Proceedings of the 6th Nordic Signal Processing Symposium - NORSIG 2004
June 9 - 11, 2004
Espoo, Finland

# Modal Synthesis of Wind Chime Sounds with Stochastic Event Triggering*

*Teemu Lukkari*[1] *and Vesa Välimäki*[2]

[1]Helsinki University of Technology
Institute of Mathematics
P.O. Box 1100, FIN-02015 HUT
Espoo, FINLAND
E-mail:Teemu.Lukkari@hut.fi

[2]Helsinki University of Technology
Laboratory of Acoustics and Audio Signal Processing
P.O. Box 3000, FIN-02015 HUT
Espoo, FINLAND
E-mail:Vesa.Valimaki@hut.fi

### ABSTRACT

*A parametric synthesis model for wind chimes is presented. A physics-based stochastic process is used to trigger a modal sound-production model for the tubes of a wind chime. Mode parameters are extracted in a semiautomatic fashion from a sound sample. An example and some possible applications are presented.*

## 1. INTRODUCTION

There is a need for modeling and synthesizing various everyday sounds. They offer a rich and intuitive alternative to musical sounds in many applications such as auditive displays [1], computer games, virtual reality systems and sonification [2]. In these applications, there are benefits to using parametric synthesis methods [3]. Parametric methods are flexible and economic because of the capability to synthezise many sounds with a single algorithm.

We present a parametric synthesis model for a wind chime. The model includes a wind force parameter that can be used to give the listener an idea of how strong a wind is blowing. This can be used for example as an auditory warning.

For our purposes, the most important parts of a wind chime are the tubes, the clapper and the windcatcher. The tubes hang in a circle. The clapper and the windcatcher are hung up in the middle of the circle. Whenever there is a wind blowing, the clapper starts swinging because of the windcatcher. When one of the tubes is hit by the clapper, a sound is produced.

Thus the task of generating wind chime sound is split into two processes. First, a way of determining when one of the tubes is hit and which tube is hit is needed. Second, the vibration of the tubes which is the actual sound production mechanism needs to be modeled.

*This work is part of the CLAPS project (Academy of Finland project #104934).

## 2. STOCHASTIC TRIGGERING

We model the motion of the clapper by using a loosely physics based stochastic process (Fig. 1). The process has states that correspond to the tubes in a wind chime, and a zero state that corresponds to the clapper hanging in the middle of the tubes.

When the process goes to a nonzero state, the clapper has hit the tube corresponding to this state. When this occurs, we trigger the sound synthesis of that tube. The state transition probabilities change depending on the energy of the clapper. The energy in turn is influenced by a wind force parameter.

### 2.1 Energy Update Rule

One can think of the motion of the clapper as the motion of a pendulum. The energy of the clapper can be changed by various causes, for example by a wind blowing or giving the clapper a push. The energy of a pendulum decays exponentially.

By the discussion in the previous paragraph, we can write an update rule for system energy as

$$E_n = R_{\mathrm{decay}}(E_{n-1} + d_n), \tag{1}$$

where $E_n$ is the system energy at time index $n$, $d_n$ is the amount of energy added between time indices $n-1$ and $n$ and $R_{\mathrm{decay}} < 1$ is the per-sample energy decay rate. For a sampling frequency of $f_s = 11025$ Hz we used the value $R_{\mathrm{decay}} = 0.9999$. If we assume that wind force is modeled as a continuous time function $f$ giving the time derivative of the system energy, $d_n$ can be approximated as

$$d_n = T_s f((n-1)T_s), \tag{2}$$

where $T_s$ is the the sampling interval.

### 2.2 Generating the State Transitions

State transitions are checked for at random intervals with length of the intervals having a uniform distri-

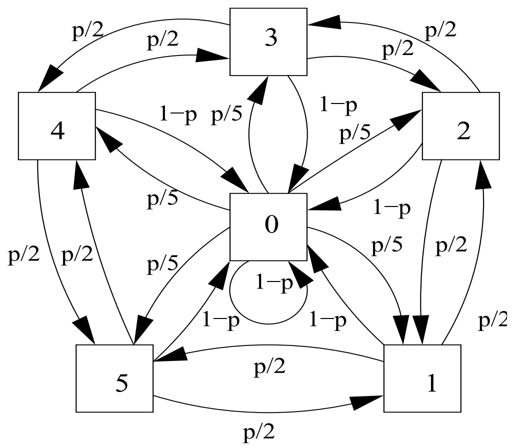


**Fig. 1.** Stochastic process for generating trigger events.

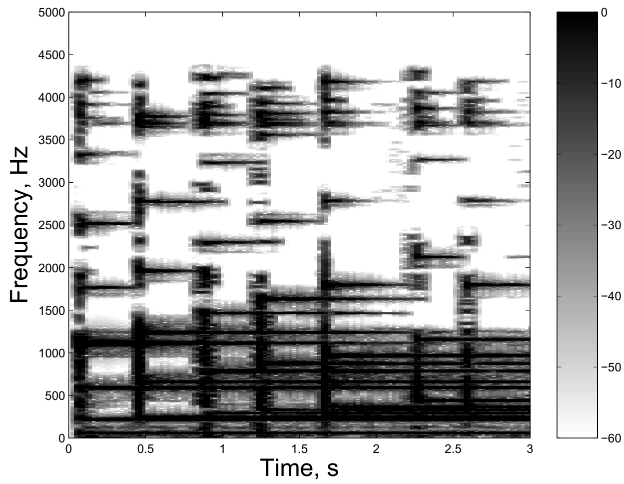


**Fig. 2.** Spectrogram of the first three seconds of the sound sample.

bution between 0.03 and 0.05 seconds. This is needed because for constant intervals the sound would be unnatural in the situation where a hit nearly always occurs.

The motion of the clapper is modeled with a stochastic process. First, the probability of a hit occuring is computed. It should be close to 1 when the clapper has lots of energy, and close to 0 when the clapper has small energy. A suitable way of achieving this is to use a sigmoid function to compute the probability $p$ from the energy $E_n$,

$$p(E_n) = \frac{1}{1 + c\exp(-2E_n)}, \tag{3}$$

where $n$ is the time index of the state transition check. Here $c$ is a constant used to give a convenient small value to $p(0)$. We used a value of $c = 99$.

Once the probability of a hit occuring is available, it should be divided between the states of the process in an appropriate manner. A fairly simple way of doing this is the following, illustrated in Fig.1.

1. If the current state is zero, the probability is divided evenly between the nonzero states.
2. If the current state is nonzero, the probability is divided between the two nonzero states that are adjacent to the current state.
3. If a hit does not occur, the process goes to the zero state.

Returning to the pendulum analogy, one can think of the zero state as a potential minimum. The fact that energy is needed to move the clapper up to the tubes is reflected in the transition probabilities.

The process generating the hits is related to a Markov chain [4]. It has the property that the next state depends only on the current state. The transition probabilities can and will vary from step to step, so the process lacks the time independence property.

## 3. SIGNAL ANALYSIS

For modeling the sound generation mechanism we chose to express the sound as a sum of exponentially decaying sinusoids corresponding to the modes of the tubes. The modes can be represented by second-order digital resonators. The transfer function of a resonator is given by [5]

$$H(z) = \frac{GA_0(1 - z^{-2})}{1 - 2R\cos(\theta)z^{-1} + R^2z^{-2}}, \tag{4}$$

where $\theta$ is related to the resonance angular frequency $\omega_0$ and the pole radius $R$ by

$$\cos(\theta) = \frac{2R}{1 + R^2}\cos(\omega_0). \tag{5}$$

$A_0$ is a normalization factor used to get maximum gain of 0 dB, and is given by

$$A_0 = (1 + R^2)\sin(\theta). \tag{6}$$

$G$ is a gain factor used to control the levels of the modes relative to one another. The parameters $G$, $\omega_0$ and $R$ need to be estimated from sound samples.

### 3.1 Extraction of the Mode Frequencies

We chose to analyze a sample from a pentatonic bass wind chime available from [6]. At the start of the sound sample each of the five tubes is hit in sequence within 2 seconds, as can be seen from the spectrogram (Fig. 2) and heard by listening to the sample. Thus we divide the signal to segments, each beginning just before a hit and ending just before the next hit. The beginning and ending times of the segments are shown in Table 1.

**Table 1.** Segmentation times of the sound sample in seconds.

| Segment | Begins | Ends |
|---|---|---|
| 1 | 0 | 0.48 |
| 2 | 0.5 | 0.88 |
| 3 | 0.9 | 1.24 |
| 4 | 1.26 | 1.65 |
| 5 | 1.68 | 2.25 |

**Table 2.** Mode frequencies in Hz.

| Mode | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Tube #1 | 219.8 | 590.2 | 1115.4 | 1766.2 | 2513.9 |
| Tube #2 | 245.8 | 657.6 | 1239.2 | 1955.3 | 2773.1 |
| Tube #3 | 293.9 | 782.4 | 1465.2 | 2311.8 | 3278.8 |
| Tube #4 | 331.6 | 875.5 | 1633.0 | 2576.5 | 3654.2 |
| Tube #5 | 366.1 | 967.5 | 1794.2 | 2831.0 | 4015.1 |

Examining the spectra of the segments reveals that the peaks corresponding to the first three modes of each tube are on clearly separate frequency bands. The fourth and fifth modes are clear only in the first two segments, then the bands start to overlap.

Thus the mode frequencies can be extracted in the following way. First the signal is bandpass filtered to isolate the modes. Since after filtering the first segment has only one distinct peak per band, it is easy to locate the peaks. Then the peaks that were located are removed from the remaining segments by using a notch filter [7].

After the equalization, the second segment has only one distinct peak per band. This means that we can just repeat the process of peak picking and equalization to get all the mode frequencies of the lowest two tubes, and the first three mode frequencies for the remaining tubes.

To obtain the remaining frequencies we compute the ratio of the fourth and third mode frequencies from the already known frequencies. Then we simply multiply the known third mode frequency with this ratio to get the fourth mode frequency. Obviously the same method works also for the fifth mode.

The peaks can be described as local maxima above some threshold value. Thus we pick the first frequency bin that is above a suitable threshold value and larger than both of its neighbors. Let the number of this bin be $k$. We fit a parabola to the points $\{(f_{k-1}, X_{k-1}), (f_k, X_k), (f_{k+1}, X_{k+1})\}$, where $f_i$ is the frequency corresponding to frequency bin $i$ and $X_i$ is the absolute value of the FFT coefficient. The frequency value for which this parabola attains its maximum is chosen as the frequency estimate. The previous processing ensures that choosing an appropriate threshold value does not present any difficulties. The values obtained are shown in Table 2.

**Table 3.** Mode gains.

| Mode | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Gain | 0.0787 | 0.1849 | 1.0000 | 0.0136 | 0.0275 |

**Table 4.** 60 dB decay times ($T_{60}$) and corresponding pole radii $R$ for sampling frequency $f_s = 11025$ Hz.

| Mode | $T_{60}$, s | $R$ |
|---|---|---|
| 1 | 40 | 0.99996867280237 |
| 2 | 7 | 0.99982100066035 |
| 3 | 2 | 0.99937364247702 |
| 4 | 1 | 0.99874767727779 |
| 5 | 0.5 | 0.99749714959934 |

### 3.2 Gain Coefficients

The gain coefficients are needed to set the levels of the modes relative to one another. We obtained them by computing the value of the parabola described in the previous paragraph at the mode frequency. We did this for the first segment, i.e. for the modes of the lowest tube. We used the same values for all the tubes. The gain coefficients are shown in Table 3. Note that the third mode is considerably stronger than the other modes.

### 3.3 Pole Radii

The pole radii were computed from 60 dB decay times $T_{60}$ assigned by hand based on inspection of the spectrogram (Fig. 2). This was done because estimating them from the signal turned out to be problematic, yielding values that were clearly too small. For a given sampling frequency $f_s$, the pole radius $R$ can be solved from

$$10 \log_{10}(R^{T_{60} f_s}) = -60. \tag{7}$$

The values used are given in Table 4.

### 3.4 Input for the Mode Filters

The resonator filters need to be driven with some suitable input. Exponentially decaying noise bursts have been previously used for this purpose [8]. This means that the input $x(k)$ at time $k = 0, 1, \ldots, N$ is given by

$$x(k) = a R_{\text{input}}^k n_k, \tag{8}$$

where $R_{\text{input}} < 1$ is the per sample decay rate of the input, $a$ is a gain constant and $n_k$ is a random noise sample drawn from a uniform distribution between $-1$ and 1. The time index value $k = 0$ corresponds to the time at which the triggering event occurs. A burst length of 20 ms was used. For a sampling frequency of $f_s = 11025$ Hz we used $R_{\text{input}} = 0.97$. To reflect the speed of the clapper the gain constant was computed as

$$a = \sqrt{E + \alpha}, \tag{9}$$

where $E$ is the system energy at the start of the burst and $\alpha$ is a constant used to ensure that the gain is never zero. A value of $\alpha = 0.1$ was used.

## 4. SOUND EXAMPLE

We generated a sound example using

$$f(t) = \begin{cases} 0 & \text{if } 0 \leq t \leq 5 \text{ or } t > 20, \\ \frac{2}{5}(t-5) & \text{if } 5 \leq t \leq 10, \\ 2 & \text{if } 10 \leq t \leq 20 \end{cases} \quad (10)$$

as the wind force function $f$ in formula (2). The results are illustrated in Fig. 3.

During the first five seconds of the example, hits occur even though the wind force parameter is zero. (See Fig. 3). This happens because formula (3) gives a positive value to the hit probability even when the energy is zero. This is useful in indicating that the system is working. In this example we wanted to make this feature clearly visible, so we changed the value of the constant $c$ in formula (3) to $c = 19$.

Between 5 and 20 seconds the rate at which the hits occur is roughly comparable to the wind force parameter. The volume of the sound also increases as the wind force increases.

As can be seen from Fig. 3, after the wind ceases at $t = 20$ seconds there are still some hits with a volume above the floor level. Thus the system has some inertia, which is to be expected because the system acts as a storage of energy.

The audio files for this and other examples are available at `http://www.acoustics.hut.fi/publications/papers/norsig04-wind/`.

## 5. CONCLUSIONS AND DISCUSSION

Signal processing techniques to synthesize wind chime sounds were described. We were able to obtain accurate values for the synthesis parameters from the sound sample in a straightforward fashion using just conventional signal processing techniques. Only the pole radii required some manual tweaking. This is remarkable because the sound sample analysed was not recorded with parameter extraction in mind.

The fact that the model described above is parametric makes it flexible. The way in which the model responds to the wind force parameter gives a good idea of what the wind force is on a general level. The resonator parameters could be modified to give an impression of some other tube material, for example wood or even glass.

The virtual wind chime descibed here offers pleasant sounds for many computer based applications. These include games, mobile phones and other portable devices and virtual reality systems. The chime could be used as an auditory indicator, for example so that no wind signifies that the situation is normal, moderate wind signifies a need for increased awareness or caution and strong wind signifies necessity of immediate action.

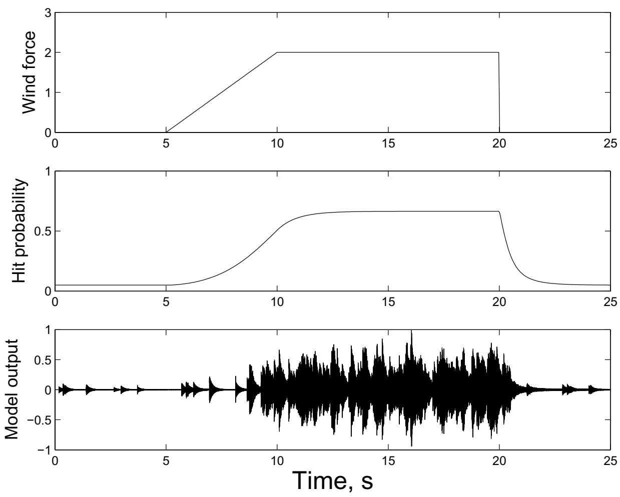


**Fig. 3.** Wind force parameter (top), hit probability (middle) and model output (bottom) as functions of time.

## REFERENCES

[1] C. Thornton, A. Kolb, F. Gemperle, and T. Hirsch, "Audiocentric interface design: A building blocks approach," in *Proc. 2003 Int. Conf. Auditory Display*, Boston, MA, USA, July 2003, pp. 180–183. Available: `http://www.icad.org/websiteV2.0/Conferences/ICAD2003/paper/44%20Thornto%n.pdf`.

[2] P. R. Cook, *Real Sound Synthesis for Interactive Applications*, Natick, MA: A K Peters, 2002.

[3] M. Fernström, "Sound objects and human-computer interaction design," in *The Sounding Object*, D. Rocchesso and F. Fontana, Eds. Firenze: Edizioni di Mondo Estremo, 2003, pp. 33–45. Available: `http://www.soundobject.org/SObBook/SObBook_JUL03.pdf`.

[4] S. M. Ross, *Introduction to Probability Models*, San Diego: Academic Press, seventh edn., 2000.

[5] K. Steiglitz, *A Digital Signal Processing Primer with Applications to Digital Audio and Computer Music*, Addison-Wesley, 1996.

[6] Wind Chimes Inc., "Wind chimes and more," [cited Jan. 22, 2004]. Available: `http://101windchimes.com/`.

[7] P. A. Regalia and S. K. Mitra, "Tunable digital frequency response equalization filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 35, pp. 118–120, Jan. 1987.

[8] P. R. Cook, "Physically informed sonic modeling (PhISM): Synthesis of percussive sounds," *Computer Music Journal*, vol. 21, pp. 38–49, 1997.