measure movement smoothness, which decreases when the delay increases [21], [14], [19], [1]. Finally, we also wanted to assess the variability of the behavior, as an enhanced variability can reveal zones of qualitative changes in the control of visually guided movements [20]. We computed the circular variance of the catch angle (see appendix C). Small variances are associated with series of catches performed with similar angles, larger variances are associated with a less regular behavior in catching. The circular variance has been chosen as indicator of dispersion of catch angles, instead of linear variance in order to get the desirable advantage of having a measure invariant with respect to the reference system and changing without discontinuities with continuous variations of input samples.

F. Statistical Analysis

The effect of the delay on the performance and on the movement organization of movement was assessed through analyses of variance with repeated measures (ANOVA). We ran an ANOVA for each dependent variable (i.e., performance, movement time, normalized jerk and number of sub-movements per catch, catching angle variability). The effect of the delay was deemed significant for probability $p < 0.05$.

III. RESULTS

We first address the effect of the delay on the performance at catching the balls, that is the dynamic target tracking task; second, we address the effects of the delay on movement organization, the behavior.

A. Performance

Figure 3 shows the mean performance as function of each added delay. A one way analysis of variance with repeated measures (ANOVA) indicates that movement performance is significantly impaired for delays greater than 60 ms ($F_{(8,72)} = 23.65, p = 0.01$). Yet, the shape of the curve has an inflection point for about 50 ms of added delay. This change in slope of the performance-delay relation is taken as an indication that users were able to cope with up to 70-80 ms of total delay (that is, when adding 50 ms to the unavoidable 28 ms delay of the system).

B. Movement organization

Having found that the delay has a significant impact on user performance, we now want to check whether it could have an impact on movement organization as well. Movement time as a function of the added delay is shown in figure 5 and the variance of catch angle as a function of the added delay is shown in figure 4.

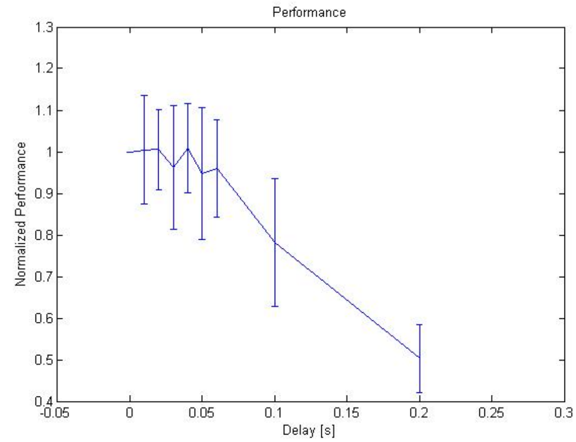

Fig. 3. Mean user performance as a function of delay. The performance is normalized to the one without added delay. Error bars represent inter-subject standard deviation.

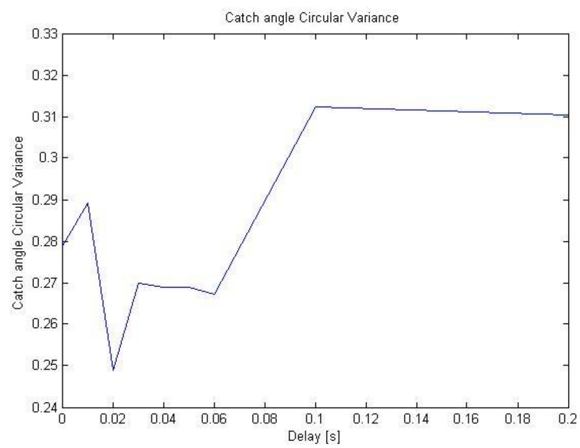

Fig. 4. Circular variance of catch angle variation with added delay. For each added delay condition the circular variance has been computed over all the catches perfromed by all the users.

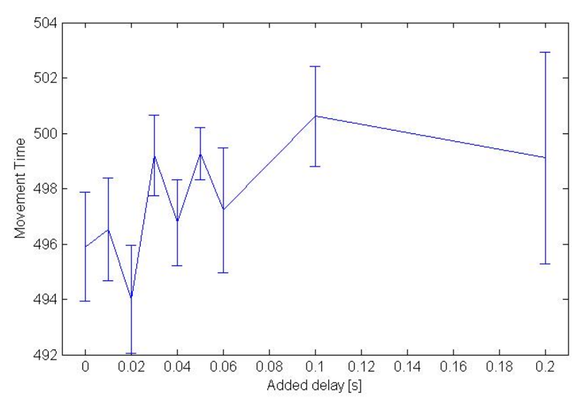

Fig. 5. Average time spent in hand movement between two catches, error bars represent estimated variance among users

IV. Discussion

A one way analysis of variance (ANOVA) indicates that movement time is not significantly influenced by the added delay (with probability $p > 0.05$). This result was surprising, because we expected that increasing the delay would lead to an increase in the coefficients of the Fitts's law, as reported for pointing tasks [13]. We ran a more detailed examination of the coefficients of the Fitts's law[6], to distinguish the *informational* and *non informational* aspects of the behavior [23]. The former relate to the slope of the linear regression of movement time on difficulty, and the latter to the intercept. These two aspects of the behavioral answer to the environmental constraints might differ as a function of the added delay: Small delays induce a change in the intercept only, larger delays elicit changes in both slope and intercept (see e.g., figure 2 in [13]). We found that the added delay did not change the informational aspects of the catching task: The slope of the Fitts's law was not significantly influenced by the added delay (with probability $p > 0.05$), with a throughput of 31 $\pm$ 11 bits/s on average. Similarly, the non informational aspects of the Fitts's law (i.e., the intercept that represents a movement time with an informational constraint of zero bits) were not significantly influenced by the added delay (with probability $p > 0.05$), with an intercept value of 383 $\pm$ 89 ms on average. As the added delay could result in less smooth movements [1] due to more corrective sub-movement units [14], we then computed the normalized jerk per catch (see appendix B). and assessed the number of sub-movements from the number of peaks in the velocity profile [14]. The ANOVA indicated that the added delay did not affect the normalized jerk cost (with probability $p > 0.05$) and neither the number of sub-movements (with probability $p > 0.05$). Finally, as far as the variability is concerned, a one way ANOVA indicates that the catching angle variability is not significantly influenced by the delay (with probability $p > 0.05$). Yet, we notice that the variability of the catching angle looks higher for delays over 60 ms, which is a pattern of change similar to that of the performance. All in all, our analyses indicate that the organization of the catching movement towards the ball is similar whatever the added delay in the interaction loop. We think that this lack of significant effect on the behavioral variables is due to the specific constraints imposed by the task. Because a new ball falls every 500 ms, one has to make catch movements with a similar tempo, hence a constant movement time (on average 498 ms, see figure 5). Indeed, an increase of the necessary time to catch a ball over 500 ms would obviously result in missing one of the next falling balls. Similarly, a higher error rate would probably hide any further increase in the variability of the catching angle over 60 ms.

V. CONCLUSIONS AND FUTURE WORKS

Our results confirms that user performance in a demanding perception-action task such as catching falling balls is not affected by delays under 70-80 ms in the human machine interaction loop. This result is important because the value of 70-80 ms is also the lower delay that is consciously perceived [16]. The fact that conscious perception of the delay is concomitant with the drop in performance suggests that the latter might cause the former. If confirmed experimentally, this would give support to the idea that human perception is rooted in action capabilities, as proposed by the psychologist J. J. Gibson [8].

Also important is the lack of significant effect the added delay on the behavioral variables, a conclusion that contrasts with previous reports about the effect of delay on bouncing balls in VR [16] or tracking targets [19]. This lack of visible effect on movement organization seems the consequence of the tempo imposed by the very nature of the task: The imposed rhythm of one catch each 500 hides probable changes in movement organization. This calls for more sophisticated movement analysis tools, but also highlight how task-dependent is the organization of human behavior. The result should hence be compared, in a future work, with the one obtained with a task avoiding user constant behavior. This can consist in the same task with balls falling at random times.

Finally, our results are important as an assessment of the platform itself [12]. Now that the impact of delay on a ball catching task has been quantified, further analysis should be applied to more complex tasks, such as the whole three-ball cascade juggling task. This would allow to test the effects of the delay on other components of user performance such as timing, balance, hand trajectory regularity and coordination. In particular the test would be performed in the range of delays not creating effects on the tracking task itself, as suggested by [16].

Future works would also take in account other the effect of other parameters besides the delay itself. The number of the balls falling simultaneously and their speed should be varied to obtain a more general test.

Appendix

A. LIGHT WEIGHT JUGGLING VIRTUAL ENVIRONMENT

The *Light Weight Juggling* (LWJ) application allows the user to perform juggling (three ball cascade) in virtual reality [12]. The specific outcome of training with the juggling demonstrator is to provide a non-juggler with the skills required to successfully juggle the three ball cascade and similar level tricks.

The LWJ application is implemented with the Mathworks® simulation tool (Simulink®) and XVR[4]. It has a modular structure: the virtual reality rendering and the user interface is implemented through the XVR code, the environment physics and the state machine representing the task are implemented as a Simulink simulation model. The two components communicate through an UDP network connection, exchanging information about the physical entities and the visual elements to be displayed. The user interacts with the simulated environment through the *Pohlemus* three dimensional tracker. It is used to get the hand position in real time. It works getting the position and

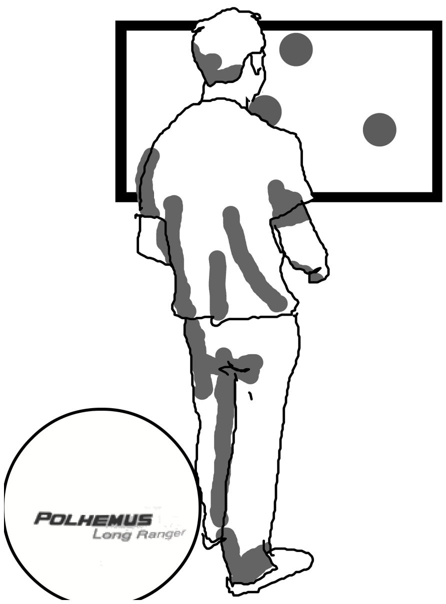

Fig. 6. Light weight juggling: user standing between the screen and the electromagnetic tracker antenna

the orientation (not used in this application) of two sensors linked to the user hands. The virtual environment consist of a three dimensional representation of Juggling portraying two hands controlled in position by the user and the three balls to be juggled. Anyway, the simulated dynamics is two dimensional. This is because the system is thought to work without any stereoscopic viewing system and then it would be difficult to control the depth of hand position. There is not any kind of haptic feedback. The tosses are triggered by hand acceleration and the catches are triggered by hand position.

B. NORMALIZED JERK COST

Movement fluency and ease of performance can be assessed by measuring movement smoothness [9]. Easy to perform reaching movements are organized with a straight and smooth path in space, and with a single and velocity peak. Such features of movement organization are well summarized with a measure of the jerk cost for the movement of interest.

The jerk J is the third time derivative of the position time series, with units of cm/s^3 here.

The jerk cost JC over a duration T is :

$$JC = \frac{1}{2} \int_0^T J_t^2 dt \qquad (1)$$

In the present paper, the method to compute the normalized jerk cost was as follows:

- compute the jerk time series J_t: this is the third time derivative of the recorded position time series. Differentiation was done using the first central difference method, combined with low-pass filtering (dual-pass second-order Butterworth filter with a cutoff at 8 Hz).
- compute the jerk cost JC for the duration of the movement lasting from t_{beg} to t_{end}, using equation 1 (with $T = t_{end} - t_{beg}$)
- make the jerk cost a dimensionless measure, using equation 2

To compare movements with different amplitude and duration, the jerk cost is expressed in normalized units of time and distance [9]:

$$NJC = JC * \frac{T^5}{L^2} \qquad (2)$$

where L is the length of the travelled path and T the duration of the travel.

C. CIRCULAR VARIANCE: DEFINITION AND APPLICATION

Statistical moments such as mean and variance, computed over angular values become dependent on the direction set as zero. This could have not sense in the general case. To extend properly these concept to angular values the following step are applied: The directions are expressed as vectors, for example the angle ϕ_i becomes the vector

$$\vec{D_{\phi_i}} = \begin{bmatrix} \cos(\phi_i) \\ \sin(\phi_i) \end{bmatrix} \qquad (3)$$

The mean is then computed as the mean of the vectors $\vec{D_{\phi_i}}$

$$\vec{M} = \frac{1}{n} \sum_{i=0}^{n} \vec{D_{\phi_i}} \qquad (4)$$

The variance is computed as

$$1 - \left\| \vec{M} \right\| \qquad (5)$$

Variance V is always $0 < V < 1$.

For an extensive exposition of this topic see[10].

REFERENCES

[1] A Beuter, D Larocque, and L Glass. Complex oscillations in a human motor system. *J Mot Behav*, 21(3):277–289, Sep 1989.

[2] U. Bockholt, S. Webel, T. Engelke, H. Graf, and J. Behr. Virtual and augmented reality training platforms for transfer of skills. 2009.

[3] Steve Bryson. Effects of lag and frame rate on various tracking tasks. In *Proceedings of the SPIE - The International Society for Optical Engineering, Vol. 1915 of Stereoscopic Displays and Applications IV*, pages 155–166, 1993.

[4] Marcello Carrozzino, Franco Tecchia, Sandro Bacinelli, Carlo Cappelletti, and Massimo Bergamasco. Lowering the development time of multimodal interactive application: the real-life experience of the xvr project. In *ACE '05: Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology*, pages 270–273, New York, NY, USA, 2005. ACM.

[5] J.Y.C. Chen, E.C. Haas, and M.J. Barnes. Human performance issues and user interface design for teleoperated robots. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 37(6):1231 –1245, nov. 2007.

[6] Paul M Fitts. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47(6):381–391, 1954.

[7] A J Foulkes and R C Miall. Adaptation to visual feedback delays in a human manual tracking task. *Exp Brain Res*, 131(1):101–110, Mar 2000.

[8] James J. Gibson. *The Ecological Approach To Visual Perception*. Houghton Mifflin, Boston, 1979.

[9] Neville Hogan and Dagmar Sternad. Sensitivity of smoothness measures to movement duration, amplitude, and arrests. *J Mot Behav*, 41(6):529–534, Nov 2009.
[10] Peter E. Jupp Kanti V. Mardia. *Directional Statistics*. Academic Press Inc. (London), 1972.
[11] Stas Krupenia. Example of m24 juggling demonstrator analysis. Report for the project SKILLS, 2008.
[12] Julien Lagarde, Gregory Zelic, Carlo Alberto Avizzano, Vittorio Lippi, Emanuele Ruffaldi, Hagar Zalmanov, Vered Erev-Yehene, Nick Gillian, Sile OModrain, Denis Mottet, and Daniel Gopher. The light weight juggling training system for juggling. In *International Conference on Multimodal Interfaces for Skills Transfer*, 2009.
[13] I. Scott MacKenzie and Colin Ware. Lag as a determinant of human performance in interactive systems. In *INTERCHI '93: Proceedings of the INTERCHI '93 conference on Human factors in computing systems*, pages 488–493, Amsterdam, The Netherlands, The Netherlands, 1993. IOS Press.
[14] R C Miall, D J Weir, and J F Stein. Visuomotor tracking with delayed visual feedback. *Neuroscience*, 16(3):511–520, Nov 1985.
[15] R C Miall, D J Weir, D M Wolpert, and J F Stein. Is the cerebellum a smith predictor? *J Mot Behav*, 25(3):203–216, Sep 1993.
[16] Antoine H P Morice, Isabelle A Siegler, and Benoît G Bardy. Action-perception patterns in virtual ball bouncing: combating system latency and tracking functional validity. *J Neurosci Methods*, 169(1):255–66, Mar 2008.
[17] Antoine H P Morice, Isabelle A Siegler, Benoit G Bardy, and William H Warren. Learning new perception-action solutions in virtual ball bouncing. *Exp Brain Res*, 181(2):249–265, Aug 2007.
[18] WM Smith, JW McCrary, and KU Smith. Delayed visual feedback and behavior. *Science*, 132(3433):1013–1014, Oct 1960.
[19] Nigel Stepp. Anticipation in feedback-delayed manual tracking of a chaotic oscillator. *Exp Brain Res*, 198(4):521–525, Oct 2009.
[20] P. Tass, J. Kurths, M. G. Rosenblum, G. Guasti, and H. Hefter. Delay-induced transitions in visually guided movements. *Phys. Rev. E*, 54(3):R2224–R2227, Sep 1996.
[21] J L Vercher and G M Gauthier. Oculo-manual coordination control: ocular and manual tracking of visual targets with delayed visual feedback of the hand motion. *Exp Brain Res*, 90(3):599–609, 1992.
[22] Martin Valere Ya-weng Tseng, John P. Scholz. Effects of movement frequency and joint kinetics on the joint coordination underlying bimanual circle drawing. *Journal of Motor Behavior*, 38:383 – 404, 2006.
[23] Shumin Zhai. Characterizing computer input with fitts' law parameters–the information and non-information aspects of pointing. *International Journal of Human-Computer Studies*, 61(6):791 – 809, 2004.