AN EXPRESSIVE REAL-TIME SOUND MODEL OF ROLLING

Matthias Rath

Dipartimento di Informatica
Università degli Studi di Verona
rath@sci.univr.it

ABSTRACT

This paper describes the structure and potential of a real-time sound model of "rolling". The work has it's background and motivation in the *ecological* approach of psychoacoustics. Scope of interest is the efficient and clear (possibly exaggerated) acoustic expression, *cartoonification*, of certain ecological attributes rather than realistic simulations for their own sake.

To this end, different techniques of sound generation are combined in a hybrid hierarchical structure. A physics-based algorithm (section 2) of impact-interaction at the audio-core is surrounded by higher-level structures that explicitely model macroscopic characteristics (section 5). Another connecting audio-level algorithm, the "rolling-filter", reduces the (3-dimensional) geometry of the rolling-contact to the one dimension of the impact-interaction-model (section 3).

1. INTRODUCTION: BACKGROUND AND MOTIVATION

The use of enhanced auditory display in ambitious Human Computer Interaction has been more and more widely recognized to be of major importance; this also holds for real-time reactive non-speech audio in particular. If such non-speech auditory display is to be intuitively understandable for non-expert listeners, models that acoustically convey ecological attributes are a natural, promising choice (e.g. compared to abstract acoustic signals). While psychoacoustic studies of *everyday listening* (as opposed to *musical listening*[1]) form an increasingly broad basis, the respective development of ecologically expressive, flexible and computationally affordable sound models has gained much less attention so far. In the field of sound synthesis, conventional techniques (such as additive, subtractive or FM synthesis) are based on signal-theoretic parameters that are known to closely relate to traditional musical terms (such as pitch and volume). *Physical modeling* on the other hand, that naturally connects to ecological, physical attributes, has made remarkable progress, but most works focus on the possibly realistic, hence often rather inflexible, computationally demanding simulation of highly specialized systems, usually musical instruments. Instead, our work aims at effective, not necessarily perfectly realistic, interactive real-time "sound cartoons" that express aspects of familiar everyday scenarios. To this end physics-based models in the direct sense (i.e. based on the numerical solution of differential equations) are combined with rather perception-oriented structures, that remind of and try to exploit the flexibility, efficiency and intuitive accessibility of older methods of sound synthesis.

Collisions of solid objects form an important class of sonic events in everyday listening. The perception/estimation of ecological information from contact sounds, *structural invariants* [1], i.e. attributes of involved objects such as size, shape, mass, elasticity, surface properties or material, as well as *transformational invariants*, such as velocities, forces and position of interaction points, is common experience. The sound produced by a rolling object is generally particularly rich in ecological information. In addition to characteristics of involved objects that are generally reflected in contact sounds, rolling-sounds may carry further details of form or surface as well as transversal velocity, gravity or acceleration/deceleration. Also, rolling-scenarios form a category that seems to be characteristic from the auditory viewpoint, in the sense that the produced sound is often recognizable as such, and in general clearly distinct from sounds of slipping, sliding or scratching interactions, even of the same objects. All this suggest acoustic modeling of rolling to be a rewarding goal under the various demands of auditory display.

Physical models of rolling however, tend to get highly complex and computationally demanding; the derivation of abstractions that simplify the process and allow realtime implementation, yet keep (or better: stress) main characteristics, in the abovementioned sense of "*cartoonification*", becomes thus a rewarding challenge. We use a physical model of impact interaction that can reflect attributes as mass or hardness in complex transients currently not fully covered by signal-theoretic approaches. Instead of expanding this one-dimensional physical model to the much more complex rolling-interaction, at costs of computation, specialization and control not suitable in our context of flexible and clear realtime-*cartoonifications*, we use higher-level structures to reduce a rolling-scenario to impact-interaction. Certain macroscopic features, like the global geometry and the transversal velocity, then have to be accounted for explicitely under perceptual considerations, since they are not "automatically" reproduced by the abstracted model.

2. A PHYSICS-BASED MODEL OF IMPACT AS LOW-LEVEL BASIS OF CONTACT SOUND

The mentioned distinctive character of rolling-sounds may be partly due to the nature of rolling as **the** continuous interaction process, where the mutual force on the involved objects is basically that of an impact perpendicular to the contact surface: in contrast to slipping, sliding or scratching actions, additional friction forces parallel to the surface are comparatively small [1]. It seemed thus promising, to model rolling-sounds using higher-level structures around a model of impact-interaction [2], that has successfully been used to generate sounds of hitting, bouncing and breaking [3].

The basis of the algorithm is a physical model of a (1-dimen-

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

sional) impact interaction force

$$f(x(t), \dot{x}(t)) = \begin{cases} -k(-x(t))^{\alpha} - \lambda(-x(t))^{\alpha} \cdot \dot{x}(t), & x < 0, \\ 0, & x \geq 0, \end{cases} \qquad (1)$$

where x is the difference of two (as well 1-dimensional) variables connected to each object. In the standard case of examined movements in one spatial direction, x is the distance variable in that direction; $x > 0$ is the case of no contact. k is the elasticity constant, i.e. the hardness of the impact. α, the exponent of the non-linear terms accounts for the local geometry of the contacting objects, while λ weighs the dissipation of energy during contact, accounting for friction loss, [2] contains detailed information.

The behavior of interacting objects can be described in different ways; in all our present modeling efforts both resonating objects are in modal description ([4, 5]), which supports particularly well our main design approach for its physical generality and, at the same time, for its intuitive acoustic meaning [6].

The physical model involves a degree of simplification and abstraction that implies efficient implementation as well adaption to a broad range of impact events. At the same time, the algorithm is reactive and dynamical, in contrast to signals generated with other synthesis techniques: 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.

3. REDUCTION OF LOCAL ROLLING-GEOMETRIES TO ONE (IMPACT-) DIMENSION

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 plain (on which it is rolling) had perfectly smooth surfaces — or at least, no other than a possible "bouncing"-like vibration that can as well occur in a pure impact contact along one axis. In fact, as an object rolls, the point of contact moves along it's surface and along the plain. These "tracked" surface profiles are the source of the acoustic vibration in rolling-interaction.

If we restrict our view on the scenario to the one dimension perpendicular to the plain, the tracked surface profiles, exactly their difference, give rise to a time-varying distance-constraint on the interacting objects (i.e. the rolling object and the plain). This constraint takes the form of a temporarily changing distance-offset that adds to the distance variable x in equation 1 as it would emerge from the movement of the interacting objects. In other words, the surface profiles are the origin of a dynamic offset signal that has to be fed into the impact model, namely added to the distance-variable x, thus causing vibration of the contacting objects. Exact investigation however reveals, that the appropriate offset signal is **not** simply the difference of the surface curves, as scanned along the rolling trajectory: not all these surface points (along the trajectories) are possible points of contact. Figure 1 shows the principle of rolling-typical "bridging" of surface details. The rolling object is here assumed to be locally perfectly spherical without microscopic details; this simplification is possible, since deviations from that ideal geometry can be carried forward to the associated profile of the plain. It is seen, that only certain surface "peaks" are potential contact points. The **hypothetical** trajectory of the rolling object, i.e. precisely its center, as depicted in figure 2, as it would

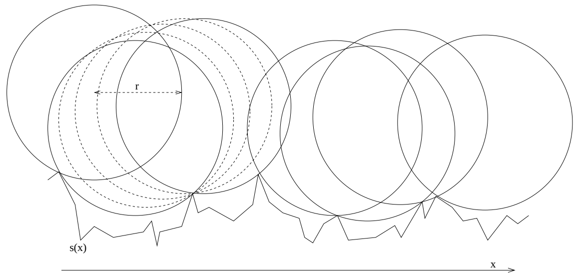

Figure 1: *Sketch of the fictional movement of a ball, perfectly following a surface profile s(x). Relative dimensions are highly exaggerated for a clearer view. Note that this is* **not** *the de-facto movement; this idealization is used to derive the offset-curve to be used by the impact-model.*

move along the plain at constant distance 0 contacting the plain exactly at these peaks (without "bouncing back" or "enforced contact", i.e. distances ≤ 0, figure 1), is finally the offset curve that expresses the constraint on the objects. **The actual movement of the rolling object differs from this idealized trajectory due to inertia and elasticity. It is exactly the consequences of these physical properties, which are described by, and substantiate the use of, the impact model**.

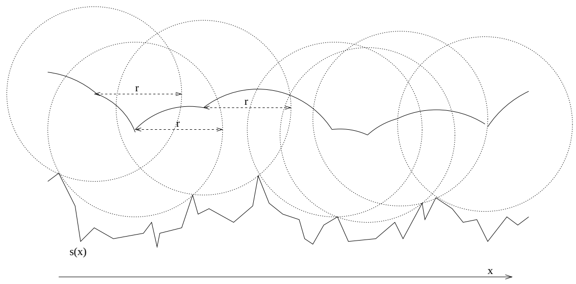

Figure 2: *Sketch of the effective offset-curve, resulting from the surface $s(x)$. The condition on the surface to be expressible as a function of one curve parameter x is clearly unproblematic in a "rolling" scenario.*

Implementation of the "rolling-filter"

In a straight approach, the calculation of contact points, necessary for the subsequent generation of the offset signal, is computationally highly demanding: in each point x along the surface curve, i.e. for each sample-point in a discrete implementation at audio rate, the following condition, which describes the momentary point of contact p_x, would need to be solved.

$$\begin{aligned} f_x(p_x) &\stackrel{!}{=} max_{q \in [x-r, x+r]} f_x(q) \qquad where \qquad (2) \\ f_x(q) &\triangleq s(q) + \sqrt{r^2 - (q-x)^2}, \quad q \in [x-r, x+r] \end{aligned}$$

The ideal curve would then be calculated from these contact points. E.g. for a diameter of $10cm$, a transversal velocity of $1m/s$ and a spatial resolution according to an audio sampling rate of $44100Hz$

at this tempo [2] the above operations, maximum/comparisons and calculus, had to deal with $44100 * 0.1m/1m = 4410$ values at each sampled position, i.e. 44100-times per second. Of course these computational costs are high in a real-time context for standard hardware, especially in our context of sound cartoons to be used within wider (also multi-modal) environments of human-computer interaction. The computations might be executed offline, which would however restrict the realtime reactivity of the model; object radius and surface structure had to be fixed and could not be easily changed dynamically.

The solution comes in form of a recursive algorithm that solves the described task with a highly reduced number of operations, to the order of 10 and therefore minimizes the computational load enabling realtime implementation. Computational costs are here comparable to that of a lowpass filter or other simple approximations that have been developed and tried by the author (figure 3 sketches an example). In fact, lowpass filtering appears to have been suggested and used to simulate the acoustic effect of rolling but sound results are often quite different. This is not surprising when remarking that the offset-curve as in figure 1 can contain strong high-frequency components (connected to its "edges"); such high frequencies may in some cases even be stronger than in the originating surface-profiles, contradicting the idea of lowpass filtering.

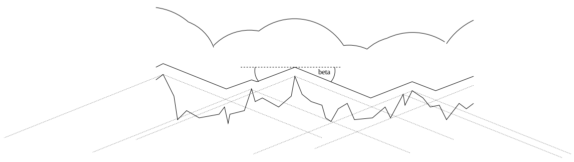

Figure 3: *A simple approximation of* rolling-filtering.

4. SURFACE PROFILE

Various origins may be thought of for the surface signal that is to be fed into the impact-model, at a rate related to the actual transversal velocity and after being processed by the *rolling-filter*. One possibility would be the scanning/sampling of real surfaces and use of such stored signals. From our standpoint of *cartoonification* and realtime interactivity we prefer the statistics-based generation of "surface-signals" of varying attributes. A commonly used model in computer graphics is fractal noise; in the 1-dimensional case this is noise of a $1/f^{\beta}$ power spectrum, where β reflects the fractal dimension or roughness. Typical surfaces of objects involved in rolling-interaction however, are usually smoothed and treated in various ways, which we reflect through band-limiting. In fact smoothing or polishing of surfaces may be seen as related to lowpass filtering, while global adjustment e.g. of tiles appears as a sort of highpass filtering. On that background, a global $1/f^{\beta}$ characteristic showed to be rather secondary in practical sound results. Therefore, white noise filtered with a bandpass of adjustable characteristics appears as an advantageous choice, combining efficiency and flexibility.

[2]…i.e., if we assume the surface profile to be resolved with a resolution such that when tracing the surface at the velocity of $1m/s$ samples appear at $44100Hz$, a canonical choice…

5. EXPLICIT MODELING OF MACROSCOPIC CHARACTERISTICS

Typical rolling-sounds usually show periodic patterns of timbre and volume that are of high perceptual importance. Periodicities that originate from macroscopic deviations of the rolling-shape from perfect sphericity — or more general, asymmetry of the object with respect to its center of mass — appear to form one important auditory cue for the recognition of rolling-sounds from similar sounds of contact, e.g. sliding. Also, the frequency of such periodic patterns strongly influences the perceived transversal velocity of the rolling object. Global asymmetries lead to modulations of the effective gravity force, that holds down the rolling object, an effect that gets stronger with increasing velocities (as motivated below). Usually less dominant is the simultaneous oscillation of the momentaneous velocity (of the point of contact along the plain). In our model, such effects have to be explicitely accounted for by according parameter modulations, since the physics-based core is one-dimensional and does not cover higher macroscopic geometries.

Figure 4 sketches an asymmetric rolling object in different positions. Its center of mass is accordingly at different heights giving different terms of potential energy. In a free rolling-movement

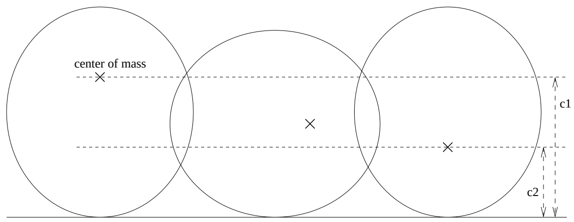

Figure 4: *Sketch of a rolling object at different instants, (strongly) asymmetric with respect to its center of mass.*

these oscillating terms of height of the center of mass $c(t)$ and potential energy are coupled to accordingly oscillating terms of kinetic energy and thus momentous velocity. This periodic energy transfer is connected to a periodic term of force acting between the rolling object and the plain (in addition to the constant gravity force). The exact terms of forces and velocities effective in this (free rolling-) situation could be found as solutions of the differential equation given by stating the principle of energy conservation; they can of course only be determined if the shape of the object is known exactly. However, in the context of effective *cartoonification*, we derive a simple example-approximation in the following, that reflects the general behavior. (With our goal in mind, 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? A general idea of "asymmetry" however may be given acoustically.)

We assume that the oscillating (in the sketch of figure 4 between the extrema of $c1$ and $c2$) height of the center of mass $c(t)$ is approximately described by a sinusoid [3] .

$$c(t) = c2 + (c1 - c2) \cdot \sin(\omega t), \omega = 2\pi \cdot f \qquad (3)$$

[3]This is e.g. the case for a spherical object rolling with constant angular velocity (which may in free rolling be approximately the case for small asymmetry or a forced condition) whose center of mass is located outside the geometrical center.

The offset force-term between the two contacting objects (the rolling and the plain) is then connected to the acceleration perpendicular to the plain through Newton's law $F(t) = M \cdot \ddot{c}(t)$, where M is the overall mass of the rolling object. The acceleration is the second derivation of equation 3.

$$\ddot{c}(t) = (c1 - c2) \cdot \omega^2 \cdot \sin(\omega t) \tag{4}$$

This sinusoidal force modulation term proportional to the square of the velocity in fact gives convincing sound results despite all involved approximations; a constant modulation amplitude sounds unnatural for changing velocity. In the model, a parameter of asymmetry, in these equations $c1 - c2$, allows to express an overall amount of deviation from perfect spherical symmetry. The modulation frequency f is related to the transversal velocity v and the (average) radius r of the rolling object, through $\omega = v/(2\pi \cdot r)$.

6. CONCLUSIONS

The development of and the ideas behind a dynamic real-time sound model of rolling have been explained. The model can be tuned in a wide range of ecological attributes; these include characteristics of material, size, surface structure, shape, velocity and direction. The implementation, realized as plugins and patches in sound software *pd* [4], runs comfortly on standard PC hardware under different operating systems. Since further on all parameters can be changed freely in real-time, the model is particularly useful for various, possibly multi-modal, environments of human-computer interaction. Applications with gestural or graphical control have been prototyped, others are in development.

7. ACKNOWLEDGMENTS

This work has been supported by the European Commission under contract IST-2000-25287 [5].

8. REFERENCES

[1] W. W. Gaver, *Everyday listening and auditory icons. PhD thesis*, University of California, San Diego, 1988.

[2] Federico Avanzini and Davide Rocchesso, "Impact," in *the Sounding Object*, Davide Rocchesso and Federico Fontana, Eds., pp. 125–129. Edizioni di Mondo Estremo, 2003.

[3] Matthias Rath, "Sound design around a real-time impact model," in *the Sounding Object*, Davide Rocchesso and Federico Fontana, Eds., pp. 149–161. Edizioni di Mondo Estremo, 2003.

[4] J. M. Adrien, "The Missing Link: Modal Synthesis," in *Representations of Musical Signals*, G. De Poli, A. Piccialli, and C. Roads, Eds., pp. 269–297. MIT Press, 1991.

[5] Federico Avanzini and Matthias Rath, "Modal resonators," in *the Sounding Object*, Davide Rocchesso and Federico Fontana, Eds., pp. 120–125. Edizioni di Mondo Estremo, 2003.

[6] Matthias Rath, "Modal synthesis," in *the Sounding Object*, Davide Rocchesso and Federico Fontana, Eds., pp. 252–255. Edizioni di Mondo Estremo, 2003.

[4] `http://iem.kug.ac.at/pd/`

[5] Project "SOb - the Sounding Object": `www.soundobject.org`)