AUDIO FEEDBACK IN PREDICTIVE HCI METHODS

Amalia de Götzen
University of Verona
Dept. of Computer Science

Davide Rocchesso
University of Verona
Dept. of Computer Science

ABSTRACT

This paper is an attempt to study a well known HCI predictive model in an audio perspective: a derived law from the Fitts' law model will be analyzed providing an audio interactive display in which the user has to perform a simple tuning task by hitting a button. The idea is to simplify as much as possible the interaction in order to find out its invariants when the feedback is just the audio one. An experiment is carried out in order to evaluate this conjecture.

1. INTRODUCTION

The human-computer interaction and the predictive models which are used to evaluate interfaces are mostly based on visual feedback: very often the user is supposed to have just one sense which is not helped nor substituted by any other one. New interfaces need to be designed in order to realize a multimodal interaction between the user and her machine: this necessity can be tested and explored by studying well known predictive laws in other sensory domains. Sensory substitution [2] as well as sensory illusion and multimodality can be useful in designing and experiencing a new interface. The main idea is to develop an interface where the information can be conveyed through different modalities. The effectiveness of the multimodal interaction will be evaluated using predictive models which have been demonstrated to be useful in this kind of context too. Gesture interaction with audio feedback, and in general non verbal interaction, can be considered the natural application of our study.

Gesture and sound seem naturally connected in a clear and obvious way: the image of instrument players learning to use their body in order to produce sound is indeed widespread and compelling enough. While each instrument needs specific gestures to be played in a correct and pleasant way, invariant laws regulating gestures across all instruments may be found. A computer can be considered as a musical instrument, perhaps not new but still far from having a coded tradition related to musical gesture. Musical gesture can be simply thought as a gesture that produces sounds in a continuous feedback loop: this is a general definition that can be used in many interactive contexts besides the musical one.

This kind of studies involves several research domains which are now getting to communicate and work together: human performance, auditory perception and signal processing are all involved in this investigation area. This work is just the first step in this direction: before introducing the gesture component in our interaction, it is worthwhile to study the audio feedback by itself, in comparison with HCI predictive models which are not usually applied to the audio domain.

2. FITTS' LAW

The origins of the Fitts' performance model, so useful in human-computer interaction, must be kept in mind when considering the Fitts' law. The law takes its name from its author whose innovative idea, in 1954[5], was to apply information theory to human-motor systems.

Even if this law is 50 years old it still plays a central role in an active community: HCI researchers started to apply this law to their needs in the seventies [15] and Fitt's law is currently used in the ISO 9241-9 standard on the evaluation of pointing devices. The debate about the mathematical formulation of the law and its interpretation are still a hot research topic: a whole issue of the International Journal of Human Computer Studies has been recently devoted to it [1].

The model is based on time and distance. It enables the prediction of human movement and human motion based on rapid, aimed movement (i.e. not drawing or writing). An intuitive idea is that movement time is affected by distance and by the precision required by the size of the target towards which one is moving.

2.1. Fitts' law and the theory of information

Fitts discovered that movement time was a logarithmic function of distance when target size was held constant, and that movement time was also a logarithmic function of target size when distance was held constant. Mathematically, Fitts' law is stated as follows:

$$MT = a + blog_2(2A/W) \qquad (1)$$

where:
MT = movement time
a, b = regression coefficients
A = distance of movement from start to target center
W = width of the target

There is an evident analogy with Shannon's Theorem 17[14]: the accomplishment of movement in Fitts' model is analogous to the transmission of "information" in electronic systems; movements have an index of difficulty which can be expressed in bits and the motor system transmit "bits of information" while moving. To explicitly show the formal derivation above mentioned, we pick up Shannon's formula:

$$C = Blog_2(S/N + 1) \tag{2}$$

where:
C = information capacity[bits/s] of the channel
B = bandwidth[Hz] of the channel, the counterpart of 1/MT
S = signal
N = noise

Here lies the innovative aspect of Fitts' law: a quantitative way to measure the difficulty of a motor task becomes available through it and a "new" way to transmit information is implicitly described through the definition of a human channel. Fitts defined also some other indexes that clarify the analogy with the Shannon formulation:

$$ID = log_2(2A/W) \tag{3}$$

$$IP = ID/MT \tag{4}$$

Where ID is the index of difficulty while IP is the index of performance, analogous to channel capacity C. [8] features a detailed analysis of all the variations on Fitts' law, (i.e., Welford, MacKenzie formulation) derived from the need to correct the approximation by Fitts of the Shannon theorem, avoiding:

1. the theoretical problem of a negative rating for task difficulty,
2. the problem of the disproportion of the relative contribution of A and W in the prediction equation: similar changes in target amplitude and target width don't affect similar but inverse changes in movement time, as suggested in the original Fitts formulation

Skipping the detailed analysis of the data performed by MacKenzie, the final result by direct analogy to the Shannon formula is shown in 5. This is also the most frequently used one because of its better fit with empirical data.

$$MT = a + blog_2(A/W + 1) \tag{5}$$

2.2. Fitts' law and dynamic systems

The information theory is a way to describe the derivation of Fitts' law, but a few others are possible[11, 10]. In particular the first order lag can be used as a model for human movement [4]: it can provide a basis for predicting both movement time and also the time history of movement. The response of the first-order lag to a step input of amplitude A with a gain factor k, can be shown to be:

$$Output = A - Ae^{-kt} \tag{6}$$

The solution of the previous equation is the time that it would take the first-order lag to cross into the target region. As shown in [11] the result is:

$$\frac{ln2}{k} log_2(\frac{2A}{W}) = t \tag{7}$$

The similarities with Fitts' law are quite evident: this is a logarithmic relation in which the movement time is directly proportional to the amplitude and inversely proportional to the target width. Though the similarity is striking, the second order system seems to provide a better fit for the velocity profile for human movements (a gradual increase in velocity) and for Fitts' law as well [7]. Both first and second order lags can be considered as continuous control methods: the error is continuously monitored and the adjustment is provided according to the detected error. Another possible model to explain the derivation of Fitts' law is a system which can provide discrete control: such iterative correction model is proposed by Crossman [4]. This model is also shown to be consistent with Fitts' law. Fitts' law is a very open field: considerable work has been carried out in HCI on the Fitts' law model [8, 3, 16] but the literature on Fitts' law with sound feedback appears to be very scarce [17, 6, 9, 6]. This investigation is also suggested by the big role played by multi-modality and multi-sensory communication in the design of next generation interfaces: non-speech communication will play an important role inside the information stream established between machines and users.

3. SCHMIDT LAW: THE LINEAR SPEED/ACCURACY TRADE OFF

This paper investigates the Schmidt law [13] which is often studied in comparison with the Fitts' law: the main difference between the two approaches is that Fitts treated movement amplitude and target width as independent variables and movement time (MT) as a dependent variable, while Schmidt chose to manipulate amplitude and movement time and to measure the movement variability of the effective target.

In the Schmidt case, the tradeoff, formally the impulse variability model, forecasts that the standard deviation in end-point coordinates (viz., accuracy), is a linear function of velocity, calculated as distance over time:

$$W_e = a + bA/MT \tag{8}$$

It is interesting to note that the previous equation and Fitts' law contain the same three parameters (except that We is the standard deviation of end-point coordinates and is $4.133 \times SD$ in Fitts' adjusted model). Although this equation can be rearranged with MT as the predicted variable, it is still fundamentally different from the Fitts' law since the relationship is linear rather than logarithmic, and because the information-theoretic analogy is absent. This linear trade off can be explained using two hypothesis:

- the movement-brevity hypothesis

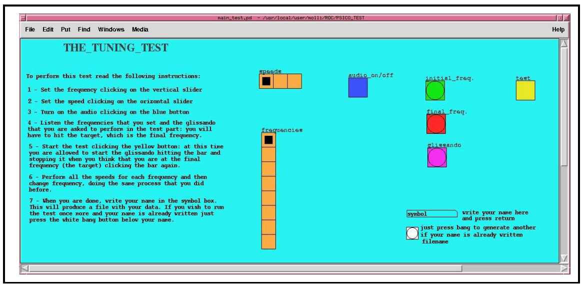

Figure 1. Software environment

- the temporal precision hypothesis

It seemed interesting for our research to investigate this law with audio feedback since temporal constraints are very often present when we are dealing with music and performance tasks.

The aim of this paper was to explore the speed/accuracy trade off when there are temporal constraints in order to discover if the Schmidt law is still reliable with audio feedback. This can be a first step in the investigation of Fitts' law in an audio perspective: it seems easier and more natural to start without involving multimodal feedback by just trying to figure out if it is possible to apply HCI models designed for visual feedback to audio feedback too. In this experiment the variables have been reduced as much as possible: the only thing that we want to observe is the response of our hear.

4. EXPERIMENT

4.1. Participants and stimuli

Ten subjects (between 26 and 48 years old) participated in the experiment performing 24 trials. All participants reported normal hearing and sight, and normal motor capabilities in their hands. All of them were naive as to the purpose and hypotheses of the test and all of them volunteered. The test has been done using Pure Data software and the user interface is shown in fig. 1.

4.2. Procedure and design

Auditory stimuli were obtained using simple pure tones: the user has to reach a target frequency from a starting one. The 8 different tasks which correspond to 8 different intervals in 4 different frequency ranges have to be performed at different speeds (time constraints). The intervals have been chosen in order to be non-musical intervals and to explore the whole audible spectrum. For each interval the user will start training his hear: she/he will hear the starting frequency, the final frequency (the target) and then the glissando performed with a given speed. After that he will start to perform the test. The user does not have any kind of visual feedback and she/he has to just press a button in order to start and stop the frequency which she/he is controlling while is continuously hearing the target frequency which is the frequency that she/he has to reach. Table 1

Initial freq. (Hz)	Final freq (Hz)
340	1310
340	2430
340	3270
1240	2510
1240	3700
2170	3580
2170	4000
3300	4020

Table 1. Initial and target frequency of each trial

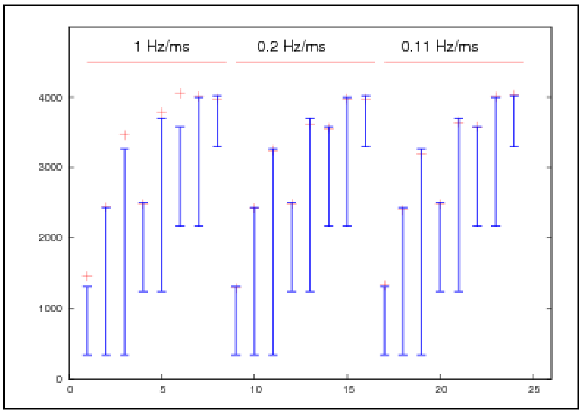

Figure 2. Data collected

shows the list of intervals that have been used for the three different speeds.

4.3. Data analysis

In fig. 2 all the data collected from one subject are plotted: the blue line is the distance (A) in Hz between the initial frequency and the target one, while the green cross is the reached target. The users performed 8 trials with 3 different velocities (1Hz/ms, 0.2 Hz/ms 0.111 Hz/ms). A one factor ANOVA test showed the significance of the mean value of the collected data: the computed p factor was definitely below the significance threshold (p=0.05).

Fig. 3 shows that the mean deviation between the reached frequency and the target frequency is quite evident (here there is just the behaviour of one subject). Also, even though each participant has her own characteristic behavior some common patterns may be noticed: the task seems to be easier when the velocity is higher. The different curves for each participant show that the mean deviation is bigger for lower velocities, where participants often stop the glissando before reaching the target frequency.

This observation is quite important since the Schmidt law is supposed to work in the same way (for quick and short movements): assuming that we can compare short movements to fast played intervals, audio feedback may thus follow the law.

Fig. 4 shows the mean values for each trial. The width of each bar represents the distance between the starting

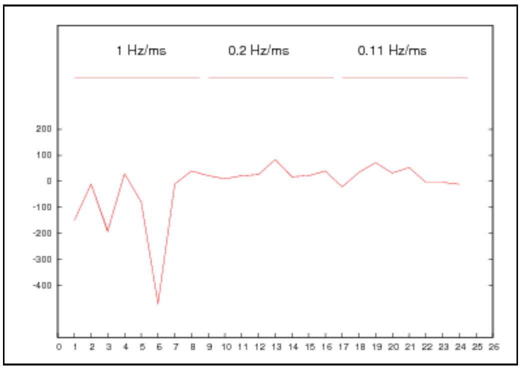

Figure 3. Mean deviation between the reached frequency and the target frequency

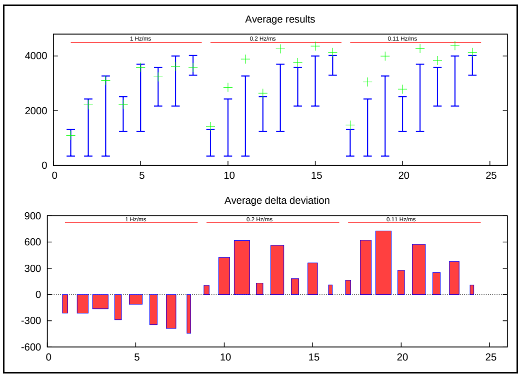

Figure 4. Reached frequency Mean values

frequency and the ending frequency.

This data plot is quite interesting, because it shows several things at the same time:

- the performances may be subdivided in two different behaviors: participants tend to stop the glissando before reaching the target when the performance is at high speed (temporal constraint), while they tend to stop it *after* when the performance is at slow speeds;
- the standard deviation is lower for higher speed and the results have better uniformity for these 8 trials;
- the standard deviation is increasing with speed and the results are quite different depending on the frequency interval;

What appears quite clearly observing and analyzing the data is that the relation between W_e and A/MT is not verified for every time constraint. Schmidt law states that the relationship between We and A/MT is nearly linear and for very different amplitudes and movement times, but with the same ratio A/MT, the We is reported to be about the same. Looking at Fig.5 the change of We for each A/MT value may be observed: each point is specified by the amplitude A (the distance between the initial frequency and the target frequency) and the standard deviation (We).

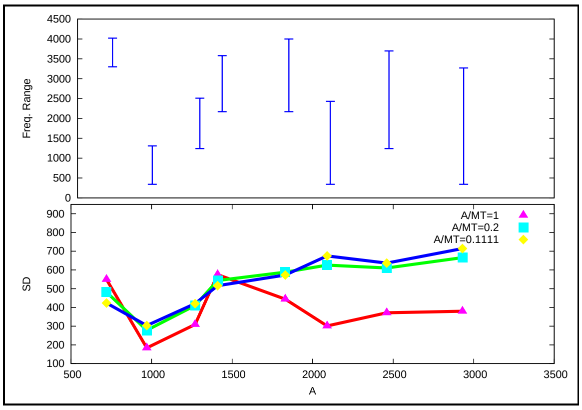

Figure 5. Effective width for the three A/MT values

The comparison between the three curves is interesting: the effective target in the audio domain seems to be strictly related to the distance A and to the frequency range in which our "gesture" is located: for each speed we have better results with smaller distances. Incidentally, the high SD value for the smallest distance sticks out: here, the frequency range is an important factor, since this distance is performed between 3300 Hz and 4020Hz, and it is the only one of this frequency range.

4.4. Results

The linear speed/accuracy trade-off was observed in very rapid actions where there was probably not enough time to detect errors and issue a correction, so we expected better data with higher speeds. This observation was also made by the users who found it easier to reach the target when the glissando speed was higher. The linear trade-off was originally observed using controlled MTs (longer than minimum MTs), whereas in the Fitts' paradigm the goal MT was to be as fast as possible while maintaining a high accuracy. Both these differences appear to influence the behavior. Regarding this question, Schmidt himself [12] asserts that:

- a linear trade-off seems to occur for movement tasks that are pre-programmed - under open loop control
- a logarithmic trade-off seems to occur for movement tasks that are governed by feedback-based corrections - under closed loop control

So, while the question of whether the relation is linear or logarithmic may seem rather abstruse, it has a functional significance in regard to whether a movement is carried out in an open loop or is subject to feedback-based corrections. These observations fit quite well with our experiment results: with audio feedback, the Schmidt law is more effective with higher speed while the frequency range factor has to be still deeply investigated.

5. CONCLUSIONS AND FUTURE WORK

The findings of this experiment clearly suggest the need of deep investigation on predictive laws in HCI with audio feedback: many variables and components of audio perception are at work in this kind of experiments, and it is necessary to analyze them and their effects on the user performance.

Therefore, future work will involve:

1. finding out whether the predictive model will work in different ways according to the sensibility of our hear to different frequency range and different frequency ratios (e.g. intervals);

2. studying the model in real-world musical applications (e.g. in musical playing);

3. finding out whether expressivity modifies the predictive model sensibly and whether expressiveness can be conveyed by these open loop models

4. finding whether the beating effect which occurs during the tuning test can modify the model, particularly in slow movements. This aspect is much more important in auditory perception than in the visual one: only audio feedback sports a proximity indicator (beating).

6. REFERENCES

[1] Fitts'law 50 years later: application and contrbutions from human-computer interaction. *International Journal of Human-Computer Interaction Studies*, 61:747–904, 2004.

[2] L. Etienne A. Lécuyer, J. Burkhardt. Feeling bumps and holes without a haptic interface: the perception of pseudo-haptic textures. In *CHI2004*, Vienna, Austria, April 2004.

[3] J. Accot and S. Zhai. Beyond fitts' law: Models for trajectory-based hci tasks. In *Proc. of CHI'97*, Atlanta, Georgia, USA, 1997.

[4] P.J. Goodeve E. Crossman. Feedback control of hand-movement and fitts' law. *Quarterly Journal of Experimental Psychology*, 1983.

[5] Paul M. Fitts. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology*, Vol. 47(No.6), June 1954.

[6] N. Friedlander, K. Schulter, and M. Mantei. Bullseye! when fitts' law doesn't fit. In *CHI1998*, Los Angeles CA USA, 1998.

[7] J.A. Foulke G.D. Langolf, D.B. Chaffin. An investigation of fitts'law using a wide range of movement amplitudes. *Journal of Motor Behavior*, 8:113–128, 1976.

[8] I. MacKenzie. *Fitts' law as a performance model in human-computer interaction*. Doctoral dissertation, University of Toronto, Toronto, Ontario, Canada., 1991.

[9] A. Pirhonen, S. Brewster, and C. Holguin. Gestural and audio metaphors as a means of control for mobile devices. In *CHI2002*, Minneapolis Minnesota USA, 2002. http://www.dcs.gla.ac.uk/ stephen/papers/-CHI2002.pdf.

[10] A.M. Alimi R. Plamondon. Speed/accuracy tradeoffs in target-directed movements. *Behavioral and Brain Sciences*, 20:279–349, 1997.

[11] J.M. Flach R.J. Jagacinski. *Control Theory for Humans - Quantitative Approaches to Modeling Performance*. Lawrence Erlbaum associates, London, 2003.

[12] Robert Schmidt and T. Lee. Motor control and learning. a behavioral emphasis. *Human Kinetics*, 1999.

[13] Robert Schmidt, A. Zelaznik, B. Hawkins, J. Frank, and J. Quinn. Motor output variability: A theory for the accuracy of rapid motor acts. *Psychological Review*, 86:415–451, 1979.

[14] C.E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, pages pp. 379–423, 623–656, July-October 1948.

[15] B.J. Burr S.K. Card, W.K. English. Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a crt. *Ergonomics*, 21:601–613, 1978.

[16] E. Sllnas and S. Zhai. Collaboration meets fitts' law: Passing virtual objects with and without haptic force feedback. In *Proc. of INTERACT 2003*, Zurich Switzerland., 2003.

[17] R. Vertegaal. *An Evaluation of input devices for timbre space navigation*. Doctoral thesis, University of Bradford, 1994. http://citeseer.ist.psu.edu/vertegaal94-evaluation.htm.