On the Relevance of Auditory Feedback for Subjective and Objective Quality of Control in a Balancing Task

Matthias Rath

Deutsche Telekom Laboratories
Technical University Berlin, Germany
matthias.rath@telekom.de

Abstract

A tangible audio–visual interface based on the metaphor of balancing a ball on a tiltable track allows the measurement of human control movements under different conditions of sensory feedback. This specific scenario of human–system interaction forms an example for the definition of various measures of performance and quality of interaction. The dependence of these measures on specific configurations of the interface with regard to the employed audio–visual feedback, and their relationship is discussed.

1. Introduction

The notion of introducing a more intuitive or natural quality to human–machine interaction has recently received strong interest. The concept and field of *Embodied Interaction* [1][2] is maybe one of the most prominent in this respect but the same scopes appear to be driving forces behind developments of multimodal interfaces in general. The terms of "intutiveness" and "naturalness" however are far from being defined in a general, fixed manner, but are conceived and understood rather ad hoc in very specific contexts of interfaces and tasks. Other, related, attributes of human–machine interaction are measures of performance, such as times needed to perform a certain task or, more basic, a user's ability to fulfill a task in question. User experiments conducted with the "Ballancer" [3], a tangible audio–visual interface based on the metaphor of balancing a ball on a tiltable track, demonstrate several possibilities of defining and measuring aspects of perceptual quality in the user–device interaction. These quality measures are based on objective measurements of control gestures and performance times on one hand and on rather conventional techniques of "self-observation" in the form of questionaires on the other. They depend in different ways on the sensory conditions of the interaction, specifically the presence or absense of sonic feedback, and correlate among each other in various ways.

The work presented here has its background in the development of sound generation algorithms that aim at an enhanced use of the auditory perceptual channel in human–machine interfaces [4][5]. These interactive *sound models* may form a basis for a use of sound in human–machine interaction adequate to the significance of continuous sonic information flows in everyday human environments [6]. While most related existing studies and applications focus on short, discrete acoustic events, mostly of notification character, one particular interest behind the work of sound generation and design and the experiments discussed in the following is the relevance of continuous sonic feedback for human–system interactions. The experiments described in this contribution started off as an evaluation of one of the mentioned sound models, one of rolling interaction of solid objects [7], and its capability to spontaneously convey ecological information such as velocity or physical attributes of the involved objects such as size or mass to a user. Results however have a wider psychoacoustic significance as they shed light on processes of auditory perception and the direct exploitation of sonic information in human control movements.

2. Interaction metaphor and control interface

In his classic experiments on human motor control [8] Fitts uses probably one of the most simple movement tasks to think of, that of moving a stylus from a given starting position into a target area. In this scenario the weight and size of the stylus are negligible with respect to the dimensions of the human body or arm. The stylus can thus be seen as a marker connected to the human body and in this sense Fitts' law applies to human movement as such rather than human interaction with the external world. Human movement as such however usually generates low acoustic feedback and is in its control dominated by tactile proprioception [1] under support through visual confrontation of body position with respect to external targets or references. A related but extended metaphor of control is used here to include true interaction of user and an external system and achieve a situation of stronger significance of auditory feedback. These objectives are reached by means of a scenario where the position of the controlling arm is not directly related to that of the controlled object but rather to a parameter of acceleration: in balancing a rolling ball on a track gravity acceleration along the track is approximately proportional to the sine of the angle of inclination which is in turn proportional to the vertical distance of the two hands holding the track.

2.1. The control interface

Physically, the *Ballancer* consists of a 1m-long wooden control-track that the user, or here: test subject, holds as if balancing a small marble rolling along its upper face (compare the photo of figure 1). This virtual ball movement is simulated by the Ballancer software according to the measured angle of the track and the simplified equations of the scenario. The virtual ball is displayed graphically in a schematic representation on the computer screen (compare figure 1) and acoustically through different sound models. Details of the technical realization and the fundamental notions and motivations behind the devel-

[1] I here use the really tautological (in combination with the term proprioception) adjective "tactile" to stress the fact that I'm not referring to any perception of the own body through other senses…

opment of the Ballancer are described in a dedicated overview article [3].

Figure 1: *The Ballancer with a 19" graphical display. The simultaneous use of the real glass marble on the track and the virtual ball on the computer screen only serves for demonstration purposes in the photo. All test settings contain only one of the two elements, "real" or virtual ball.*

3. Aspects of quality of interaction

Different studies have been conducted at the Ballancer interface in various sensory configurations that deal with different aspects of quality of control interaction. Relevant results in the present context are collected, the first originating from an initial series of studies that has been reported in previous publications [3][9] to which I refer for details, and that is thus only shortly summarized here. An overview of more recent work follows below.

3.1. First assessment of the Ballancer metaphor and interface

One of the sound models previously mentioned, the one of rolling [7], is capable of expressing auditorilly a wide range of ecological parameters such as the size, mass, shape and velocity of a rolling object. It is based on a simplified physical model of the inner resonance behavior as well as the force acting between the rolling object and the plane to roll on. The involved surface profiles that are "tracked" during the rolling movement form here the initial source of vibration, that is processed as part of the sound model following considerations on the physical and geometrical principles of the scenario. (I refer to [7] for details.) The model hereby follows the notion of *cartoonification* [10][3], i.e. it aims at simplification, expressivity and clearness in its sonic appearance rather than realism.

As a first step in assessing users' spontaneous understanding and adoption of the Ballancer interface and the involved sonic feedback in particular, the first evaluation study with 10 participants [3][9] involved a free identification test. Blindfolded test subjects were presented sound examples of the employed model of rolling and of a real ball rolling on the control stick in its mechanical configuration and asked to answer the question of "What is going on here?". Later they were given (still blindfolded) access to the device in the same two configurations ("real" and virtual ball and sound), followed by the same question. Answers as discussed in [3] reveal that the sonic behavior of the device with synthetic, cartoon-sound feedback from the virtual ball was recognized as "synthetic" or virtual but represents more clearly the intended metaphor than the actual "real" mechanical scenario. The presented synthetic sound was in fact described by 9 of the 10 subjects in terms of a rolling object and all 10 subjects described the intended scenario of an object rolling on a tiltable track when handling the Ballancer interface blindfolded in this configuration. The sound of the real ball in contrast turned out to be more ambiguous as it was attributed to *several* rolling objects by 4 subjects (out of 10) along with other discrepant interpretations ("something like a toy car"...). This shows that the attributes of "realism" and "transparency" in virtual scenarios and clearness or unambiguousness are not identical and may even not correlate.

The second (and longer) part of the same study [9] consisted of a performance experiment where the same 10 subjects were asked to perform the target reaching task as described above with and without acoustic feedback from the rolling model. Analysis of recorded movement trajectories and resulting "task times", needed to accomplish the task, reveal significant (in the sense of t-test comparison) improvements of the average task times and several other parameters of the movement with sonic feedback. As displayed in table 3.1, at a $19''$ graphical display subjects concluded the task in average ca. 9% faster when they had in addition auditory feedback from the rolling model. Summarizing shortly further analyses of recorded control movements, it can be stated that subjects in average accelerated and stopped the controlled ball more efficiently with the additional auditory feedback [9]. These results prove that the acoustic feedback from the rolling model conveys additional information about the system that users spontaneously and in general unconsciously exploit for optimizing their control behavior. More precisely, the performance improvement (in the sense of the noted effects on average task times and control movements) with sonic feedback must be attributed to information of velocity of the controlled ball contained in, and spontaneously perceived from, the sound, since only this and *no* positional information is reflected in the acoustical behavior. [2] In this sense the sonic feedback contributes to an objective quality of the Ballancer interface as a means to convey information or supply control over a system. Subjective quality aspects such as possible differences in "pleasantness" of the handling of the device while trying to accomplish the task were not yet systematically examined in the first study. Some users however made spontaneous remarks of the kind that the task is "more fun" with sound or "more boring" without.

Table 1: Average times needed to complete the "target reaching"-task at a $19''$ display, with and without sound. The additional columns contain the relative difference of the values δ, "without sound" to "with sound" in $\%$ and the statistical p-value for the two compared groups of measurements.

average task time (ms) with (+) and without (–) sound, percentual difference (δ) and statistical significance (p)			
+	–	δ (%)	p
5217	5675	8.8	0.031

[2] For a more detailed discussion of this notion the reader is referred to [9].

3.2. Effects of “abstraction” of sonic feedback

After the observation from the study reported above of the informative and clear character of the synthetic sound in parallel with it being perceived as “non-real” the notion of further “abstracting” the sonic feedback and possible consequences on different aspects of quality has been addressed again in recent works. To this end, a second, very simple and rather abstract sound model has been derived, by widely ignoring any idea of realism or even immediate similarity with the mechanical sound of a rolling object. This sonic feedback however still aims at expressing in a possibly intuitive way what is considered the main parameter of interest for auditory display in our context, the one of velocity of a controlled movement. To this end the processing that accounts for the physical interaction in rolling is strongly reduced and the model is stripped down to tracing, at audio rate [3] a chosen surface profile. This strategy may be compared to replacing the object *rolling* along the surface with an ideal needle as of a record player that follows what would be the “essential” trajectory of the movement (of the center of the virtual ball). Finally, an optimized, highly unrealistic surface profile is chosen, one of the shape of a lowpass-filtered sawtooth signal, to optimize the low-level psychoacoustic properties of the resulting signal: The Fourier spectrum of the sawtooth spreads over a very wide range of the frequency sensitivity of human hearing, up to its upper limit even for fundamental frequencies at the lower end of the hearing range. The sawtooth is perfectly periodic and thus stimulates a clear and strong sensation of pitch, e.g. in contrast to filtered noise. The signal is slightly lowpass filtered in order to minimally smoothen the otherwise extremely harsh esthetic appearance of the sawtooth. Summing up the consequences of the described derivation its frequency follows proportionally the sonified velocity.

In order to examine the effects of the more abstract, very simple sonic feedback on users’ handling and perception of the Ballancer device, as compared to the original sound model of rolling, a second experiment analog to the one described above was conducted, containing both conditions of sonic feedback (again along with a “no sound” condition). This performance experiment was accompanied by a questionaire addressing the subjective influence of the different sonic conditions. 6 subject participated in this pilot experiment, all of them students at the Technical University of Berlin, aged between 21 and 27, four male and two female. Despite this rather small number of participants (that is augmented in current experiments continuing the pilot) some statistically significant and interesting results were found, the most important of which are shortly reported below. Again, participants were asked to perform the target reaching task as fast as possible, this time under the three different conditions of *a)* sound feedback from the rolling model, “rolling sound (rs)”, *b)* feedback from the rather abstract model, “abstract sound (as)” and *c)* without sonic feedback, “no sound (no)”. The different conditions appeared in *sets* of 20 *games* (trials) each. The order of the sets/conditions was counterbalanced across subjects and the whole series of all conditions was repeated once for each subject so that the whole test consisted of six sets, e.g. for subject 1 of the form: “rs, as, ns, rs, as, no”, subject 2: “rs, no, as, rs, no, as”. . . Due to the repetition of the whole series, each condition appeared twice for each subject, as one set in a less “trained” state and again in “trained” circumstances in the second half of each test. Together with counterbalancing the order of conditions we can thus assume that any training effect during performance of the test should not introduce artefacts in the comparison of performance values at different conditions of feedback measured through all subjects.

After having concluded the performance test, subjects were asked to describe in free words the “felt influence of the different conditions of feedback (with the different types of, resp. without sound) during task performance”. Other points of the questionaire consisted of judging eventuell differences in difficulty of the task under the different sonic conditions and in estimating the influence of sonic conditions on each subject’s own average performance time.

3.2.1. Results

As in the initial experiment with the Ballancer in the 19″-screen configuration (3.1), subjects in average performed the task faster with sonic feedback than without. Table 2 shows the times subjects needed to complete the task in average — over all subjects and games — under the different conditions of feedback, in the first (“untrained”) and second (“trained”) series. One interesting new observation with respect to the initial

Table 2: “Average task times”, i.e. times subjects needed to complete the task in average — over all subjects and games — with rolling sound (*rs*), abstract sound (*as*) and without sound (*no*). The third line shows the relative difference of task times in the “trained” series with respect to the respective “untrained” set of equal feedback condition. Line 4 gives the p-value resulting from t-test comparison of these according sets.

	Average task times (ms)		
	rs	*as*	*no*
“untrained”	7623	8388	9286
“trained”	7149	6613	7366
$\delta(\%)$	-6.21	-21.16	-20.67
p	0.261	0.000	0.001

Ballancer experiment (3.1) is the fact that a training effect, i.e. the improvement of performance in doing the task over time — here depicted in line 3 of table 2 is comparatively small with the rolling sound and does not reach a statistical significance of 5%. The training effect is however much higher and about equally strong under the conditions of abstract sound and without sound feedback. In accordance with this phenomenon, average performance in the trained series gets better with the abstract sound than with the rolling sound, while it is best with rolling sound in the untrained series. Table 3 shows a comparison of these different values in the form of relative difference and according statistical significances, i.e. p-values resulting from t-test comparison of the two respective sets of measurements. It can be seen that in the untrained series task performance is significantly faster with rolling sound than without sonic feedback and still faster with abstract sound than without but the latter improvement does not reach statistic significance. In the trained series rolling sound and abstract sound somewhat “switch roles”. Of course these interesting first observations will have to be rechecked with a larger set of subjects.

Also interesting is the comparison of the objective performance measurements as represented in tables 2 and 3 with the subjective judgements of the subjects: Without going into details, as a general summary, 5 out of the 6 subjects liked the rolling sound and stated their conviction that it made the task

[3] i.e. with a temporal precision high enough to represent a continuous process to the human auditory system, exactly: at “cd audio rate” 44,100 Hz. . .

Table 3: Differences in average task times (in %) under the different conditions in the untrained and trained set. Below each difference value the according statistical significance, p, is given.

Differences in performance, statistic significance	"untrained"		"trained"	
	as	*no*	*as*	*no*
rs				
$\delta(\%)$	10.04	21.83	−7.50	3.04
p	0.134	0.005	0.141	0.598
as				
$\delta(\%)$		10.71		11.40
p		0.183		0.021

easier to perform while only 2 subjects claimed the same supporting effect for the abstract sound. In general, the abstract sonic condition was rated as "not fitting", "annoying" or "confusing".

3.3. Summary

The main points with respect to the aspect of perceptual quality of the interaction with the Ballancer in the different feedback conditions shall be shortly summarized. Following the results of the blindfolded identification task (3.1) we can say that the sonic feedback from the rolling model represents the underlying scenario intuitively in the sense of spontaneous, untrained identification. This synthetic cartoon sound is perceived as non-mechanical, electronic or artificial. It can however be attributed a strong clearness in its representation of rolling in the sense of showing less ambiguity in its identification judgements than the direct natural sound of the specific used mechanical realisation of the scenario. The sound of the rolling model is further shown to have an objective informative quality, again intuitive in the sense of occurring without conscious explanation or training, as it is seen to lead to faster performance and optimized control movements that must be attributed to additional velocity information gained from sound. The abstract sonic feedback following a record–needle metaphor turns out to be even stronger in its effect of objective information convection (3.2), however only for a certain trained state of subjects. It can in this sense be seen as less intuitive for unprepaired subjects, however more informative after training. At the same time subjective judgements of the "pleasantness" and "helpfulness" are clearly in favour of the less abstract rolling sound, the abstract sonic feedback was repeatedly judged as "annoying" and "not helpful in performing the target reaching task".

4. Conclusions

This paper shortly summarizes experiences in motor control in a balancing task under different conditions of audio–visual feedback. The example shows that objective performance of an interface configuration and its subjective judgement (by users) may strongly diverge. Also, it is seen that perceptual conditions that are clearly advantageous for unprepaired, untrained users — in this sense more intuitive — may be outperformed by "less natural" feedback configurations after only a relatively short period of usage. The results suggest that an approach of interface design based on ecological considerations and following a notion of cartoonification is of promising potential.

5. References

[1] Kenneth P. Fishkin, Thomas P. Moran, and Beverly L. Harrison, "Embodied user interfaces: Towards invisible user interfaces," in *Proceedings of the Seventh Working Conference on Engineering for Human-Computer Interaction, EHCI 98*, (Heraklion, Crete, September 13-18 1999, pp. 1–18, Kluwer, B.V.

[2] Kenneth P. Fishkin, Anuj Gujar, Beverly L. Harrison, Thomas P. Moran, and Roy Want, "Embodied user interfaces for really direct manipulation," *Communications of the ACM*, vol. 43, no. 9, pp. 75–80, 2000.

[3] M. Rath and D. Rocchesso, "Informative sonic feedback for continuous human–machine interaction — controlling a sound model of a rolling ball," *IEEE Multimedia Special on Interactive Sonification*, pp. 60–69, April 2005.

[4] M. Rath, *Interactive Realtime Sound Models for Human–Computer Interaction — A Sound Design Concept and Applications*, Ph.D. thesis, Department of Computer Science, University of Verona, 2004, available online under <http://www.di.univr.it/dol/main?ent=pubbdip>, "research reports".

[5] "*The Sounding Object (SOb)*," European research project (IST-25287, < http//:www.soundobject.org >) as part of the *Disappearing Computer (DC)* proactive initiative (< http://www.disappearing-computer.org/ >).

[6] W. W. Gaver, "What in the world do we hear? an ecological approach to auditory event perception," *Ecological Psychology*, vol. 5, no. 1, pp. 1–29, 1993.

[7] M. Rath, "An expressive real-time sound model of rolling.," in *Proceedings of the 6th "International Conference on Digital Audio Effects" (DAFx-03)*, London, United Kingdom, September 2003.

[8] P. M. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement," *Journal of Experimental Psychology*, vol. 47, pp. 381–391, 1954.

[9] M.Rath, "Gestural exploitation of ecological information in continuous sonic feedback — the case of balancing a rolling ball," in *Proceedings of the 8th International Conference on Digital Audio Effects (DAFx'05)*, Madrid, Spain, 2005, pp. 98 – 103.

[10] W. W. Gaver, *Everyday listening and auditory icons*, Ph.D. thesis, University of California, San Diego, 1988.