# Informative sonic feedback for continuous human–machine interaction — controlling a sound model of a rolling ball

Matthias Rath, Davide Rocchesso
Dipartimento di Informatica
Università degli Studi di Verona
Strada le Grazie 15
37134 Verona, Italy

***Abstract*— In this contribution, the use of dynamic, interactive cartoon models of everyday sound scenarios together with a familiar, concrete control metaphor, is proposed as a promising approach in sonification. The idea, that aims at a more intuitive, less training-intensive access and control of data-processing is exemplified through a realtime implementation of a audio-visual–tangible system. In the *Ballancer* a virtual ball is balanced on a tiltable track, and the user receives continuous sonic feedback from a realtime sound model of rolling interaction, with or without a simple graphical display. The system is controlled by holding and moving a wooden "balancing-track" whose relevant position is measured with an accelerometer. Evaluation tests demonstrate that users spontaneously understand the sound model and the overall metaphor and perceive and exploit the continuous sonic information to (mostly) unconsciously optimize their control movements.**

## I. INTRODUCTION

ONE widespread way today of improving a user's access to a computerized system and making the procedures of interaction more easily understandable, more intuitive and efficient, is by means of employing a familiar concrete control metaphor. Such metaphors are usually communicated to a user through abstracted but clear representations on in- and output-devices. An example is the common desktop-metaphor for computers, symbolized by the "desktop" screen, with graphical representations of a trashcan, disks, CDs…, with drag-and-drop actions controlled through the mouse. Such employed metaphors and their representations can optimize human–machine interaction and change the way we perceive, handle, manage and exchange stored data and its processing. Even new procedures of data-processing, -acquisition and -control can become possible and emerge through the use of new representations and/or new metaphors for user–machine interaction.

Control metaphors in human-computer interaction are so far allmost exclusively based on graphical representations, just as electronically stored data and its processing in general are usually represented, supervised and controlled graphically. The use of sound is basically restricted to short, distinct signals, mainly affirmations or warnings. This situation is in strong contrast to our "natural", every-day surroundings where sounds constantly inform us about ongoing processes and in particular contiuously give feedback about our own actions and interaction with the environment. Sonification aims at the representation of data through the auditory chanel, which is often achieved through the mapping of variables of interest on properties of a representing acoustic signal, often resulting in abstract sound signals. We believe that one promising approach of encoding and transmitting information acoustically is through enhanced, dynamic, sonic representations of familiar metaphors or scenarios, or the development and adoption of new, rather sound-based or "sound-friendly" metaphors. The perception of ecological information from sound in everyday scenarios has been (and is still further) revealed and examined in a growing number of works of a younger branch of psychoacoustic research. From the *ecological* viewpoint everyday sounds might be seen as the "natural sonification" of information about our environment, expressing attributes of objects and processes that surround us; and human auditory perception "by its nature" is adapted to extract the information transmitted through everyday sounds. In the course of the European project "The Sounding Object (SOb)" this *ecological* approach has been applied to the development of *sound objects* that are clear, *cartoonified* acoustic realtime-models of typical everyday scenarios as "hitting", "bouncing", "breaking", "rolling" or "rubbing". These incorporate a dynamic complex sonic behavior rather than fixed isolated signals and are thus predestined for the sonification of dynamic, in particular unpredictable, e.g. interactive, processes. One conceivable strategy of sonification would be to map the state of an abstract system onto an everyday sounds scenario in the form of a suitable realtime-model; this connection might be established in the form of an intuitively understood metaphor. The work presented here forms an example for this notion and its practicability has been shown in evaluation tests.

Among the scenarios of contacting solid objects, maybe the most important class of everyday sound-emitting processes, the sound of rolling-interaction is particularly rich in ecological information: in addition to attributes generally reflected in contact sounds, such as material or size, it can convey information ([1]) about the direction or velocity of the movement and the shape and surfaces of the contacting objects. All these parameters can be changed dynamically in our realtime sound model of rolling and are immediately reflected in the resulting sound. Following the ideas described above, the rolling-model has been embedded into a simple control metaphor, that of balancing a ball on a tiltable track. The system is accessed through a physical representation of the balancing-track, a 1m wooden stick, whose relevant position is measured through an accelerometer. The complete tactile–audio-visual device, the "*Ballancer*" has been evaluated in performance tests that demonstrate how users spontaneously understand sound and control metaphor and how they perceive and exploit the conveyed ecological information.

## II. A SOUND MODEL OF ROLLING

As an example of the sound models developed in "The Sounding Object" project, our model of rolling is realized through a hybrid approach to sound design ([2] gives an overview). At the core we make use of an algorithm that is based on a physical model of contacts of solid objects and does not rely on prerecorded/predefined sound samples. This choice is advantageous because:

- The generated sound is free of repetitions and reacts continuously and dynamically upon variations of the driving parameters such as position and velocity. Since the model runs and reacts in realtime, parameters may come from unpredicted sources, e.g. user input as in the audio-visual–tactile device presented in section III.
- There is no need to store, maintain and process large banks of sound samples.
- Ecological attributes such as the size, mass or regularity of the rolling object, as well as the material or fine structure of the supporting surface can be varied along a continuum and are directly reflected in the acoustic output.

While most works of physical modeling aim at the possibly realistic simulation of highly specialized physical systems, mostly musical instruments, at high costs of implementation and specialization, we follow a rather untypical "cartoon"-approach. We try to be rather general and flexible to cover a wide range of everyday scenarios and prefer a clear, possibly exaggerated, cartoon-like, expression of ecological attributes at a lower computational burden to enable realtime realization in an interface context. For these reasons, we surround the central physics-based model with higher-level structures that remind of traditional techniques of sound synthesis. As a result we can combine the abovementioned advantages of physics-based modeling with

- flexibility and efficiency in implementation, and
- a sharper elicited ecological impression.

The main stages of this *hybrid architecture* are sketched in the following, details have been described in previous papers [2], [3].

### A. A Physics-based model of impact interaction

The distinctive character of rolling-sounds may be partly due to the nature of rolling as the most prominent continuous interaction process where mutual contact forces are exerted perpendicularly to contact surfaces: in contrast to slipping, sliding or scratching actions, additional friction forces parallel to the surface are comparatively small [1].

We thus base our work on a physics-based model of impact interaction [4], that has successfully been used to generate sounds of hitting, bouncing and breaking [5]. In our work, as in several related works, the resonating objects are described according to the modal formalism [6], [7] which supports particularly well our main design approach for its physical generality and, at the same time, for its intuitive acoustic meaning. To excite the modes, we use a dynamic term of interaction force that depends on the instantaneous values of position and velocity of the interacting objects [4]:

$$f(x(t), \dot{x}(t)) = \begin{cases} -k(x(t))^{\alpha} - \lambda(x(t))^{\alpha} \cdot \dot{x}(t), & x > 0, \\ 0, & x \leq 0. \end{cases} \quad (1)$$

Here $x$ measures the compression of bodies in contact ($x \leq 0$ corresponding to no contact), $k$ is the elasticity constant (the main parameter that controls the hardness of the impact), $\alpha$ accounts for the local geometry of the contacting objects, while $\lambda$ weighs the dissipation of energy during contact, accounting for friction loss.

The physics-based model involves a degree of simplification and abstraction — all quantities in the interaction term are one-dimensional — that implies efficient implementation as well as adaption to a broad range of impact events. In contrast to previous works on synthesis of contacts sounds [7] that focus on the resonance-/decay behavior of the interacting objects, the algorithm is reactive and dynamical: complex transients are produced that depend on the parameters of interaction (such as hardness) as well as the attributes and momentary states of the contacting objects. That dynamical quality is particularly important in situations of repeated, frequent or constant contact, as in the case of rolling.

### B. Reduction of rolling-geometry

The previously described formula of interaction force and the modal equations of the interacting objects are transferred into a discrete-time algorithm to achieve a realtime implementation of the impact model. Expanding this straight physics-based approach — description in differential equations, discretization, implementation — to the full 3-dimensional rolling scenario in all details would lead to high computational complexity. Therefore, instead of following a brute-force strategy, we reduce the global rolling geometry to the one dimension of the developed impact model and take care of specific macroscopic characteristics explicitely in a separate stage. The process is sketched in the following, [3] contains more detailed explanations.

The acoustic vibration in a rolling-scenario has it's cause in the structures of the contacting surfaces; no sound would emerge if the rolling object and the plane (on which it is rolling) had perfectly smooth surfaces. As an object rolls, the point of contact moves along it's surface and along the plane. These "tracked" surface profiles are the source of the

---

[1] Probably the main notion behind the invention of the wheel.

acoustic vibration in rolling-interaction. If we restrict our view on the scenario to the one dimension perpendicular to the plane, we see a time-varying distance constraint on the interacting objects (i.e., the rolling object and the plane). This constraint takes the form of a temporarily changing distance offset that adds to the compression variable $x$ in (1) . The offset signal however is not simply the difference of the two surface profiles, since only certain peak points on both surfaces are possible points of contact. The position constraint on the rolling object is given by the hypothetical trajectory of motion along the plane at zero constant distance, with contact at peaks, without “bouncing back” or “enforced contact”. Figure 1 shows this final offset-signal, as it is derived from the surface profile. The actual movement of the rolling object differs from this idealized trajectory due to inertia and elasticity. Notice that it is exactly the consequences of these physical properties that are described by and substantiate the use of the impact model.

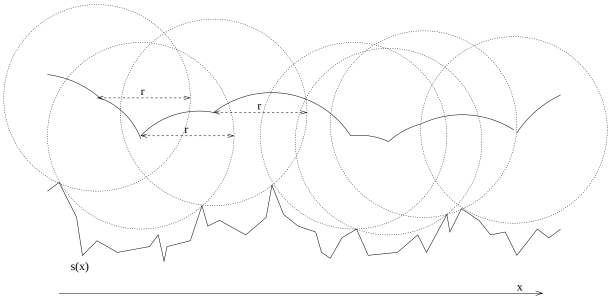

Fig. 1. *Sketch of the effective offset-curve, resulting from the surface* $s(x)$.

The derivation of the offset-curve from a given surface profile is, in a straight approach, highly demanding for real-time processing. An efficient recursive algorithm has been developed that solves the task at a strongly reduced cost. As origin for the surface profile we again preferred a dynamic, controllable cartoon-type source over a prestored table, even though it is always possible to hold a scanned real surface profile. One surface model, frequently used in computer graphics, is fractal noise, i.e. noise with a $1/f^{\beta}$ power spectrum, where $\beta$ reflects the fractal dimension or roughness. Low- and high-pass filtering however showed to be efficient for the neccessary consideration of various smoothing processes of typical rolling-surfaces. Under these circumstances the fractal parameter $\beta$ is less influential and band-pass filtered white noise proved to be a simple and efficient choice.

### C. Higher-level modeling

Of high perceptual importance are periodic patterns of timbre and intensity in typical rolling sounds. They may originate from periodic features of the rolling-surface (e.g., tiles) or from irregularities of the shape of the rolling object. These periodic modulations appear as a strong perceptual cue of rolling, as opposed to e.g. sliding, and their frequency is particularly important for the perception of size and speed [1]. Macroscopic deviations of the rolling-shape from perfect sphericity — or more general, asymmetry of the object with respect to its center of mass — lead to modulations of the effective gravity force that holds down the rolling object. In our model, such effects are explicitely accounted for by according parameter modulations, since the physics-based core is one-dimensional and does not cover higher macroscopic geometries.

If a rolling object does not show perfect circular symmetry with respect to its center of mass, the height of the center of mass will vary during the movement (see figure 2). This varying height is related to according variations of the potential energy of the object in the gravity field, and it is reflected by variations of the effective force that is holding down the rolling object. The exact modulating terms of energy, forces and velocities can only be determined if the shape of the object is known exactly (as solutions of the differential equation given by stating the principle of energy conservation). However, if the goal is ecological expressiveness rather than simulation for its own sake, we have to consider that the exact shape of a rolling object is rather not perceived from the emitted sound. But a general idea of “asymmetry” may be given acoustically.

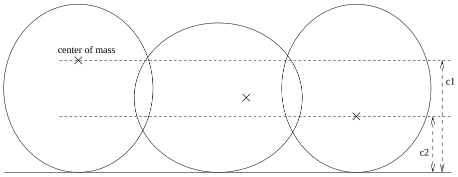

Fig. 2. *Sketch of a rolling object with weight and shape asymmetries at different instants*

One simple assumption is that the oscillating (in the sketch of figure 2 between the extrema of $c1$ and $c2$) height of the center of mass $c(t)$ is approximately described by a sinusoid [2]:

$$c(t) = (c2 + c1)/2 + (c2 - c1)/2 \cdot \sin(\omega t) \ . \tag{2}$$

The effective force-term between the two contacting objects (the rolling and the plane) in addition to gravity is then proportional to vertical acceleration, which is the second derivative of (2):

$$\ddot{c}(t) = -(c2 - c1)/2 \cdot \omega^2 \cdot \sin(\omega t) \ . \tag{3}$$

We use such a sinusoidal force modulation term whose frequency is related to the transversal velocity and radius of the object. Equation 3 shows that the amplitude is proportional to the square of the angular velocity. The proportionality constant $(c1 - c2)$ allows to express an overall amount of deviation from perfect spherical symmetry.

## III. CONTROL METAPHOR AND INTERFACE

For a control interface to access the developed sound-model of rolling we chose a particularly simple metaphor, that of balancing a ball on a tiltable track. The (virtual) ball is free to move along one axis over the length of the track, beeing stopped or bouncing back when reaching the extremities. The acceleration of the ball along the length of the track is directly related to the vertical angle (as sketched in figure 3). More exactly, if the track forms an angle $\alpha$ with the horizontal plane,

---

[2] This is e.g. the case for a spherical object whose center of mass is displaced from the geometric center.

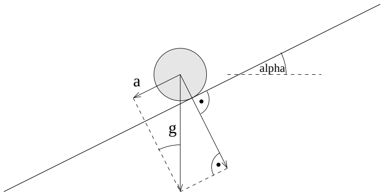

Fig. 3. *Scheme of a ball rolling on a tilted track. The gravity acceleration is split into two terms parallel and perpedincular to the track, according to the track-angle.*

the acceleration $a$ along the track results from the vertical gravity acceleration $g$ via

$$a = g \cdot \sin(\alpha)\ . \tag{4}$$

We neglect any effects of the changing vertical ball position induced by tilting the track. Further, all damping of the ball movement through friction on the track and in the air is modeled by one term of friction force $f$, proportional to the instantaneous velocity $v$ along the track: $f = -k{\cdot}v$ Finally, in considering the ball's displacement along the track, all effects of rotation, such as the moment of inertia, are ignored. The position $x$ of the ball on the track is described by the following differential equation:

$$\ddot{x} = \sin(\alpha) \cdot g - k \cdot \dot{x}\ . \tag{5}$$

The rationale of the system metaphor is substantiated by the following points:

- The simplicity of the idea supports a robust realization. In fact, a previous attempt to embody a more complicated metaphor was only partially successful due to sensitivity to imperfections in the practical implementation [8].
- The general principle of the balancing-metaphor, as well as its haptic control, is familiar from everyday experience. It should thus be easy to understand for an average user, even without explanation and after very little training.
- Working on the same general balancing notion, the system could easily be expanded, e.g. to a two-dimensional plane.
- The metaphor can be adapted to a wide range of control tasks. The system can be seen as a possibly simple representation of a controlled system that is reacting with inertia.

Another strong advantage is that the physical, purely mechanical realization of the metaphor is straightforward. For instance, in our implementation the control track can also hold a real ball moving on its top surface. In this way the virtual system can be directly compared to its mechanical pendant, to measure how far it is from the "real thing".

### *A. Implementation*

The complete software part of the tangible-audible interface runs with low computational load on a standard personal computer. The sound component is realized with `pd` ("*pure data*"[3]), a free software that allows to combine predefined sound processing blocks and custom-made plugins using a graphical interface. Specific signal processing algorithms, which are discrete-time realizations of the *impact*, the *modal resonators*, and the *rolling-filter*, are programmed in `C` as `pd` plugins. The motion of the ball (5) is also transferred into discrete time, but at a rate (in the range of $100$ Hz) much lower than audio sampling rate. The resulting calculation as well as higher-level structures of the rolling-model are defined by means of the `pd`-GUI. A schematic graphical representation of the balancing track and the rolling ball is implemented in `gem`, an `openGL` rendering-extension for `pd`.

The interface is physically controlled by holding and tilting the rolling track, a $1.05$m wooden bar. This tangible controller has an aluminium track attached to its upper face which can hold a glass marble of $2.5$ cm diameter rolling along the track. Fixed to the rolling track is an accelerometer that measures acceleration in the direction of the length of the track. This measured acceleration is the fraction of gravity in this direction, as described in equation (4). We can thus calculate the tilt angle from the accelerometer output, again using the `pd` environment. The data-transfer from the (analog) accelerometer to the software is established through a `Kroonde`[4] sensor wireless interface, connected to the computer via a UDP socket-connection.

Fig. 4. *The "rolling-track" with a glass marble rolling in its upper-face aluminium track.*

## IV. EVALUATION

A user evaluation has been done to test and prove the capability of the tangible–audio-visual device to convey information, following the mechanisms of *ecological perception* as outlined in the introduction. One important point of our work is the scope of achieving intuitive understandability through the use of everyday (instead of abstract) sound as well as a familiar control metaphor. This requires at first rank, that subjects associate the modeled sound and the whole device in its feedback behavior in the intended way. The initial (shorter) part of the evaluation test adresses this aspect in a short questionaire following listening- and controlling-trials. In a second step it is demonstrated that subjects also perceive and exploit the continuous information conveyed through the

---

[3] http://www.pure-data.org

[4] http://www.la-kitchen.fr

interface and through the sonic channel in particular. To that end, an indirect strategy is used, rather than an analysis of provoked verbal or in some other way formalized user response. The reason for this is, that direct questions about the perceived continuous attributes — in the point of focus here are velocity and position of the ball — might reveal a conscious reaction of the test subject on the question rather than a spontaneous perception. As an example, one might think of a sorting- or scaling-task with generated sounds of various velocity or explicit questions about the (development of the) velocity of an acoustically modeled rolling object. It is possible that subjects connect a sound with a faster modulating amplitude to a faster moving object when they are suggested (or even forced) to make a choice, although they would not spontaneously have this association without being asked. We believe that such processes of perception may often be unconscious and thus hard to verify through questions resulting in a conscious answer. The approach here is therefore somewhat indirect and more complicated. Subjects are asked to perform a specific control task, and their movement while solving the task (with and without acoustic/graphic feedback) is recorded and analyzed. From systematic differences in the subjects' movements under the different sensory conditions we can conclude that the information they (the subjects) perceive depends on the stimulated sensory chanels, and finally support the conjectures given above. This "indirect" strategy of using a performance task enables us to illuminate processes of perception and information processing that may be unconscious; in fact reactions during the tests underline the unconscious nature of the process. We can furtheron, besides proving the absorption of different sensory information, demonstrate their immediate active exploitation in human control gestures and give some useful [5] quantitative measures.

A total of 10 subjects participated in the evaluation. They were 6 men and 4 women of age between 23 and 32, chosen among students of electrical engineering at the Technical University of Sweden (KTH, Kungl. Tekniska Hoegskolan). The subjects had no or little (3 subjects) experience with electronic sound production and no knowledge about the work presented here. Each session (with one subject) consisted of a shorter, ca. 20 mins, listening/trying+interview-part, and a longer, ca. 1h, performance-part as explained above. For their participation in the experiment, subjects were paid 80 Swedish Crowns (ca. 9 Euro) each.

### *A. Recognition of sound and metaphor*

To examine if and, if any, which spontaneous association, the sound generated by the rolling-model provokes by itself, subjects were at the beginning of their testing session played two short sound examples of the model, each followed by the question "What do you hear?". The sounds were presented through headphones without previous information about their origin and background whatsoever. Both sounds were generated with parameters according to a small hard ball of 2.5cm diameter rolling on a hard surface rather fast in the beginning, then subsequently slowing down to a stop as if being rolled on a horizontal surface. One of the two sounds also contained a few accelerating initial bounces as if the ball was being dropped on the surface and rebouncing for some period before finally rolling. The order of presentation of the two sounds (with the subsequent question) was varied, "no-bouncing – bouncing" for one half of the subjects, opposite for the other 5. The motivation of this choice of two sound examples was to test if such a typical starting dropping incident contributes to the identification of the scenario. Previous informal experience had suggested this conjecture.

Next, blindfolded subjects were given access to the balancing-track and asked to carefully move up and down their arm holding the track. After a free period of testing the device in its sonic reaction to the movement, they were again asked the same question to identify "what they heard".

The direct output of the model was taken, i.e. the modeled vibration of the object at one point without any consideration of spatial sound propagation[6]. As the only form of post-processing, the right-to-left movement was acoustically displayed through simple amplitude panning.

Later, the previous test procedure was repeated, this time with a real glass marble of ca. 3cm diameter rolling on the track (replacing the virtual ball and synthesized sound and display). Blindfolded subjects were made listen to the sound of the small marble and again asked "What do you hear?", finally given access to the track as before followed by the same question.

The main results of this identification test are:

- Overall association of the synthetic sound with rolling was very high: For the sound example with initial bouncing, 9 of the 10 subjects immediately described what they heard in terms of a scenario of a rolling object. Only one subject heard "a starting engine". Surprisingly, after the informal expectations mentioned above, even all 10 subjects identified the "no-bouncing" sound example as "rolling".
- The sound of the small glass marble rolling on the track in front of blindfolded subjects turned out to be more ambiguous than the synthesized sounds. While clearly hearing something moving in front of them, 4 of the 10 subjects perceived several (not only one) objects rolling on the track, or at least were not sure about the presence of only 1 or several objects. One subject heard "something like a toy car", 2 others heard the object "inside a tube". Also, the diameter of the marble was regularly guessed much smaller than the real size of 3cm, typically around 0.5 – 1cm. The size of the virtual ball instead was described to lie between 1 and 3 cm, much closer to the intended diameter of 2.5cm.
- When controlling (blindfolded) the tangible–audible device with the synthesized sound-feedback, all 10 subjects clearly described an object rolling from side to side, steered by the height of the held end. 8 of the 10 subjects even described the construction of a tiltable track or pipe (without having seen any parts of it or other hints...).

---

[5] ...e.g. in concrete implementations,

[6] A mechanical pendant would be the signal as picked up by a contact microphone.

- The ambiguity in the (purely acoustic, blindfolded) perception of the mechanical scenario did not diminished when subjects were given access to the track and were allowed to control it. Remarkably, the identification of the scenario changed for some subjects when they were allowed to control it, but overall the recognition of the de–facto scenario did not improve.

Summarizing the results of the questions about the sounds and the tangible–audible device, we can say that the modeled metaphor is intuitively understood. The combination of modeling everyday sounds and using a familiar control metaphor here exhibits the advantage that virtually no explanation and learning are necessary. As opposed to what happens with abstract sounds/controls [9], users may immediately understand and react to transported information without being instructed. The emergence of a mental model is even more clear for the tangible–audible interface than for the actual mechanical device that provides a physical realization of the metaphor. This demonstrates how effective the *cartoonification* approach to sound modeling can be: although the device is perceived as fictitious, nevertheless it can quite reliably elicit an intended mental association, even more clearly than the real thing.

### B. *Performance measurement*

The second part of the evaluation test adresses the question if users, in addition to “understanding”, i.e. identifying and appropriately using, soundmodel and interface, perceive dynamic ecological attributes contained in the sound and make use of this information. In the system under test, the dynamic attributes are the position and velocity of the ball since other parameters, such as the size, weight, hardness and *sphericity* [7] of the ball, the 2-dimensional direction of its movement or the structure of the surface are kept fixed. All these additional variables might be changable dynamically in future enhancements to increase the informative potential in sonification tasks. We here focus on the two parameters that are considered the most salient, in accordance with the consideration that also in physical reality the situation of a rolling object that changes in form and size during the movement is rather unfamiliar.

Subjects were asked to perform a specific task, consisting of moving the virtual ball from a resting position at the left end of the track to a marked area of 15cm length slightly on the right of the center. The edges of the target area were located 10cm and 25cm right from the center, i.e. 60cm and 75cm from the left track end. In order to examine if and how subjects perceive and use information about the movement (position and velocity) of the controlled ball through different sensory chanels the task was presented and performed with various configurations of sensory feedback and the subjects' movements were recorded.

The target area was indicated in the graphical representation by a different color, and a rougher surface with a “furrowed” structure was used in the sound model. During the movement of the ball the surface profile at its momentary position is constantly reflected in the sound; vice versa, the emitted sound informs about the ball's position on the surface, since the latter does not change in form. More exactly, after the above description of the *Ballancer* it can be heard [8] if the ball is currently moving inside or outside the target area. In particular, an abrupt change of the surface structure, further underlined by a little “step” due to the different depths of surface profiles, marks the moments when the ball enters and leaves the target area. As the model does not consider any spatial sound propagation, the momentary position of the ball is furtheron expressed only through simple stereo amplitude panning between left and right. From this behavior of the sonic feedback, the position of the ball can be perceived with much less precision than it can be perceived in a good [9] visual representation. However, the following tests show that subjects do generally understand the position information contained in the sound of the *Ballancer*, at least to the extent necessary to perform the test task with purely auditory feedback.

Subjects were asked to try and accomplish the task as fast as they could. They were not informed anyhow about the measured time needed in the trials, in order to minimize effects of conscious adaptation to the test conditions and isolate the effects of unconscious mechanisms applied by the subjects trying to optimize (subjectively) their performance.

Feedback about the position and velocity of the virtual ball during the trials was given acoustically through sound from the rolling-model as described above (section II) and/or visually on the computer screen, as a schematic representation of the ball on the track . The graphical display, with the ball represented as a monochrome sphere on a line representing the track and the target area marked by a different color, was realized in 4 different sizes. Scaling factors for graphical display were ranging from 12, the largest, with the track horizontally filling the computer screen, over 4 and 2 to the smallest, 1. In the latter, smallest, size, the moving sphere (representing the ball) could not always be visually detected due to the bounderies of the screen resolution .

Each test started with 20 training runs (10 + pause + 10) with the largest display (“full screen”, scaling factor 12) and sound feedback, to minimize possible training effects. In the following runs, the needed time was measured with display sizes of 12, 4, 2 and 1 (in this fixed order). Again 20 measurements were made for each size, 10 times with and 10 times without sonic feedback. The order of the measurements, “with–without” resp. “without–with sound”, was switched after half of the subjects to test for, and eventually counterbalance an effect of the order of performance on the results.

The main results of the performance task can be summarized as follows:

- For all display sizes, the average time needed to perform the task improves significantly with the auditory feedback from the model. Table I shows the average

---

[7] The meaning of this parameter is explained as part of the description of the rolling-model in section II-C.

[8] This informal examination is in fact proven by the results of the following tests.

[9] E.g. in comparison to a graphic display spanning a standard computer screen. . . It is not the subject of this text to further specify the quality of graphical displays or compare the resolution of position in possible graphical displays with the sound model.

TABLE I
AVERAGE TIMES NEEDED TO PERFORM THE TARGET–REACHING TASK BY THE TWO GROUPS OF SUBJECTS WITH OPPOSITE PRESENTATION ORDER (“WITH–WITHOUT SOUND” AND VICE VERSA) AND THE SET OF ALL 10 SUBJECTS, TOGETHER WITH THE PERCENTUAL DIFFERENCE “WITHOUT SOUND” TO “WITH” AND THE STATISTICAL P-VALUES FOR THE TWO ACCORDING SETS OF MEASUREMENTS.

| sub- jects no. | average task time (ms) at various display sizes, with (+) and without (–) sound, percentual difference ($\delta$) and statistical significance (p) | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | scale factor 12 | | | | scale factor 4 | | | | scale factor 2 | | | | scale factor 1 | | | |
| | + | – | $\delta$ (%) | p | + | – | $\delta$ | p | + | – | $\delta$ | p | + | – | $\delta$ | p |
| 1 – 5 | 5387 | 5805 | 7.8 | 0.203 | 5968 | 6771 | 13.5 | 0.135 | 6157 | 7392 | 20.1 | 0.015 | 8340 | 14242 | 70.8 | 0.000 |
| 2 – 6 | 5048 | 5545 | 9.9 | 0.063 | 5320 | 5941 | 11.7 | 0.086 | 6228 | 7174 | 15.2 | 0.095 | 8003 | 12346 | 54.3 | 0.018 |
| 1 – 10 | 5217 | 5675 | 8.8 | 0.031 | 5644 | 6356 | 12.6 | 0.029 | 6192 | 7283 | 17.6 | 0.004 | 8172 | 13188 | 61.4 | 0.000 |

performances for the two groups of subjects with different presentation order and the set of all subjects at the various display sizes, with and without sound. The two respective neighboring columns contain the relative difference, “no sound” to “with sound” (in %, $\delta$) and the statistical p-value for the according set of measurements. p-values of ($\leq 0.05$) or near ($\leq 0.1$) statistical significance are highlighted in green. It can be seen that the average task time for the set of all subjects as well as for both subgroups improves (i.e. gets shorter) with the auditory feedback for all display sizes, corresponding to only positive $\delta$ values in table I. These performance improvements, ranging from around 9% for the largest to around 60% for the smallest display, are always statistically significant for the whole set, while they reach statistical significance for the subgroups only for the smaller displays. Since significance is reached for the whole set of subjects, we can expect that it would be found independently of the order of presentation with a larger set of measurements, using more subjects or more trials per subject.

- An analysis of the recorded user movements reveals that the improved performance is, at least partly, explainable through a different, more efficient behavior of acceleration and stopping of the virtual ball already before reaching the target area. This disproves the belief that the performance improvement could be due only to an additional notification through the change of sound when the ball enters the target area. If this was the case, the continuous feedback might not be neccessary, but be replaced through any simple short signal of notification. Again, the measured/derived indicators showing the different quality of user movement are statistically significant. Exact details are being prepared for publication [10].
- No statistically significant difference between the performance times of groups 1 and 2 is found. This means that any possible training effect that might still be present after the $2 \times 10$ trials is not strong enough to reach statistical significance. In other words, subjects adapt very quickly to the usage of the device, as intended by the used familiar metaphor and sound scenario.

## V. CONCLUSION

We propose models of every-day sounds to be a valuable complement to the common use of abstract signals for sonification tasks. On one hand, the mapping of given data to everyday sounds may be more delicate than the creation of suitable abstract sonifications, since it is restricted by the chosen class of sounds. Also, human auditory perception in terms of signal parameters like fundamental-frequency/pitch has been better formalized through longer dedicated psychoacoustic research than mechanisms of *everyday listening*. On the other hand, everyday sounds may offer a high degree of intuitive understandability, i.e. the potential of being spontaneously identified in a desired way and of transporting information in terms of perceived ecological attributes without or with very little training and explanation. Information may be absorbed and processed by a user without awareness/attention which may set free cognitive resources and decrease the need of conscious learning. This latter advantage is supported by the use of clear, familiar embedding control metaphors.

We have demonstrated these potentials of everyday sound models at our example of rolling-interaction. The sound model has been embedded in a simple control metaphor with a tangible audio-visual interface. User evaluation tests have proven the proposed intuitive understanding of the sound, the overall metaphor and conveyed information, and shed some light on the unconscious absorption and processing of acoustic ecological information.

## VI. ACKNOWLEDGMENTS

This work has been partially supported by the European Commission under contract IST-2000-25287 (project “SOb - the Sounding Object”: `www.soundobject.org`). The experiment has been conducted by Matthias Rath as part of his visit at the Royal Institute of Technology (KTH) of Stockholm, Department of Speech, Music, and Hearing, supported with the Marie Curie TMR grant HPMT-CT-2000-00119.

## REFERENCES

[1] M. M. J. Houben and C. N. J. Stoelinga, “Some temporal aspects of rolling sounds,” in *Journée design sonore à Paris*, Paris, France, 2002, < http://confs.loa.espci.fr/DS2002/ >.

[2] M. Rath, F. Avanzini, N. Bernardini, G. Borin, F. Fontana, L. Ottaviani, and D. Rocchesso, “An introductory catalog of computer-synthesized contact sounds, in real-time,” in *Proc. Colloquium of Musical Informatics*, Florence, Italy, May 2003, pp. 103–108.

[3] M. Rath, “An expressive real-time sound model of rolling,” in *dafx03*, 2003, pp. 165–168.

[4] F. Avanzini and D. Rocchesso, “Modeling collision sounds: Non-linear contact force,” Dec. 2001, pp. 61–66.

[5] F. Avanzini, M. Rath, and D. Rocchesso, “Physically–based audio rendering of contact,” Lausanne, Aug. 2002.

[6] J. M. Adrien, “The missing link: Modal synthesis,” in *Representations of Musical Signals*, G. D. Poli, A. Piccialli, and C. Roads, Eds. Cambridge, MA: MIT Press, 1991, pp. 269–297.

[7] K. van den Doel, P. G. Kry, and D. K. Pai, “Foleyautomatic: Physically-based sound effects for interactive simulation and animation,” in *Proc. ACM SIGGRAPH*, August 2001, pp. 537–544.

[8] D. Rocchesso, F. Avanzini, M. Rath, R. Bresin, and S. Serafin, “Contact sounds for continuous feedback,” in *Proc. Workshop on Interactive Sonification*, Bielefeld, 2004.

[9] S. Brewster, “Non–speech auditory output,” in *The Human-Computer Interaction Handbook*, J. Jacko and A. Sears, Eds. Lawrence Erlbaum Publishers, 2002, pp. 220–239.

[10] M. Rath, “Interactive sound-models for human-computer-interaction — a sound design concept and applications,” Ph.D. dissertation, Dipartimento di Informatica - Università di Verona, 2004.

**Matthias Rath** Matthias Rath received his master in mathematics (algebraic geometry) in 1998 from the University of Bonn, Germany. From 1999 to 2001 he has been studying electronic composition at the Folkwang-Hochschule Essen, until the intermediate diploma. Since 2001 he is doing PhD studies in informatics at the University of Verona, Italy, where he has also been working for the SOb European project. In 2003, Rath worked at the Department for Speech, Music and Hearing of the Royal Institute of Technology in Stockholm, Sweden, under a Marie-Curie scholarship. His main interests are sound sysnthesis, physical modeling, and sound and interaction design.

**Davide Rocchesso** Davide Rocchesso received the Laurea in Ingegneria Elettronica degree from the University of Padova, Italy in 1992, and the Ph.D. degree from the same university in 1996. In 1994 and 1995 he was visiting scholar at the Center for Computer Research in Music and Acoustics (CCRMA) at Stanford University. Since 1991 he has been collaborating with the Centro di Sonologia Computazionale (CSC) at the University of Padova as a researcher and live-electronics designer. Since 1998 he has been with the Dipartimento di Informatica at the University of Verona, Italy, where he is now Associate Professor. His main interests are in sound processing, physical modeling, and human-computer interaction.