

BIOACOUSTIC MODELING FOR SOUND SYNTHESIS: A CASE STUDY OF ODONTOCETI CLICKS

Cumhur Erkut

Helsinki University of Technology
Lab. of Acoustics and Audio Signal Processing
P.O. Box 3000, FIN-02015 HUT, Espoo, Finland
Cumhur.Erkut@hut.fi

ABSTRACT

This paper documents the design of a computer instrument that is based on a simplified bioacoustic model of odontoceti clicks. The components of the basic model are described, and its implementation is carried out by using simple filters. By imposing restrictions, this model has been converted to a real-time computer instrument. Related material and sound examples are available online at <http://www.acoustics.hut.fi/~cerkut/catacea/>.

1. INTRODUCTION

Model-based sound synthesis provides tools for simulation of the sound production mechanisms in musical instruments [1, 2]. More recently, the same tools are being used for simulation of bioacoustical systems [3]. The influence of the anatomy on the sound production mechanism, and in turn, on the emitted sounds is an important topic in the study of sound in non-human animals [4, 5]. Although the musical potential of animal sounds has been recognized for a long time, their musical usage has been limited.

This paper documents the design of a virtual instrument, which is based on a simplified sound-source model derived from the acoustics of odontoceti (toothed whales or dolphins) clicks. The design was motivated by a computer music composition.¹ The synthesis model runs in real-time in CsoundAV [6], and can be controlled via a GUI.

The organization of the paper is as follows. Sec. 2 provides the background on the acoustics of catacean clicks. In Sec. 3, a simple three-component model of click production mechanism is reviewed. This simplified model is originally proposed by Aroyan [7], and it provides the basis of the synthesis model introduced in this paper. Section 3.4 documents the extrapolation strategies to extend the model to different odontoceti, and Sec. 4 concentrates on the real-time implementation of the instrument. Finally, Sec. 5 draws the conclusions and indicates the future directions.

2. ACOUSTICS OF ODONTOCETI CLICKS

Odontoceti can detect small targets at ranges from several meters for smaller porpoises to over a hundred meter for larger dolphins. An increasing number of studies are being carried out to understand the principles of the biosonar system enabling such an ability, see [5] for a review. In general, echolocation signals of different species have common features, but some characteristics of these signals differ depending whether or not the species produce whistles. The whistles are continuous narrowband FM signals that usually accompany the echolocation clicks. Recent theories suggest that the whistles and echolocation pulses are produced by the same mechanism [5], however in this paper we concentrate on the echolocation pulses and leave the whistles aside. The common features

¹PCM 0355+53 by Shinji Kanki for the Helsinki Computer Orchestra.

of the pulses of different species are their short duration, high mean-frequency, and high intensity (150-227 dB). Whistling odontoceti produce brief (40-70 μ s, typically 4-8 cycles) broadband signals. The mean frequency of these signals are intensity-dependent. For high-intensity signals, the main frequency is around 100 kHz. Most dolphins belong to this class, the bottlenose dolphin (*tursiops truncatus*) being a typical example. In addition, there is evidence that the riverine dolphins also produce whistles [8]. Their echolocation pulses exhibit similar characteristics to that of the whistling dolphins. Smaller dolphins and porpoises that do not whistle emit narrowband echolocation signals with at least 12 cycles and a duration greater than 100 μ s. The harbor, finless, and Dall's porpoises, the Commerson's and Hector's dolphins, as well as the pygmy sperm whale belong to this group.

Although once believed that the source of the echolocation pulses were the larynx as in the other mammals, more recent studies point out a flow of air that is pressurized within the bony nares and passed upward into the nasal sac system during production of echolocation clicks and whistles [9]. A particular biological structure within the nasal system just inside the blowhole, the *monkey lips - dorsal bursae* (MLDB) complex has been widely accepted as the source of the sonar clicks [10, 9]. This complex includes a small pair of fatty bursae embedded in a pair of connective tissue lips, a cartilaginous blade, a stout ligament, and an array of soft tissue air sacs.

A physically plausible source model is based on the pneumatically driven self-oscillations of the monkey lips [11, 9], which in turn cause the bursae to accelerate and radiate. However, the radiation due the surface accelerations of the bursae is neither efficient nor directional enough to form a biosonar signal. The melon, which is a fat-filled cavity in the dolphin forehead, focuses and transmits the sound generated within the forehead tissues and acoustically couples it to the surrounding water. The melon has long been regarded as the primary acoustic lens, however using numerical simulations based on finite differences, Aroyan et al. concluded that the the geometry of the skull and air sacs play the central role in beam formation, hence they restricted the function of the melon to a mild focus [12].

More recently, based on a simplification of his previous works, Aroyan suggested that, if the dolphin's click generator can produce impulses of progressively higher frequency as the driving air pressure (and muscular tension) is increased, and if this source is located in soft tissues of the nasal passageways that are partially surrounded by air sacs, then the biosonar characteristics may simply result from a variable source being filtered by the resonances of a soft tissue resonator and a projector [7]. Aroyan's simple model of the dolphin echolocation emission system is such a three-component model that will be further discussed in the next section.

3. A SIMPLE SOUND SOURCE MODEL

Aroyan's simple model consists of the following three components: a source mechanism of broadband pulses, a damped resonator, and a projector of the signal produced by the first two components. The source model assumes the region of the nasal passageways as the origin of echolocation pulses. The exact physical mechanism for click production is not explicitly modeled, rather a pulse train is considered as a source signal. A Gaussian-windowed sinusoid represents a close match to the source characteristics, as will be discussed in Sec. 3.1.

A resonant component is needed because of the the reverberant characteristics of echolocation signals to broadband pulse inputs. Such characteristics are the result of multiple reflective structures that partially surround the soft tissues within the dolphin's nasal passages. In addition, the natural resonances of the air sacs themselves exhibit low-frequency peaks around 3-4 kHz in typical delphinid echolocation spectra. These low-frequency resonances are excluded in the Aroyan model, since they fall far below the predominant spectral frequencies. However, they are within the human hearing range and are important for sound synthesis purposes. Modeling of these resonances will be discussed in Sec. 3.2, and modeling of the damped resonator component will be discussed in Sec. 3.3. The signal produced by a nasal click source (and resonator) must pass through the dolphin's forehead tissues before entering seawater. The forward-reflective properties of the skull and nasal sacs and the refractive properties of the melon and other forehead soft tissues combine to produce additional filtering of the signal. For this reason, a projective element is considered to be the final component of the dolphin emission model, as will be discussed in Sec. 3.3.

The following subsections are devoted to implementation of these components in MATLAB environment. All the parameters in the implementation are tuned roughly for matching the bottlenose dolphin's (*tursiops truncatus*) click characteristics.

3.1. Broadband source signal

The source is modeled as a Gaussian-modulated sinusoidal signal. The mean frequency of the Gaussian pulse f_0 is set to 105 kHz, and the bandwidth is set to $\pm 0.75f_0$. The signal is truncated at -10 dB. These values reproduce the typical spectratemporal characteristics of the high-intensity click sources of the bottlenose dolphin, as discussed previously in Sec. 2. A synthetic example is illustrated in Fig. 1.

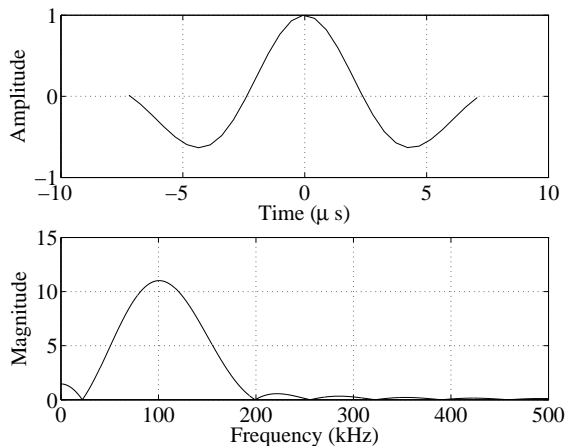


Figure 1: *The spectral and temporal properties of synthesized source signals.*

3.2. Air-sac resonances

Although not included in the Aroyan model, the air-sac resonances fall within the human hearing range and thus are important for the recognition of the clicks by humans. The sac resonances are roughly modeled by a Helmholtz resonator. The frequency of an Helmholtz resonator is given by

$$f_0 = \frac{c}{2\pi} \sqrt{\frac{A}{Vl_o}} \quad (1)$$

where A is the cross-sectional area and V is the volume of the air contained in the cavity, and l_o is the vertical offset. Note that substituting this variables, one can conclude that $f_0 \propto 1/L$, i.e., the air-sac resonance is inversely proportional to the length of the dolphin. Furthermore, since both the f_0 and L are known for the bottlenose dolphin, the proportionality constant can also be calculated. This proportionality constant can be used for scaling the model parameters given the length of the odotocete. Although useful for modeling purposes, it should be kept in mind that the actual geometry and acoustical properties of air sacs are far more complex than a simple Helmholtz resonator. Within these limitations, the basic resonator is realized with a second-order bandpass Butterworth filter. Only the resonance frequency is matched with the data, the Q-factor and relative spectral height are not calibrated.

3.3. Resonator and projector components

The damped resonator component models the partially air-bounded spherical enclosure that has monopole and dipole radial radiation modes. The monopole modes are given by [7]

$$f_{0,n} = n \frac{c}{2r} \quad (2)$$

where r is the radius of the air-bounded enclosure and is about 1.25 cm for the bottlenose dolphin. Note that the monopole modes are in harmonic ratio. The dipole radial modes do not have harmonic ratio. Since the enclosure is excited by a monopole source from its center, the monopole modes would be expected to dominate so that the resonance structure of such an enclosed source should be predominantly harmonic. Therefore the model includes only the monopole modes. These monopole components are realized with a FIR comb filter. A simple FIR filter is also included to simulate losses due to leakage from the partially air-bounded enclosure. The transfer function of the resulting filter is given by

$$H_r = \frac{1 - g_r}{2} \frac{1 + z^{-1}}{1 + z^{-L_r+1} - g_r z^{-L_r}}, \quad (3)$$

where (2) is used to calculate the comb length L for the given sampling frequency f_s

$$L_r = \frac{f_s}{f_{0,1}}. \quad (4)$$

The projector component models the bony structure as a parabolic reflector with source input at focal point. The delay of the reflected components can be calculated by [7]

$$\Delta t = 2 \frac{a}{c} \quad (5)$$

where a is the parabola focal distance. For the bottlenose dolphin, this distance is $a \approx 1.3$ cm. The multiple reflections will occur at the integer multiples of this basic time-offset. The refracted components, on the other hand, have a basic high-pass character. These two components, in the main axis of symmetry, may be realized with the following lossy FIR high-pass comb filter

$$H_p = \frac{1 - g_p}{2} \frac{1 - z^{-1}}{1 + z^{-L_p+1} - g_p z^{-L_p}}, \quad (6)$$

where L_p can be derived from (5).

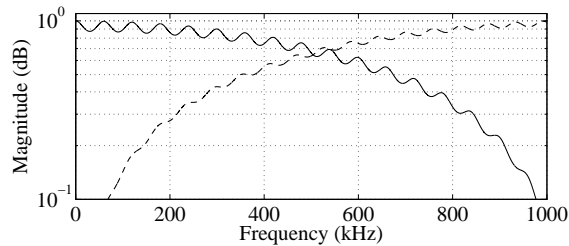


Figure 2: The magnitude responses of the reflector (solid) given by (3) and the projector (dashed) given by (6).

The magnitude responses of the reflector and projector components are illustrated in Fig. 2 for the reflector gain $g_r = 0.067$ and the projector gain $g_p = 0.033$. Since the sampling frequency is very high, many peaks can be seen in the figure.

The model is constructed by cascading the damped resonator and projector components and adding the air-sac resonator in parallel. The final synthetic click signal and its magnitude spectrum are shown in Fig. 3.

3.4. Different Odontocete Objects

After testing the match of the available specifications of the bottlenose dolphin with the basic model, the model parameters are extrapolated to other species. The air-sac extrapolation is carried out by using the proportionality factor obtained from (1), as discussed in Sec. 3.2. A similar proportionality is obtained in the main frequency of

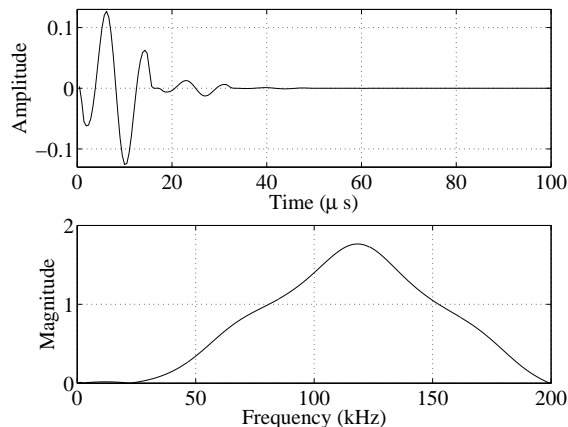


Figure 3: *The synthetic click signal (top) and its magnitude response (bottom).*

| Odontoceti | Length | AirSac | Click | Width | Monopole | Projector |
|--------------------|---------------|---------------|--------------|----------------|-----------------|------------------|
| Delphinus delphis | 2 m | 5.25 kHz | 157.5 kHz | 40.07 μ s | 90 kHz | 86.54 kHz |
| Tursiops truncatus | 3 m | 3.5 kHz | 105 kHz | 60.21 μ s | 60 kHz | 57.69 kHz |
| Orcanis orca | 8 m | 1.31 kHz | 39.38 kHz | 160.20 μ s | 22.5 kHz | 21.64 kHz |

Table 1: *Parameters for selected odontocete objects.*

the clicks. With the mean frequency at hand, the click width is obtained by calculating the effective time support of the Gaussian pulse that is attenuated by -60 dB. The monopole harmonics are calculated by (2), where the radius is assumed to be proportional to the length. Similarly, the focal distance of the parabola is assumed to be proportional to the length in (5). The values obtained with these assumptions are tabulated in Table 1.

4. REAL-TIME IMPLEMENTATION OF THE SSM

An investigation of Table 1 reveals that the most interesting properties of the simplified model occurs in the ultrasonic frequencies. However, most of the recorded clicks are sampled by $f_s = 44100$ Hz. Therefore, in the real-time implementation, the monopole and projector components of the model have been discarded except the killer whale (orcanis orca), but the other parameters are kept in their original values. The models are implemented in CsoundAV [6]. For each object, the source signal is obtained by the `grain` opcode that creates a Gaussian-modulated sinusoidal source signal of main frequency `krfreq`, in accordance with Table 1. The implementation instantiates 3 similar objects with the parameters listed in Table 1. A GUI consisting of simple widgets has been implemented to assist to real-time control of the model.

An option in real-time synthesis would be to map the ultrasonic range to the audible range. A linear mapping preserves the quasi-harmonic structure and beating of the synthetic clicks, but results in very low-frequency pulses. Alternatively, frequency-warping could be used [13]. This mapping alters the sound characteristics and produces inharmonic clicks. Further experimentation on these sonic possibilities are left out as future work.

5. CONCLUSION AND FUTURE WORK

This paper has documented the design of a computer instrument that is based on a simplified bioacoustic model of odontoceti clicks. The components of the basic model are described, and its implementation is carried out by using simple filters. By imposing restrictions, this model has been converted to a real-time computer instrument,

which can be controlled with ease during a performance. Related material and sound examples are available online at <http://www.acoustics.hut.fi/~cerkut/catacea/>.

There are many possible directions for the future work. The source signal component can be replaced by the pneumatic-mechanic source model described in [11]. More realistic modeling of the air-sac resonances would certainly improve the simulation. In particular, the detailed geometry of the non-whistling odontoceti sacs needs to be incorporated into the basic model. In the harbor porpoise, for example, the vestibular sacs have prominent connective tissue ridges along their floors, and the relative size of these sacs is larger than in most delphinids [7]. The Commerson's dolphin and the franciscana also exhibit relatively large vestibular sacs.

Modeling of melon as well as the directional components are further future directions. Finally, the comparison of the model output with recorded click patterns could help to fine-tune the basic model parameters.

6. REFERENCES

- [1] J. O. Smith, "Principles of digital waveguide models of musical instruments," in *Applications of Digital Signal Processing to Audio and Acoustics* (M. Kahrs and K. Brandenburg, eds.), pp. 417–466, Boston, Massachusetts, USA: Kluwer Academic Publishers, 1998.
- [2] P. R. Cook, *Real sound synthesis for interactive applications*. Natick, MA, USA: A. K. Peters, 2002.
- [3] T. Smyth and J. O. Smith, "Applications of bioacoustics in physical modeling and the creation of new musical instruments," in *Proc. Int. Symposium on Musical Acoustics (ISMA)*, (Perugia, Italy), Sept. 2001.
- [4] N. H. Fletcher, *Acoustic systems in biology*. New York, USA: Oxford University Press, 1992.
- [5] J. A. Thomas, C. F. Moss, and M. Vater, eds., *Echolocation in bats and dolphins*. London, UK: The University of Chicago Press, 2004.
- [6] G. Maldonado. CsoundAV is freely available at <http://www.csounds.com/csoundav/index.html>.
- [7] J. L. Aroyan, "Simple models of the dolphin echolocation emission system," 2002. Available online at <http://members.cruzio.com/~jaroyan/research.html>.
- [8] W. Ding, B. Würsig, and S. Leatherwood, "Whistles of boto, inia geoffrensis, and tucuxi, sotalia fluviatilis," *J. Acoust. Soc. Am.*, vol. 109, pp. 407–411, Jan. 2001.
- [9] T. W. Cranford and M. Amundin, "Biosonar pulse production in odontocetes: The state of our knowledge," in *Echolocation in bats and dolphins* (J. A. Thomas, C. F. Moss, and M. Vater, eds.), ch. 4, pp. 27–35, London, UK: The University of Chicago Press, 2004.
- [10] T. W. Cranford, M. Amundin, and K. S. Norris, "Functional morphology and homology in the odontocete nasal complex: implications for sound generation," *Journal of Morphology*, vol. 228, no. 3, pp. 223–285, 1996.
- [11] N. A. Dubrovskiy and L. R. Giro, "Modeling of the click-production mechanism in the dolphin," in *Echolocation in bats and dolphins* (J. A. Thomas, C. F. Moss, and M. Vater, eds.), ch. 10, pp. 59–63, London, UK: The University of Chicago Press, 2004.
- [12] J. L. Aroyan, T. W. Cranford, J. Kent, and K. S. Norris, "Computer modeling of acoustic beam formation in *Delphinus delphis*," *J. Acoust. Soc. Am.*, vol. 92, pp. 2539–2545, Nov. 1992.
- [13] A. Härmä, M. Karjalainen, L. Savioja, V. Välimäki, U. K. Laine, and J. Huopaniemi, "Frequency-warped signal processing for audio applications," *J. Audio Eng. Soc.*, vol. 48, pp. 1011–1031, Nov. 2000.