Rolling Sound Synthesis : Work In Progress

Simon CONAN, Mitsuko ARAMAKI, Richard KRONLAND-MARTINET and
Sølvi YSTAD⋆

Laboratoire de Mécanique et d'Acoustique, MARSEILLE, FRANCE
{conan , aramaki , kronland , ystad}@lma.cnrs-mrs.fr

Abstract. This paper presents a physically informed rolling sound synthesis model for the *MétaSon* synthesis platform. The aim of this sound synthesis platform will be shortly described. As shown in the state of the art, both in terms of sound effects and proposed controls, existing models can be improved. Some details on asymmetric rolling objects will be given and the sound synthesis model will be exposed. Perspectives for further studies and work in progress will be discussed.

Keywords: Rolling Sounds, Sound Synthesis and Control, Sound Invariants, Physically Informed Synthesis, Rolling ball, Environmental Sounds Synthesis

1   Introduction

This study is part of a larger project (*MétaSon*) whose aim is to build a real-time sound synthesis platform that offers a perceptual control of sounds. To this purpose we need to :

- Build real-time synthesis models that reproduce sound features of real objects (plates, shell, water...) and interactions between them (rolling, rubbing...).
- Associate perceptual control strategies to these sound synthesis models control. By perceptual control, we mean that the synthesis model should be controllable by words which describe the sound as the user perceives it (e.g. "I want to produce a sound which rolls and with a liquid texture" or "I want a wind that sounds metallic") or by gestures.
- Identify sound *invariants* in order to achieve such perceptual controls. These *invariants* are either *structural invariants*, i.e. sound morphologies responsible for the recognition of an object and its properties, or *transformational invariants*, i.e. sound morphologies responsible for the recognition of the action on the object (see for example [1] for a study on breaking and bouncing invariants, and [2–4] for a more general approach). We are convinced that these *invariants* are strong enough to evoke both actions and objects and that it is possible to build and control sounds from perceptual categorizations in order to construct sound metaphors, like "bouncing water" or "rolling wind" for example.

In this paper, we propose to examine rolling sound synthesis and possible control strategies associated to such sounds. Different approaches to the synthesis of rolling sounds can be found in the literature.

One approach is the physical modeling of the phenomenon and the computation of equations with finite difference scheme. Stoelinga et al. derived a physical model that produce rolling sounds [5] from previous studies on impact sounds on damped plates [6, 7]. This model can reproduce effects like Doppler which is also found in the measures. However, sound examples aren't fully convincing, i.e. the sounds don't evoke rolling balls without ambiguity. This can be explained by the fact that no amplitude modulation is present in the case of continuous contact as the rolling object is considered as a perfect sphere (i.e. the mass center is the geometrical center). They also simulated very special cases of rolling like periodic bouncing and their simulations are comparable to their measures. It is important to note that these models cannot be computed in real time.

Another approach is the physically informed modelling. Here, the aim is to reproduce the "sound effect" produced by a rolling object. The sensation of a rolling object can then be modified by acting on specific acoustic features. These models are generally source-filter synthesis models, i.e. a source excitation which passes through a filter-bank (resonant object), informed by phenomenological considerations. The user can hereby control the synthesis parameters to act on size and speed of the rolling balls for example, but the control of these parameters can sometimes be difficult because the synthesis model has been constructed empirically. Van den Doel et al. [8] proposed a model where modal resonators were fed with a noise whose spectral envelope was defined by $\sqrt{1/(\omega - \rho)^2 + d^2}$ where ρ and d are respectively the frequency and the damping of the resonance, in order to enhance the resonance near the rolling object's mode.

In order to extract parameters from real recorded sounds for a sound synthesis model (i.e. parameters of a source-filter model), Lagrange et al. [9] and Lee et al. [10] proposed an analysis/synthesis scheme. The aim was to extract the excitation pattern and the object's resonances (the resonance of the rolling object and the surface on which it rolls were not separated). Depending on the recording that is analyzed, this can yield good resynthesis, but no general model of rolling objects with associated controls can be derived from such methods.

Both in terms of sound effects and proposed controls, the previously presented models can be improved. In fact, if we examine sound features that seem important for the perception of rolling sounds, we can conclude that more cues are necessary to perceive a wide variety of object sizes and rolling speeds. Houben et al. studied the auditory ability to distinguish the largest or the fastest ball between two recorded sounds. They also attempted to identify acoustic cues that characterize the size and speed of rolling balls, like auditory roughness or spectral structure that could be used to identify size and speed of rolling balls. They showed that at constant velocity (respectively at constant size) listeners can distinguish the largest (respectively the fastest) rolling ball with good results. Performance is impaired when the two factors (i.e. velocity and size) are crossed [11]. The influence of spectral and temporal properties was studied in [12] by

crossing the temporal content of a stimulus with the spectral content of another stimulus and using the obtained sound (the obtained stimulus has its spectrum very close to one stimulus and its temporal envelope very close to the other stimulus) in a perceptual experiment. It is shown that only the spectral structure is used to determine the fastest or largest ball and that results are better for the size judgement than for the speed judgement. However only recordings without clear amplitude modulation (due to an unbalanced ball or a deviation from perfect sphericity) were used in the experiment. This can explain why no temporal cues were found. The influence of this amplitude modulation is addressed in [13]. Artificial amplitude modulations were added to the recordings used in the previous experiments. Perceptual experiments showed that amplitude modulations clearly influence the perceived size and speed.

We would like to improve the synthesis of rolling by including the modulation effects. This problem was already addressed by some authors. In [14], Hermes proposed a synthesis model. It consisted of a series of impacts following a Poisson law amplitude modulated to account for the asymmetry of the ball. This pattern was further convolved with a sum of gamma-tones to represent the impulse response of the object on which the ball rolled. He justified this form of impulse response in comparison to the classical representation that uses a sum of exponentially decaying sinusoids by the fact that the collisions between the ball and the plate are "softer". The control of this model is quite complicated. In [15], Rath described a physically informed rolling sound synthesis model. Impacts on the surface were modeled by modal resonators. These resonators were fed with a low-pass filtered noise (which represents the surface profile) which was further filtered by the "rolling filter" to simulate the irregularities encountered by the ball. The force was modulated by a sinusoid to account for asymmetry of the rolling ball and the whole model could be run in real-time. Nevertheless, the modulation force was only derived for a constant velocity.

In the rest of this paper, we will describe the rolling sound synthesis model we developed. First, a simplified model of amplitude modulation due to an unbalanced rolling ball will be exposed and then the global synthesis model will be presented. In the last section, the work in progress and the perspectives will be presented.

2   Asymmetric rolling object

The aim of this section is to get an idea of the modulation profile generated by an imperfect rolling ball (i.e. a ball for which the mass center differs from the geometrical center). Our model is very simple and does not reflect the real motion of the ball because we impose a given velocity profile of the ball's geometrical center to get the modulation profile (in fact, the mass center's eccentricity impose a velocity to the geometrical center). We use this approximation because the kinematics of an unbalanced ball is not a trivial problem and we cannot easily derive solutions from studies on unbalanced rolling objects (see the studies on loaded hoops in [16] for example).
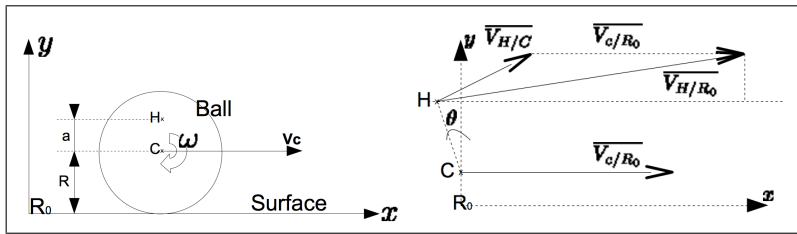
Fig. 1. Left : Representation of a ball of radius R rolling with a transversal velocity V_c. Its mass center H is off centered by a distance a from its geometrical center C. Right : Representation of the velocities.

So let us consider a ball with C its geometrical center and R its radius that rolls with a transversal velocity V_c (angular velocity ω) in the reference frame R_0. The mass center of this ball is situated at the point H, at a distance a from C (see figure 1). Assuming a pure rolling motion, the movement of the point H toward y is :

$$h_y(x) = R + a \sin\left(2\pi \frac{x}{2\pi R}\right) = R + a \sin\left(\frac{x}{R}\right) \tag{1}$$

The amplitude modulation is due to the height variation of the mass center (which is related to potential energy) with respect to time :

$$h_y(t) = R + a \sin\left(\frac{x(V_H(t))}{R}\right) \tag{2}$$

with $V_H(t)$ the velocity of the point H within the frame of reference R_0. The component of $\overline{V_{H/R_0}}$ on the $\overline{x}$ axis (see figure 1 for notations) is (assuming the ball is rolling without sliding : $\|\overline{V_{c/R_0}}\| = R\omega$ and $\|\overline{V_{H/c}}\| = a\omega$) :

$$\overline{V_{H/R_0}} \cdot \overline{x} = \|\overline{V_{c/R_0}}\|(1 + \alpha \cos(\theta)) \tag{3}$$

with $\alpha = a/R$. For more clarity, we will write $\overline{V_{H/R_0}} \cdot \overline{x}$ and $\|\overline{V_{c/R_0}}\|$ respectively V_H et V_c. And θ follows :

$$\theta(t) = \int \omega(t)dt = \int \frac{V_c(t)}{R} dt \tag{4}$$

We can see the modulations for two different asymmetries in figure 2.

3 Sound synthesis model

A general framework of the model is shown in figure 3 with its associated controls. The sound synthesis model is based on a source-filter architecture : a noise (the source) is sculpted by successive filters, then is modulated in amplitude and finally feeds a bank of resonant filters that describe the surface on which the object rolls. The control parameters are based both on perceptual attributes
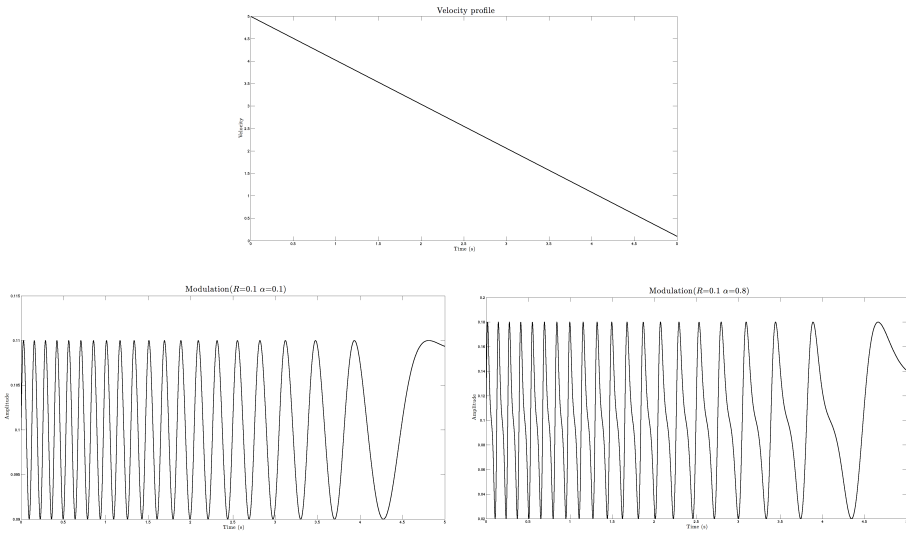
Fig. 2. Velocity profile and associated modulation for two different asymmetries.
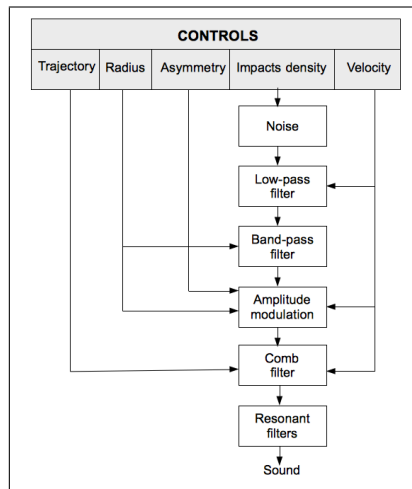
Fig. 3. General Framework of the synthesis model to produce rolling sounds.

and on physical considerations. The links between these controls and the low-level synthesis parameters are described below.

In order to simulate series of collisions between the ball and the surface (i.e. the relative surface which is "seen" by the ball as it is rolling), we use a noise. This signal consists in a sequence of impacts of different amplitudes which are more or less spaced in time. The temporal pattern is modelled as a random process : at each sample, a Bernoulli process is performed with a probability ρ. We choose $\rho \in [0.01, 0.03]$ (a value of ρ that is too high leads to a sound that is too noisy, and a value below 0.01 leads to a sound that is too discontinuous). The amplitude of each impulse is random, and follows an uniform law between 0 and 1. We already used this noise to simulate the sound produced by two continuous interactions, rubbing and scratching [17]. This study focuses on the perceptual differences between rubbing and scratching actions evoked by recorded and synthesized sounds, and shows that a density of $\rho < 0.01$ is associated to scratching and $\rho > 0.1$ is associated to rubbing and that the perception is ambiguous between these two values. This kind of noise is similar to the one used in [14].

This noise is then low-pass filtered. The cut-off frequency is related to the transversal V_c velocity of the ball, i.e. the faster the ball the higher the cut-off frequency (see [8] for further information). To account for the size of the ball, we use a band-pass filter with center frequency $f_c \propto (1/R)$ with R the radius of the ball. The assumption that motivates this filtering stage is that the plate is more excited near the modes of the rolling object.

The amplitude modulation is then applied to the filtered noise. We get the resulted noise s by computing $s^{n+1} = e^{n+1}.h_y^{n+1}$ with e the excitation noise previously described and h_y the height variation of the mass center computed as :

$$h_y^{n+1} = R(1 + \alpha \sin\left(\frac{x^{n+1}}{R}\right)) \quad \text{with} \quad \begin{cases} x^{n+1} = x^n + V_H^{n+1} dt \\ V_H^{n+1} = V_c^{n+1}(1 + \alpha \cos(\theta^{n+1})) \\ \theta^{n+1} = \theta^n + \frac{V_c^{n+1}}{R} dt \end{cases} \quad (5)$$

To account for the position-dependent excitation, we use a comb filter (see for example the explanation of Smith on the position dependent excitation on a guitar string [18]). This filtering stage is important as it gives the listener a sensation of displacement.

Finally, the obtained excitation is used to feed a resonant filter bank used to simulate the surface on which the ball rolls. These filters model the impulse response of the surface, which is related to the *structural invariant* responsible for the recognition of the surface.

4   Perspectives

This model yields convincing results but the mapping strategy needs to be improved. In fact, the height variation of the center of mass needs to be linked to a

force applied to the surface on which the object rolls. We could derive a force as Rath in [15] by applying Newton's law to the height variation of the mass center $F(t) = M.\ddot{h}_y(t)$, with M the rolling object mass and $\ddot{h}_y(t)$ the acceleration perpendicular to the plain of the mass center. However, the mapping between the obtained force and the synthesis model is not direct and we are currently investigating a mapping between the force and the rolling sounds signal model. The force should be integrated at a lower level in the model, i.e. directly in the random process of micro-impacts generation. We are also working on defining a physical model of rolling objects. Our aim is to try to recover parameters of the model from real recordings thanks to inverse problem methods. Such an approach will validate our force model or show us if further refinements (e.g., reduction of simplifying assumptions) are necessary.

Besides the construction of a synthesis model that simulates realistic rolling sounds, calibrated sounds from this model can also be used to identify perceptual cues responsible for the recognition of this specific action. Furthermore, this approach aimed at highlighting the acoustical cues, also called *invariants*, which characterized the rolling action. Perceptual tests should further validate these assumptions before they are tested as descriptors on real sounds.

Another aspect that would add realism to the rolling sound synthesis is the simulation of the ball's position on the surface. Taking into account the multiple reflections from the edges of the surface on which the object rolls by adding several comb filters for which the delays are computed by a source-image method may increase realism. In [19], Stoelinga et al. analysed the wave dispersion (i.e. the frequency dependent wave velocity) in a plate and concluded that frequency dependent comb filters add more realism when simulating a ball approaching the edge of a plate. Finally, real time implementation should conclude this work.

References

1. Warren, W.H. and Verbrugge, R.R.: Auditory Perception of Breaking and Bouncing Events: A Case Study in Ecological Acoustics. Journal of Experimental Psychology: Human Perception and Performance 10, 5, 704–712 (1984)
2. McAdams, S.E. and Bigand, E.E.: Thinking in Sound: The Cognitive Psychology of Human Audition. Oxford Science Publication. Chapter 6 (1993)
3. Gaver, W.W.: What in the World Do We Hear?: An Ecological Approach to Auditory Event Perception. Ecological psychology, 5, 1, 1–29 (1993)
4. Michaels, C.F. and Carello, C.: Direct Perception. Prentice-Hall Englewood Cliffs, NJ (1981)
5. Stoelinga, C., Chaigne, A.: Time-Domain Modeling and Simulation of Rolling Objects. Acta Acustica united with Acustica 93, 2, 290–304 (2007)
6. Chaigne, A., Lambourg, C.: Time-Domain Simulation of Damped Impacted Plates. I. Theory and Experiments. Journal of the Acoustical Society of America 109 (2001)
7. Lambourg, C., Chaigne, A., Matignon, D.: Time-Domain Simulation of Damped Impacted Plates. II. Numerical Model and Results. Journal of the Acoustical Society of America 109 (2001)
8. Van Den Doel, K., Kry, P.G., Pai, D.K.: FoleyAutomatic: Physically-Based Sound Effects for Interactive Simulation and Animation. In: Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques, 537–544 (2001)

9. Lagrange, M., Scavone, G., Depalle, P.: Analysis/Synthesis of Sounds Generated by Sustained Contact between Rigid Objects. IEEE Transactions on Audio, Speech, and Language Processing, 18, 3 509–518 (2010)

10. Lee, J.S., Depalle, P., Scavone, G.: Analysis/Synthesis of Rolling Sounds Using a Source-Filter Approach. In: 13th Int. Conference on Digital Audio Effects (DAFx-10), Graz, Austria (2010)

11. Houben, M.M.J., Kohlrausch, A., Hermes, D.J.: Perception of the Size and Speed of Rolling Balls by Sound. Speech Communication 43, 4, 331–345 (2004)

12. Houben, M.M.J., Kohlrausch, A., Hermes, D.J.: The Contribution of Spectral and Temporal Information to the Auditory Perception of the Size and Speed of Rolling Balls. Acta Acustica united with Acustica 91, 6, 1007–1015 (2005)

13. Houben, M.: The Sound of Rolling Objects, Perception of Size and Speed. Technische Universiteit, Eindhoven (2002)

14. Hermes, D.J.: Synthesis of the Sounds Produced by Rolling Balls. In: Internal IPO report no. 1226, IPO, Center for User-System Interaction, Eindhoven, The Netherlands (1998)

15. Rath, M.: An Expressive Real-Time Sound Model of Rolling. In: Proceedings of the 6th "International Conference on Digital Audio Effects" (DAFx-03). Citeseer (2003)

16. Theron, W.F.D.: Analysis of the Rolling Motion of Loaded Hoops (2008)

17. Conan S., Aramaki M., Kronland-Martinet R., Thoret E. and Ystad S.: Perceptual Differences Between Sounds Produced by Different Continuous Interactions. Acoustics 2012, Nantes (2012).

18. Smith, J.O.: website, `https://ccrma.stanford.edu/realsimple/faust_strings/Pick_Position_Comb_Filter.html`

19. Stoelinga, CNJ and Hermes, DJ and Hirschberg, A. and Houtsma, AJM.: Temporal Aspects of Rolling Sounds: A Smooth Ball Approaching the Edge of a Plate. Acta Acustica united with Acustica, 89, 5, 809–817 (2003)