The emulation of telephone-like sound is a widely used effect in music, television and film industry. This paper presents a digital model of a vintage telephone sound effect. The effect is derived based on the physical principles of a single-button carbon microphone. The model is realized by using a sandwich model, which consists of a cascade of second-order equalizer filters followed by a nonlinearity and an additive noise generator and a band-pass filter. The model can be used as an effect in singing track processing or in audio antiquing purposes by processing a piece of music with it. The resulting sound can be adjusted by changing the model parameters.

1. INTRODUCTION

The telephony at present day has changed a lot since the early days. Traditional analog telephone networks were digitized, and later on cellular phones superseded the local line telephones. The old fashioned analog telephone technology is still used in some third world countries, but is vanishing as mobile communication technique is getting more popular. Some work has been done to maintain the cultural heritage of vintage audio recordings in form of digital modeling of recording and playback instruments of such era [9].

In the early days of telephony, electrical amplification of audio signals was not possible and limited the transmission distances. Microphones started to evolve during 1870s when Bell patented the electromagnetic telephone using permanent magnets in 1876 that were followed by the invention of carbon microphone by Edison. The invention of the carbon microphone made it possible to increase transmission distances to a level where telephone transmission became an important communication method. Since then carbon microphones were commonly used in telephones for about a century. Carbon microphones were popular because of the ease of manufacturing and fairly low costs. The performance of the carbon microphone was later improved in many steps [3].

The vintage telephone sound is widely used in the music industry. From the musical perspective, the telephone sound effect is typically used in modifying the singing voice. Complete music tracks can be processed with the telephone effect to make an illusion that they have been produced a long time ago. Another application field is the television and film industry where the telephone sound effect is widely used in occasions where a telephone conversation is shown. The nature of audio signal transmission characteristics in analog telephone systems is highly nonlinear. Most of the nonlinearity originates from the carbon microphones; typically the second harmonic is the most dominant. Other sources of nonlinearity include transmission line terminations, switches, and telephone receivers. In general, telephone systems have a very limited transmission band. The first step to model the telephone sound effect is to limit the audio signal bandwidth to match properties of the system to be modeled.

The scope of this paper is to present a novel digital model for reproducing the vintage telephone sounds. This paper is organized as follows: first, the key element of a vintage telephone, the carbon microphone, is discussed and its distortion characteristics are briefly explained. Next, a digital model of a vintage telephone sound effect is presented and its elementary features are explained. Finally, the results are discussed and conclusions are drawn.

2. CARBON MICROPHONES

Carbon microphones are typically made of a cup filled with carbon granules and the cup is covered with an electrically conductive diaphragm [4]. The functionality of a carbon microphone is based on the varying resistance of the carbon granules under a variation of pressure applied to them. A sinusoidal movement of the diaphragm produces alternating current

\[ i = \frac{e}{r_{E0} + hx \sin \omega t}, \]  

Figure 1. Frequency response of a single-button carbon microphone (Type I) from the early 1930s and an improved single-button carbon microphone (Type II) from late 1930s according to Olson [4].
3. PROPOSED MODEL

Digital implementation of a telephone-like sound effect may be based on the sandwich model where a nonlinearity is preceded with a linear pre-filter and followed by a linear bandlimiting post-filter (See Fig. 2). Originally this model was used to estimate the nonlinear behavior of telephone handsets by Quatieri et al. [5] and was later applied to the vintage telephone sound modeling by Välimäki et al. [8].

![Block diagram of the proposed vintage telephone sound effect model.](image)

**Figure 2.** Block diagram of the proposed vintage telephone sound effect model.

Table 1. Pre-filter parameters for Type I and II carbon microphones.

<table>
<thead>
<tr>
<th>Type</th>
<th>ID</th>
<th>$f_c$ (kHz)</th>
<th>Q</th>
<th>$G$ (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I</td>
<td>EQ1</td>
<td>1.8</td>
<td>4.0</td>
<td>+22</td>
</tr>
<tr>
<td>Type I</td>
<td>EQ2</td>
<td>4.3</td>
<td>6.0</td>
<td>+6</td>
</tr>
<tr>
<td>Type II</td>
<td>EQ1</td>
<td>0.28</td>
<td>0.9</td>
<td>+8</td>
</tr>
<tr>
<td>Type II</td>
<td>EQ2</td>
<td>2.0</td>
<td>3.0</td>
<td>+15</td>
</tr>
<tr>
<td>Type II</td>
<td>EQ3</td>
<td>3.9</td>
<td>2.5</td>
<td>+21</td>
</tr>
</tbody>
</table>

3.1. Pre-Filter

The pre-filter is used to create the overall spectral shape and resonances of the model on the passband. The pre-filter design process depends on the complexity of the desired spectrum. The carbon microphone frequency responses presented in Fig. 1 are rather simple consisting of a few resonances. The resonances can be realized by using second-order equalizing filters in cascade [6]. The transfer function of the second-order peak filter is given by

$$H(z) = 1 + \frac{H_0}{2} [1 - A_2(z)],$$

(3)

where $A_2$ is realized using a second-order allpass filter section,

$$A_2(z) = -a_B + (d - d)z^{-1} + z^{-2}$$

$$1 + (d - d)z^{-1} - a_B z^{-2}.$$  

(4)

Parameter $d$ is used for the filter cut-off frequency $f_c$ determination,

$$d = \cos(2\pi f_c / f_b).$$ 

(5)

The resonance gain $G$ in decibels is controlled by

$$H_0 = 10^{G/20} - 1,$$ 

(6)

and the resonance bandwidth $f_b$ by

$$a_B = \frac{\tan(\pi f_c / f_b) - 1}{\tan(\pi f_b / f_b) + 1}.$$ 

(7)

The parameter $a_B$ is used for boosting. The desired responses are constructed by using two or three equalizing filters in cascade. The values of the equalizer-filter control parameters for type I and II microphones are presented in Table 1.

3.2. Nonlinearity

The structure of the carbon microphones is known to behave in a nonlinear manner [2]. The second harmonic is the most dominant distortion component [1]. Quatieri et al. [5] present the Olson model for the carbon button handset nonlinearity modeling.

The physical principles behind this model are presented in [4]. The nonlinearity model of the carbon button based on the physical properties follows the relation

$$Q(u) = \beta \frac{1 - u}{1 - \alpha u} + dc,$$ 

(8)
where \( u \) is the input signal, \( \beta \) is a constant based on the physical properties and \( dc \) is the DC offset value [5]. The nonlinearity is controlled by parameter \( \alpha \). A simplified version of Eq. 8 can be written

\[
Q(u) = \begin{cases} 
\frac{(1-\alpha)u}{1-\alpha u}, & u \leq 1 \\
1, & u > 1.
\end{cases}
\]  

(9)

When \( \alpha \) is small, the Olson model of the carbon microphone nonlinearity can be approximated by using finite order polynomials expanded using a Taylor series

\[
Q(u) = \begin{cases} 
(1-\alpha)\sum_{k=1}^{\infty} \alpha^k u^k, & u \leq 1, |\alpha| < 1 \\
1, & u > 1.
\end{cases}
\]  

(10)

where \( u \) is the input signal. The nonlinearity approximations for degrees from second to fifth can be written as

\[
p_2 = \alpha u^2 - \alpha u + u,  
\]

(11)

\[
p_3 = \alpha^2 u^3 + \alpha^2 u^2 - \alpha u^2 - \alpha u + u,  
\]

(12)

\[
p_4 = \alpha^2 u^3 - \alpha^3 u^4 - \alpha^3 u^4 + \alpha^2 u^4 
\]

\[
-\alpha u^2 - \alpha u + u,  
\]

(13)

\[
p_5 = \alpha^4 u^5 + \alpha^4 u^4 - \alpha^4 u^4 - \alpha^4 u^3 + \alpha^2 u^3
\]

\[
+\alpha^2 u^2 - \alpha u^2 - \alpha u + u.  
\]

(14)

The polynomial approximation for the second- to fifth-order polynomials and the Olson nonlinearity model for \( \alpha = 0.5 \) is presented in Fig. 3. When \( \alpha = 0.2 \), only the second order approximation deviates slightly from the Olson model while higher order approximations are following the original model curve very precisely. The corresponding squared errors are presented in Fig 4. The polynomial approximations start to deviate from the reference model when \( \alpha = 0.5 \), and only the fifth order polynomial approximation is able to follow the reference with some precision.

3.3. Noise related topics

Carbon microphones are known to generate noise [3]. The thermal noise of the carbon microphone is modeled by additive white Gaussian noise (AWGN) after the nonlinearity block, see Fig. 2. The noise power (variance) can be adjusted to meet the requirements for each case. The model was tested by using AWGN having maximum variance of 0.04.

Input signal is also known to create self-induced to carbon microphone output [7]. This is modeled by modulating the amplitude of the bandpass filtered (second-order Butterworth, transition points at 1.5 kHz and 4.5 kHz) noise with the lowpass filtered input signal envelope. The cutoff frequency of the lowpass filter is set to 700 Hz. The modulated bandpass filtered noise is added to the output of the nonlinearity block.

3.4. Post-filter

The post-filter is used to limit the output to a desired bandwidth. A typical bandwidth in the telephony is from 300 Hz to 3300 Hz. The model uses a fourth-order Butterworth bandpass filter with 400 Hz and 2800 Hz as transition points. The order of the bandpass filter can be increased to achieve narrower transition bands.

4. RESULTS

The input and output spectrograms for the music signal is presented in Fig. 5. The output is produced by using type II filter model with fifth order polynomial nonlinearity approximation, \( \alpha = 0.2 \) and noise variance of 0.05. The effect of band-limitation is clearly visible. Model nonlinearity produces new spectral components that can be seen in the output spectrogram. The corresponding time domain difference is also clearly visible, see Fig. 6. The model properties can be varied from a very mild processing (\( \alpha < 0.2 \) and small noise variance (below 0.02)) to a very extreme degradation (\( \alpha > 0.5 \) and increased noise variance > 0.05).
5. CONCLUSIONS

A model for the vintage telephone sound generation was presented in this paper. This model can be used in various phases in music production. The telephone effect is typically used to process voice signals, but the usage of the proposed model could be extended to cover the processing of the entire musical piece to create an effect of an old recording. This model could also be used in television and film industry for making the scenes containing telephone discussion even more realistic.

The proposed model is based on the approximated physical properties of the single-button carbon microphone. The model can be modified to meet frequency characteristics of a desired microphone, the nonlinearity approximation can be modified for an exaggerated distortion effect. The model could also be extended to mimic the properties of the voice transmission in walkie-talkies or in special environment like in the vintage aviation or battlefield communications. Sound examples are available online1.

6. ACKNOWLEDGMENTS

This work is supported by the Academy of Finland, project No. 122815.

1http://www.acoustics.hut.fi/go/icmc11-vintage/

7. REFERENCES


