MODELING OF THE CARBON MICROPHONE NONLINEARITY FOR A VINTAGE TELEPHONE SOUND EFFECT

Sami Oksanen, Vesa Välimäki

Department of Signal Processing and Acoustics,
Aalto University, Espoo, Finland
sami.oksanen@aalto.fi, vesa.valimaki@tkk.fi

ABSTRACT

The telephone sound effect is widely used in music, television and the film industry. This paper presents a digital model of the carbon microphone nonlinearity which can be used to produce a vintage telephone sound effect. The model is constructed based on measurements taken from a real carbon microphone. The proposed model is a modified version of the sandwich model previously used for nonlinear telephone handset modeling. Each distortion component can be modeled individually based on the desired features. The computational efficiency can be increased by lumping the spectral processing of the individual distortion components together. The model incorporates a filtered noise source to model the self-induced noise generated by the carbon microphones. The model has also an input level depended noise generator for additional sound quality degradation. The proposed model can be used in various ways in the digital modeling of the vintage telephone sound.

1. INTRODUCTION

The age of telephone started during the 1870's when Bell patented the electromagnetic telephone. The viability of electroacoustic transmission was increased when Edison patented the carbon microphone. The transmission distances grew to a level where the telephone became an important method of communication due to the fact that they were easy to manufacture at a fairly low cost. Later on, carbon microphone performance was improved in various steps. About of century later, carbon microphones were outdated and replaced with more sophisticated microphone designs [1].

The telephony in the present day has changed greatly since early days. Traditional analog telephone networks were digitized, and later on cellular phones superseded local line telephones. Old fashioned analog telephone technology is still used in some third world countries, but is vanishing as the 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 [2] and also in the field of vintage telephony [3].

The vintage telephone sound effect is widely used in the music industry. From a 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 important field is the television and film industry where the telephone sound effect is widely used is occasions where telephone discussions are shown.

The scope of this paper is to introduce a method for modeling the nonlinear features of a carbon microphone to create a vintage telephone sound effect. The proposed model is flexible and can be adjusted to meet the demands of the computational efficiency. The modeling accuracy can be increased in occasions where more realistic outcome is wanted.

This paper is organized as follows. in Sec. 2 the key element of the vintage telephone, the carbon microphone, is reviewed. Next, the measured distortion characteristics of a carbon microphone are presented. In Sec. 3 the modeling of the carbon microphone nonlinearity is inspected. A novel method for the carbon microphone nonlinearity approximation is presented in Sec. 4 and a case example of modeling a carbon microphone using the proposed model is presented in Sec 5. Finally, the conclusions are drawn in Sec. 6.

2. CARBON MICROPHONE MEASUREMENTS AND DISTORTION ANALYSIS

In general, telephone systems have a very limited transmission band and signal transmission is highly nonlinear [4]. Most of the nonlinearity is originates in the carbon microphones. The second harmonic is usually the most dominant [5]. Transmission line terminations, telephone receivers and switches are also known to cause nonlinearities.

Carbon microphones are typically made of a metallic cup which is filled with carbon granules. On top of the cup there is a diaphragm which will transform the air pressure variation into a movement which will affect the resistance of the carbon granules. The carbon microphone is equipped with a DC bias voltage source which has one pole connected to the electrically connective diaphragm and the other to the bottom of the cup. The changing resistance generates electrical signal which can then be electrically transmitted. The harmonic frequency response of a carbon microphone is typically dominated by peaks around the 2 kHz and 3 kHz areas. Due to their construction, carbon microphones are known to produce harmonic distortion which gives the sound output its unique characteristics [5].

The nonlinear behavior of carbon microphones was analyzed using the nonlinear system identification method [6] based on a Hammerstein model [7]. The nonlinear system under investigation is excited with a swept-sine input signal while the system response is recorded. Then the system response is convolved with the time reversed excitation signal which is multiplied with an exponentially decaying amplitude envelope. As a result, a series of impulse responses representing the linear response (see Fig. 1 upper part) and corresponding harmonic distortion component responses is generated (see Fig. 1 lower part). The linear contribution is the

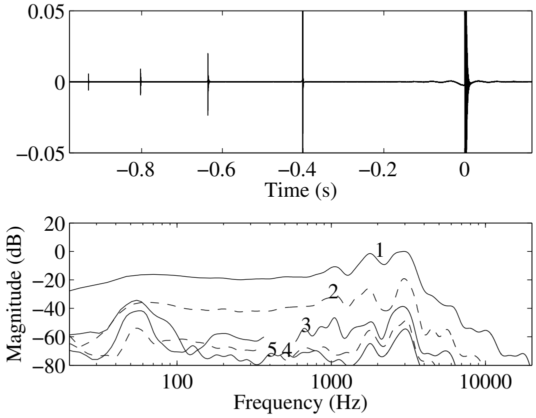

Figure 1: *Upper figure, series of impulses resulting from the logsweep analysis of a carbon microphone. The impulse of a linear contribution is at the time origin and corresponding distortion component impulse responses are on the negative time axis. Lower figure, the linear response is marked with number 1 and the harmonic distortion components with numbers from 2 to 5 (third-octave smoothing is applied).*

most dominant in the deconvolved impulse response and is preceded by the impulse responses of the distortion components.

An example of a carbon microphone response with a few lowest distortion components is presented in Fig 1. The linear response (marked with number 1) follows quite closely the single-button carbon microphone response described in [4]. The harmonic distortion components (marked with numbers 2-5) gives an approximate estimate of the total harmonic distortion as a function of frequency of the microphone under investigation. The second harmonic distortion component is the most dominant whereas the following components are more attenuated. The measurements were done for a telephone carbon button microphone from the late 60s.

The measurement setup consisted of a PC which was controlling the sample playback and recording. The measurement software was written in MATLAB [8]. The measurements were conducted in an anechoic chamber where the excitation signal was produced using an active loudspeaker with known frequency response. The reference signal was measured using B&K 4192 microphone. The carbon microphone under investigation was mounted to a LM Ericsson telephone handset from the 1970's. The carbon microphone was connected to a custom made microphone pre-amplifier which provided a 1.5-V bias voltage to the carbon microphone under investigation. The handset was mounted on a microphone stand that was placed at a 1 m distance in front of the speaker. The microphone responses were measured at approximately 80 dB SPL at the measurement point.

3. PROPOSED MODEL

The proposed model is based on the sandwich vintage telephone model presented by Välimäki *et al.* [9]. The nonlinearity of the proposed model is produced using the Chebyshev model for audio effects presented by Novak *et al.* [10]. The nonlinear components are modeled using Chebyshev polynomials together with individual impulse responses for each distortion component. The main components of the proposed model are presented in Fig 2. First, the input signal is bandlimited by using a linear pre-filter (Line EQ). The nonlinearity is realized by using the output of the pre-filter as an input signal to the distortion component generation block where each distortion component is generated individually. Carbon microphones are known to generate noise [11, 12]. The self-induced noise of the carbon microphone is realized by adding a suitable amount of filtered noise to the pre-filtered and the distorted signal. The noise level is controlled with the G_p parameter. Some extra degradation to the output signal is created by adding some input signal dependent white noise to the filtered noise. Finally, the output signal is constructed by mixing the linear response with the distortion components and the generated noise. The system output is then fine-tuned with a band-limiting post filter (Post EQ). The nonlinear system modeling is done by processing each distortion component.

3.1. Pre-Filter (Line EQ)

The linear pre-filter is constructed based on the measured carbon microphone frequency response instead of the synthetically constructed response presented in [3]. The pre-filter can be realized by using the measured impulse response as FIR filter coefficients. A computationally more efficient realization can be achieved by using for example Prony's method to construct an IIR filter approximation of the impulse response.

3.2. Nonlinearity

According to Novak *et al.* [10] the nonlinearity is produced by reflecting the input signal to the corresponding harmonic distortion component frequency by using the first-order Chebyshev polynomial definition

$$T_n(x) = \cos(n\theta), \quad \text{where} \quad x = \cos(\theta). \tag{1}$$

The corresponding Chebyshev polynomials $T_n(x)$ are defined recursively as

$$T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x), \quad n = 2, 3 \ldots, \tag{2}$$

using the initial conditions

$$T_0(x) = 1, \quad T_1(x) = x. \tag{3}$$

In the proposed model the nonlinearity is produced by using the linear response for each distortion component. However, this is not the optimal solution for modeling of the exact physical behavior of the carbon microphone, but it offers a computationally more efficient solution compared to the use of individual impulse responses for each component.

To add more realism to the model, the distortion component can be adjusted by using a separate Regalia-Mitra equalization filter [13]. This equalization filter can be used to fine-tune the computationally efficient distortion components realization, to add desired spectral tilt or to boost or attenuate certain frequency range. The transfer function of the second-order peak filter is given by

$$H(z) = 1 + \frac{H_0}{2}[1 - A_2(z)], \tag{4}$$

where A_2 is realized using a second-order allpass filter section. The output of the nonlinear section is constructed by creating a weighted sum of the distortion components using the gain coefficients $G_2 \ldots G_n$ as shown in Fig 2.

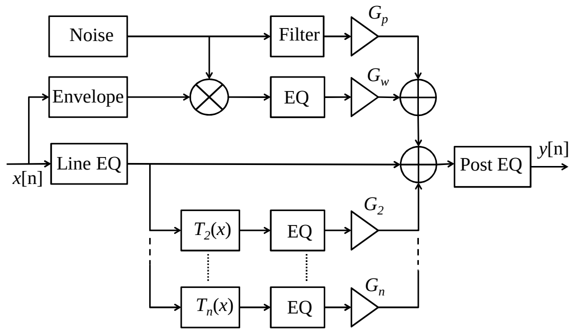

Figure 2: *Block diagram of the proposed effect model.*

3.3. Self-Induced Noise Generation

The input signal dependent noise can be used to increase the impression of failures in the microphone signal. The input signal dependency is created by modulating the white noise amplitude with the input signal envelope. The input signal envelope is obtained by applying the Hilbert transform to it and lowpass filtering the result with a second-order Butterworth filter with cutoff frequency set to 700 Hz. From the measurement results can be seen that the microphone output produces a background noise spectrum close to a $1/f$ relation. The constant background noise is realized by adding a suitable amount of filtered noise to the distorted signal.

3.4. Post-Filter (Post EQ)

The post-filter is used to reduce the bandwidth to the desired level. In telephony the transmission band is typically from 300 Hz to 3400 Hz. A post-filter meeting this requirement can be implemented by using a fourth-order Butterworth bandpass filter with 400 Hz and 2800 Hz as transition points. The sound effect can be enriched by setting the upper threshold value even higher. This would result in to having more harmonic distortion components in the sound effect output. The steepness of the band limitation can also be adjusted to meet the desired output. Finally the distortion components are summed with the linear response and the noise output to generate the sound effect output.

4. MODELING OF THE CARBON MICROPHONE NONLINEARITY

A digital implementation of the telephone sound effect may be based on the sandwich model where a static nonlinearity is preceded by a linear pre-filter and followed by a linear bandlimiting post-filter (See Fig. 3). Originally this model was used to estimate the nonlinear behavior of telephone handsets by Quatieri *et al.* [14]. Recently, this method was applied to telephone sound modeling by Välimäki *et al.* [9].

The nonlinearity approximation is based on Taylor series polynomials which are used as static nonlinearity functions [14]. The following approximation is controlled with the nonlinearity parameter α and is valid for low values

$$Q(u) = \begin{cases} (1-\alpha)u\sum_{k=1}^{\infty}\alpha^k u^k, & u \leq 1, |\alpha| < 1 \\ 1, & u > 1, \end{cases} \tag{5}$$

where u is the input signal and k is the model order. An example of a logsweep analysis of this model output using a measured microphone response is presented in Fig. 4. The model is capable of producing the linear response but fails in producing accurately the distortion components. The level of third-order and higher harmonic components is greatly attenuated compared to the measured response.

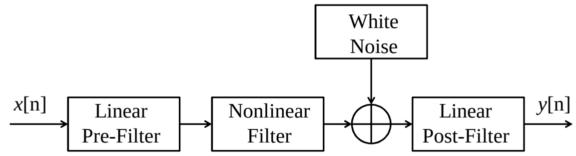

Figure 3: *Sandwich model for the telephone sound effect where the nonlinearity is produced with a static nonlinear function [9].*

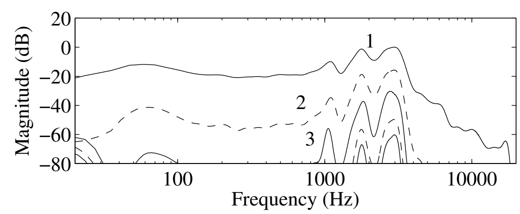

Figure 4: *Logsweep analysis of the simple sandwich model with a static nonlinearity (N=5). The linear response is marked with number 1 and the harmonic distortion components with numbers from 2 to 5 (third-octave smoothing is applied).*

5. RESULTS

The model was tested by constructing the pre-filter from the measured carbon microphone response and the nonlinearity was modeled based on the harmonic distortion analysis. The FIR pre-filter was constructed from the measured system impulse response. To limit the model computational complexity the distortion components were composed from the pre-filtered input signal. This does not exactly follow the actual carbon microphone distortion component behavior, but results in more realistic distortion characteristics (Fig. 5) compared to the simple sandwich model (Fig. 4).

The pre-filtered signal was conditioned to correct the low-frequency attenuation by designing a second-order Regalia-Mitra equalization filter by boosting the filter resonance frequency of 300 Hz by 30 dB. The resonance width was controlled with the Q-value of 0.4. The resulting logsweep analysis of the model output is presented in Fig. 5. It can be seen that the distortion components have a larger gain at the frequencies from 100 Hz to 1 kHz compared to the simplified model. The magnitude of each distortion component can be adjusted by setting the gain coefficients ($G_2 \dots G_n$) to match the desired level. In this example the gain coefficients were adjusted as follows, G_2=0.7, G_3=0.5, G_4=0.4, and G_5=0.5 as the linear part gain was 1. The background noise was controlled with the filtered noise gain setting of G_p=0.009.

The model output for a 1-kHz sinusoidal excitation signal and corresponding measurement are presented in Fig 6. It can be seen that the distortion components produced with this model follow quite closely the measured values. However, it should be noted that the distortion component spectra are approximated with the

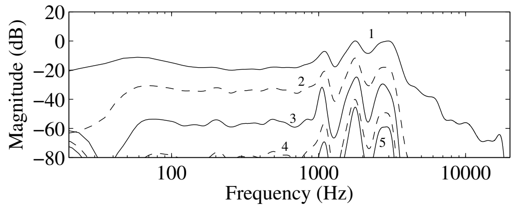

Figure 5: *Logsweep analysis of the model output. The linear response is marked with number 1 and the harmonic distortion components with numbers from 2 to 5 (third-octave smoothing is applied). compare with lower part of the Fig. 1.*

measured linear microphone response and this might result in deviation between the model output compared to the actual response to be modeled. The effect can be reduced by using the measured responses in each modeled distortion component. Sound examples are made available through a website for downloading[1].

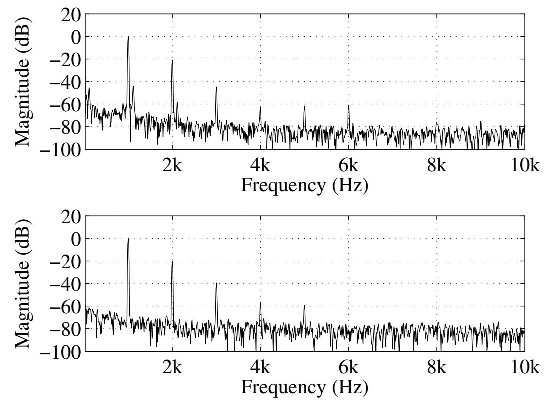

Figure 6: *Response to a 1-kHz sinusoidal signal, measured carbon microphone response in lower figure and model output in lower figure.*

6. CONCLUSIONS

A method for modeling carbon microphone nonlinearities for the telephone sound effect was presented in this paper. The proposed model consists of the modified sandwich nonlinear model and noise generation block. The nonlinearity is modeled by using Chebyshev polynomials and distortion component frequency responses individually for each distortion component. The model allows to control each distortion component and the model can be scaled according to computational limitations.

A case study of modeling carbon button microphone nonlinearity was presented and was found out that the proposed model can be used to approximate the nonlinearities in telephone sound effect. In this example the computational efficiency was improved by modeling the distortion components using the linear frequency response instead of individual responses. The input signal dependent noise can also be added to the degradation process to create more disturbances to the sound.

7. ACKNOWLEDGMENTS

The authors would like to thank Dr. Jyri Pakarinen for the assistance provided during the distortion analysis. The work of Sami Oksanen is funded by the Academy of Finland, project no. 122815.

8. REFERENCES

[1] J. K. Hilliard, "A review of early developments in electroacoustics in the U.S.A.," in *Proc. 57th Audio Engineering Society Convention*, May 1977.

[2] V. Välimäki, S. González, O. Kimmelma, and J. Parviainen, "Digital audio antiquing – signal processing methods for imitating the sound quality of historical recordings," *J. Audio Eng. Soc.*, vol. 56, no. 3, pp. 115–139, 2008.

[3] S. Oksanen and V. Välimäki, "Digital modeling of the vintage telephone sound," in *Proc. ICMC 2011*, Huddersfield, UK, Aug. 2011.

[4] H. F. Olson, *Acoustical Engineering*, Professional Audio Journals, Inc., 1991.

[5] M. T. Abuelma'atti, "Harmonic and intermodulation distortion in carbon microphones," *Applied Acoustics*, vol. 31, no. 4, pp. 233–243, 1990.

[6] A. Farina, "Simultaneous measurement of impulse response and distortion with a swept-sine signal," in *Proc. AES 108th Convention*, Paris, Feb. 2000.

[7] M. Rébillat, R. Hennequin, E. Corteel, and B. Katz, "Identification of cascade of Hammerstein models for the description of nonlinearities in vibrating devices," *J. Sound and Vibration*, vol. 330, no. 5, pp. 1018 – 1038, 2011.

[8] J. Pakarinen, "Distortion analysis toolkit - a software tool for easy analysis of nonlinear audio systems," *EURASIP J. Advances in Signal Processing, Special Issue on Digital Audio Effects*, 2010, Article ID 617325.

[9] V. Välimäki, S. Bilbao, J. O. Smith, J. S. Abel, J. Pakarinen, and D. Berners, "Virtual analog effects," in *DAFX Digital Audio Effects*, Udo Zölzer, Ed., pp. 473–522. 2011.

[10] A. Novak, L. Simon, P. Lotton, and J. Gilbert, "Chebyshev model and synchronized swept sine method in nonlinear audio effect modeling," in *Proc. DAFx-10*, Graz, Austria, Sept. 6–10, 2010.

[11] E. Schiller, "Self induced noise in carbon microphones," in *Proc. 17th AES Convention*, October 1965.

[12] J. R. Sank, "Microphones," *J. Audio Eng. Soc*, vol. 33, no. 7/8, pp. 514–547, 1985.

[13] P. Regalia and S. Mitra, "Tunable digital frequency response equalization filters," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. 35, no. 1, pp. 118–120, January 1987.

[14] T.F. Quatieri, D.A. Reynolds, and G.C. O'Leary, "Estimation of handset nonlinearity with application to speaker recognition," *IEEE Trans. Speech Audio Process.*, vol. 8, no. 5, pp. 567–584, Sept. 2000.

[1] `http://www.acoustics.hut.fi/go/dafx11-carbon`