Modeling the Sound of Astronaut Voice Communications

Aaron Geldert[1] and **Tom Railio**[1]

[1]**Aalto University**, Espoo, Finland

ABSTRACT

Reproducing the sound of vintage audio systems using modern digital signal processing (DSP) is often desirable in audio production. In order to emulate the iconic astronaut voice communications from the 1960's, the combined effects of the microphone, communications systems, and background noises of the original recordings must be efficiently contained in a DSP model. The model presented here is composed of four stages: pre-filtering, nonlinear processing, noise generation, and post-filtering. The stage parameters were manually tuned to convincingly replicate two distinct sounds from NASA's Mercury 6 and Apollo 11 missions. The final model is a simple, parametric DSP algorithm suitable for implementation as a real-time audio effect. This vintage audio effect is relevant to sound design, post-production, and a range of other artistic uses.

1. INTRODUCTION

This work explores reproducing the sonic features of historical astronaut voice communications as an audio effect. The concept of 'vintage sounds' must first be introduced. Vintage sounds evoke a sense of a particular period of history by means of the characteristic timbral features and distortions of legacy audio technology. The emulation of many types of vintage sounds is often desirable in audio production work.

This implementation is an extension of a previous work by Oksanen and Välimäki [1], which modeled the vintage sound of early carbon microphone telephones. The model is constructed of a four-stage processing chain of pre-filtering, waveshaping, multiple noise additions, and post-filtering. Our work demonstrates the versatility of this processing chain by extending it to generate a wider range of distortion effects.

The motivation to model certain astronaut voice communications from the 1960's is based on a combination of their historical and cultural relevance. Aboard Mercury 6, John Glenn became the second human to orbit the Earth. The Apollo 11 mission, broadcast across the globe, saw Neil Armstrong become the first human to walk on the surface of the moon. Armstrong's famous quote about humankind has thus made astronaut voice communications a familiar and desirable vintage sound.

Another objective when designing any audio effect using digital signal processing (DSP) is for it to run efficiently without excessive demands for computation or memory. Rather than attempting to replicate all of the analog stages of the original communications system, the four-stage processing chain seeks to form an intuitive lumped model of the entire system distortions. The waveshaping and noise additions generate harmonic and non-harmonic distortions respectively, while the pre- and post-filters affect the frequency balance of the input and output signals. Implementable with tunable parameters, this architecture is suitable for emulating a wide range of vintage sounds.

It is important to clarify that the implementation presented here does not include any virtual analog modeling, machine learning, or optimization methods. While there exist many state-of-the-art methods for modeling a particular audio system in an optimal manner [2–4], most of these methods are costlier to run in real-time, and more difficult for humans to understand or modify. Additionally, the historical astronaut communications systems to be modeled are likely too complex for proper “white-box" modeling, and lack the training data needed for “black-box" modeling. This implementation instead presents a 'musical' tuning of tools that are already known to be practical for real-time DSP implementations.

Section 2 introduces the features of the reference sounds of astronaut voice communications that guided the model tuning. Section 3 discusses the details of the implementation and model parameters. Section 4 summarizes the results of the modeling, and section 5 contains some concluding thoughts on the model and its applications.

2. REFERENCE SOUNDS

Reference sounds from the Mercury 6 (1962) and Apollo 11 (1969) missions were selected as two notable and unique examples of astronaut speech. Recordings are accessible online from NASA [5]. Critical listening and spectrogram analysis were used to examine the key features of two eras of audio samples. The spectra of both sets' reference signals showed a sharp cut at frequencies above 4 kHz, indicating that they had been band-limited. Energy present in the spectrogram above 4 kHz is assumed to have been generated during an audio coding process, and irrelevant to the sound of the original recordings.

2.1 Expected features

A proper analysis of the transfer function of the astronaut communications system would require a comparison of the

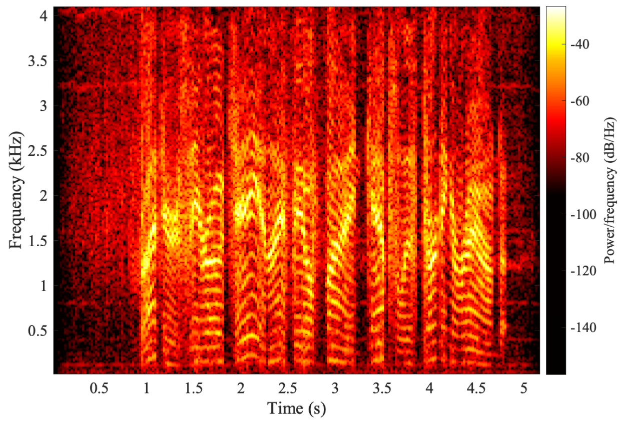

Figure 1. Spectrogram analysis of John Glenn's communication from the Mercury 6 mission in 1962. Speech: "Roger, zero-G and I feel fine. Capsule is turning around."

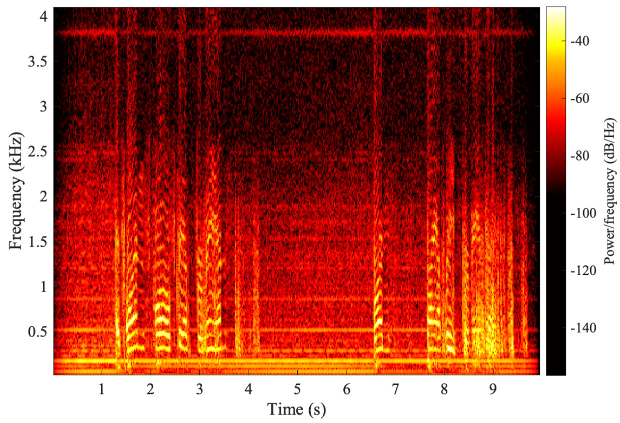

Figure 2. Spectrogram analysis of Neil Armstrong's communication from the Apollo 11 mission in 1969. Speech: "That's one small step for [a] man, one giant leap for mankind."

input and output signals. While the input signal (the acoustic waveform of the astronaut's voice) is not acquirable, we can generalize that it contains the typical features of a male speaking voice. The voice's fundamental frequency typically lies between 50-150 Hz, with a series of harmonically related upper frequencies extending higher. Both the pitch and envelope of the signal should appear time-variant, while any features that exhibit a static frequency and magnitude can be characterized as part of the constant background noise.

2.2 Mercury 6

The Mercury-era sounds exhibited a clearly "brighter" spectrum, with less energy than expected in the lower vocal harmonics and with the most energy concentrated around 1.5 kHz. The spectrogram analysis of an audio sample from Mercury 6 is displayed in Fig. 1. The background noise is composed of broadband noise, with a low cut near 1 kHz. The expected harmonic components of the voice signal are evident, and additional inharmonic noise appears to follow the envelope of the speech signal.

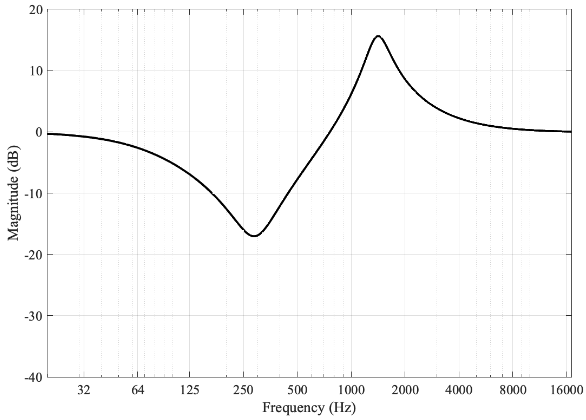

Figure 3. Magnitude response of the tuned pre-filter for the Mercury model.

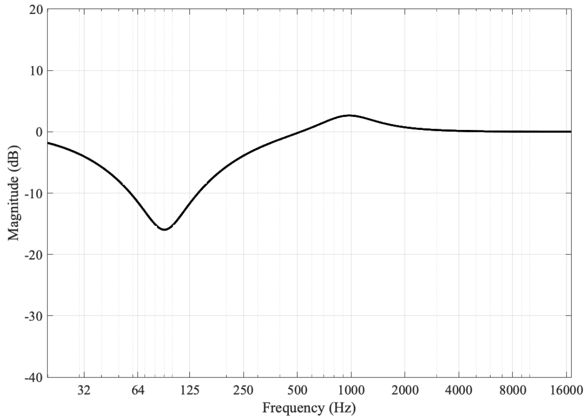

Figure 4. Magnitude response of the tuned pre-filter for the Apollo model.

2.3 Apollo 11

The spectrogram analysis of an audio sample from Apollo 11 is displayed in Fig. 2. The Apollo-era sounds display a greater amount of pitched background noise and appear to be low-pass filtered at a cutoff frequency near 2.5 kHz. The background noise contains a low hum with a fundamental frequency around 58 Hz, as well as broadband noise. There is also some inharmonic noise following the envelope of the speech signal.

3. MODEL IMPLEMENTATION

The astronaut voice distortion effect is performed on each channel of an audio signal in four processing stages: pre-filtering, waveshaping, noise addition and post-filtering. Fig. 5 shows the layout of the general model. Each stage contains tunable parameters, and the examples presented have already been manually tuned to best resemble the Mercury or Apollo-era reference sounds we selected. All equalizing filters were designed using a standard biquadratic digital filter design method [6]. All stage parameters are easily tunable, making the model suitable to recreating many other vintage sounds with similar features to astronaut voice distortion.

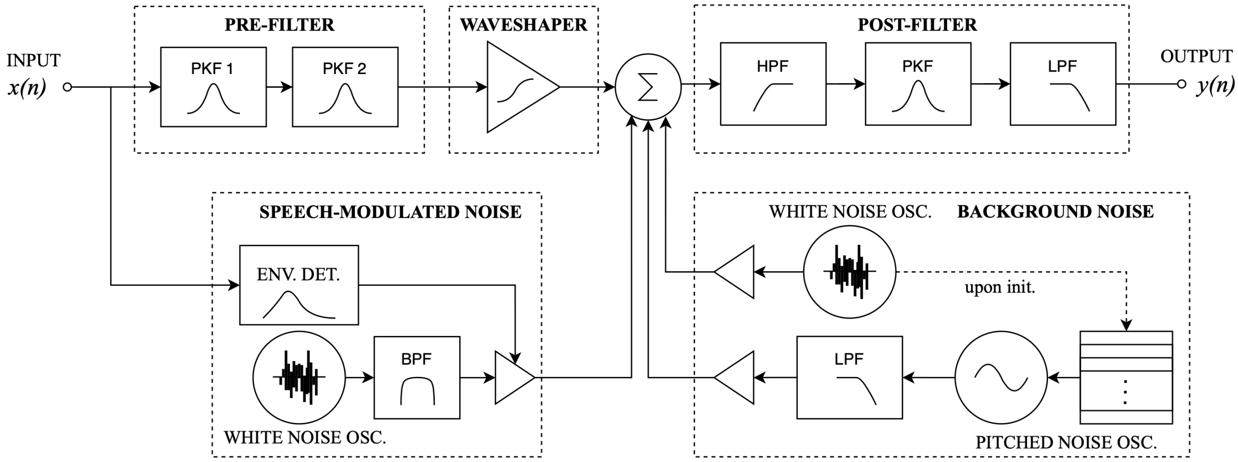

Figure 5. Block diagram of the general astronaut voice distortion effect.

3.1 Pre-Filtering

Each dry signal is first processed through a two-band pre-filter. The pre-filter is composed of two cascaded peaking filters with parameters for center frequency, peak gain, and Q factor. Each vintage model implemented a low-mid frequency cut and a high-mid frequency boost, with varied parameters.

3.2 Waveshaping

A waveshaping algorithm is applied to the signal following the pre-filter in order to create nonlinearity. Waveshaping algorithms can vary in complexity, but are always a fundamentally nonlinear function operating on the audio sample data. This nonlinearity may alter the signal by adding harmonic distortion, alter its dynamic range, or generate asymmetry. The generation of harmonic distortion also causes higher harmonics at frequencies above half the sample rate (Nyquist frequency) to alias. Though it is often appropriate to suppress these artifacts by employing oversampling [7], the current model can tolerate significant aliasing artifacts. This is because the model, by design, is band-limited with a high noise floor.

A simple memoryless waveshaping function can be applied to each input audio sample to then calculate the output sample. Our implementation applies a simple sigmoid waveshaping function that is suitable for many vintage distortion effects, especially vacuum tube amplification and magnetic tape saturation [8].

The waveshaper is completely described in Eq. 1, where $y(n)$ is the output sample, $x(n)$ is the input sample, and k is a distortion intensity parameter. Fig. 6 shows the relationship of the input and output sample values for different distortion intensities. The distortion intensity parameters of the tuned Mercury and Apollo models are set to $k = 6$ and $k = 16$, respectively.

$$y(n) = \frac{\arctan\left(kx(n)\right)}{\arctan(k)} \tag{1}$$

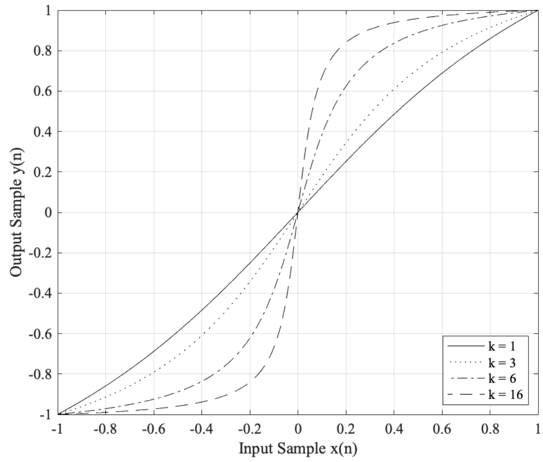

Figure 6. Input/output relationship of sigmoid waveshapers set to a range of parameter values.

3.3 Noise Additions

Two distinct types of noise need to be generated to match the noise features found in the reference audio. These noises are differentiated by whether their amplitude correlates with the speech signal or remains steady. Note that all noise additions will still pass through the post-filter stage, modifying which spectral characteristics are ultimately audible.

3.3.1 Background noise

The first noise addition is the static background noise. This noise is composed of two elements: an “unpitched" white noise and a “pitched" harmonic noise. The pitched background noise is generated by looping a buffer of white noise with a period according to the desired pitch. This noise looping is essentially equivalent to wavetable synthesis, using a single wavetable composed of random values. To attenuate the harsher high harmonics, the pitched noise is then passed through a second order low-pass filter with a 180 Hz cutoff frequency. Both the unpitched and pitched

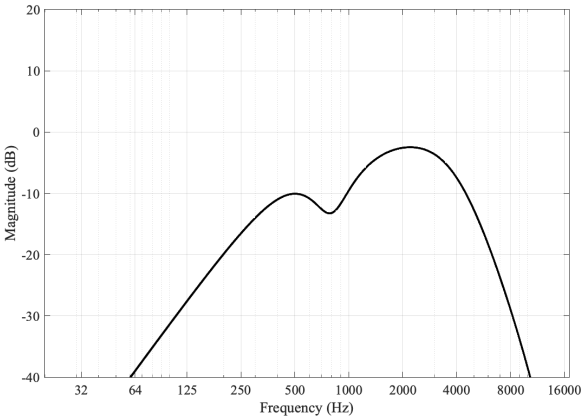

Figure 7. Magnitude response of the tuned post-filter for the Mercury model.

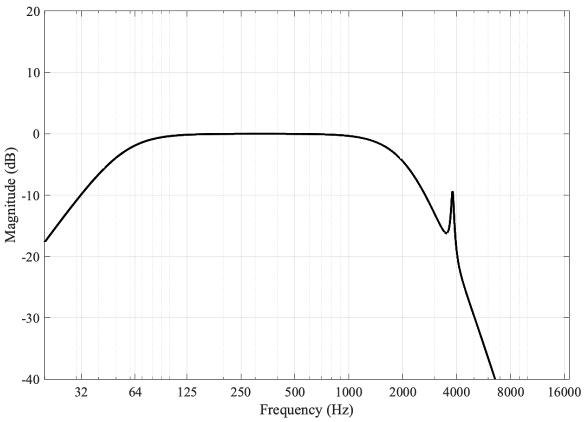

Figure 8. Magnitude response of the tuned post-filter for the Apollo model.

noises are then mixed with the speech signal.

3.3.2 Speech modulated noise

The second noise addition is the noise generated in conjunction with speech, which may correspond to the breath noise of the microphone capsule or speech coding noise. First, white noise is generated and band-passed from 0.5 to 2.7 kHz. By detecting the envelope of the speech signal, the band-passed noise can be modulated to match the envelope of the speech. This speech modulated noise is then mixed with the speech signal at a level set by a parameter.

3.4 Post-Filtering

The final processing stage is a post-filter composed of three bands. Crucially, a high-pass and low-pass filter bandlimit the signal, as well as ensuring that any DC offset that may have been generated in a previous stage is removed. The low-pass edge is fourth order with a 24 dB /octave slope, made of two identical second order filters in series. Another peaking filter band is also used to shape the final signal's spectrum to better match the reference. The Mercury-modeled peaking band uses a slight mid-frequency cut, while the Apollo model includes a sharp peak at 3.8 kHz.

4. RESULTS

Several speech samples of a male voice, recorded using a close microphone placement, were used during development and testing to tune the sound of the models. The model parameter tuning was done informally through a combination of subjective critical listening judgements and spectrum analysis.

The filter parameters of the final tuned models are specified in Tables 1 and 2. The optimal tuning depends on the frequency balance of the input audio signal, so further equalization may be necessary for novel audio samples.

Stage	Pre-Filter		Post-Filter		
Band	PKF1	PKF2	HPF	PKF	LPF
Freq. (Hz)	290	1.4k	600	800	3.8k
Gain (dB)	-18	+18	-	-12	-
Q	0.71	1.2	0.71	1.2	0.71

Table 1. Filter parameters for the tuned Mercury-era astronaut voice distortion model.

Stage	Pre-Filter		Post-Filter		
Band	PKF1	PKF2	HPF	PKF	LPF
Freq. (Hz)	90	950	55	3.8k	2.2k
Gain (dB)	-16	+3	-	+11	-
Q	0.8	1	0.71	15	0.71

Table 2. Filter parameters for the tuned Apollo-era astronaut voice distortion model.

An audio sample of male voice was used to demonstrate the spectral effects of the astronaut voice distortion algorithm. Fig. 9 shows the expected harmonic structure of the voice with little background noise. Figs. 10 and 11 show the speech after being processed with the Mercury and Apollo distortion models, respectively. In comparison with Figs. 1 and 2, the overall similarity of the spectrogram features is evident, though the exact speech content differs.

These subjective results indicate that this implementation is convincing in the context of a musical or artistic audio effect. In addition, we demonstrated the usefulness of the architecture presented by [1] for producing a variety of distortion effects. MATLAB example code, sound examples, and use cases are available online. [1]

5. CONCLUSIONS

A digital signal processing model for simulating the sound of astronaut voice communications has been presented. Building off of previous work that simulated the vintage telephone sound [1], we have adapted and tuned the model to replicate the vintage sounds of two historical NASA mission recordings. The model parameters were manually tuned to emulate the notable features of this vintage audio effect. The proposed model has only been informally evaluated through critical listening and spectrum analyses. Comparin the Mercury and Apollo missions' audio also

[1] `https://github.com/ageldert/Astronaut-Voice-Distortion`

shows how the specific communication technologies differ between these two periods of space exploration.

An astronaut voice distortion effect range is highly applicable as a creative tool for sound design and post-production in film, music, video game, podcast, and other multimedia productions.

Further research might extend this model to other characteristic voice communications sounds besides the Apollo or Mercury missions. A variety of modern space-themed audio effects may also be of interest in the future. NASA's fifth rover, *Perseverance*, landed on the surface of Mars on February 18th, 2021. *Perseverance* carries microphones on board to record sound in the thin Martian atmosphere and may soon introduce new concepts about audio samples from space. As the iconic sounds of space exploration evolve in the modern era, so will the motivation to create evocative audio effects.

Acknowledgments

Special thanks go to Doctoral Candidate Alec Wright and Professor Vesa Välimäki for proposing the project work and conference submission.

6. REFERENCES

[1] S. Oksanen and V. Välimäki, "Digital modeling of the vintage telephone sound," in *ICMC*, 2011.

[2] J. Pakarinen, V. Välimäki, F. Fontana, V. Lazzarini, and J. S. Abel, "Recent advances in real-time musical effects, synthesis, and virtual analog models," *EURASIP Journal on Advances in Signal Processing*, vol. 2011, pp. 1–15, 2011.

[3] E.-P. Damskägg, L. Juvela, V. Välimäki *et al.*, "Real-time modeling of audio distortion circuits with deep learning," in *Proc. Int. Sound and Music Computing Conf.(SMC-19), Malaga, Spain*, 2019, pp. 332–339.

[4] M. A. M. Ramírez and J. D. Reiss, "Modeling nonlinear audio effects with end-to-end deep neural networks," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2019, pp. 171–175.

[5] National Aeronautics and Space Administration, "Audio and ringtones," https://www.nasa.gov/connect/sounds/index.html, 2021, accessed: 2021-02-16.

[6] R. Bristow-Johnson, "Cookbook formulae for audio eq biquad filter coefficients," http://www.musicdsp.org/files/Audio-EQ-Cookbook.txt, 2016, accessed: 2021-02-16.

[7] J. Kahles, F. Esqueda, and V. Välimäki, "Oversampling for nonlinear waveshaping: Choosing the right filters," *Journal of the Audio Engineering Society*, vol. 67, no. 6, pp. 440–449, 2019.

[8] W. Pirkle, *Designing Audio Effect Plug-Ins in C++: With Digital Audio Signal Processing Theory*. Focal Press, 2012, pp. 497–499.

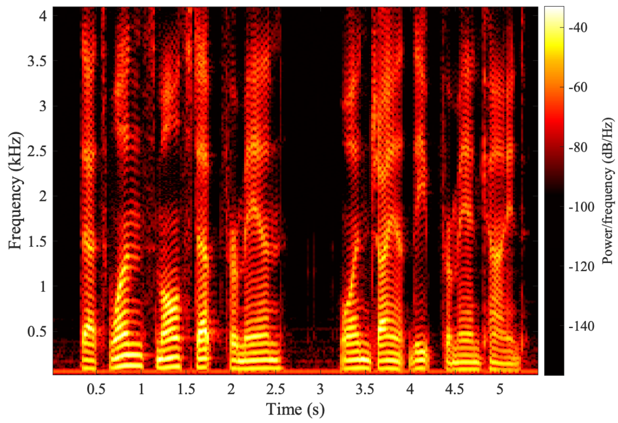

Figure 9. Spectrogram analysis of male speech without any effects processing. Speech: "That's one small step for man, one giant leap for mankind."

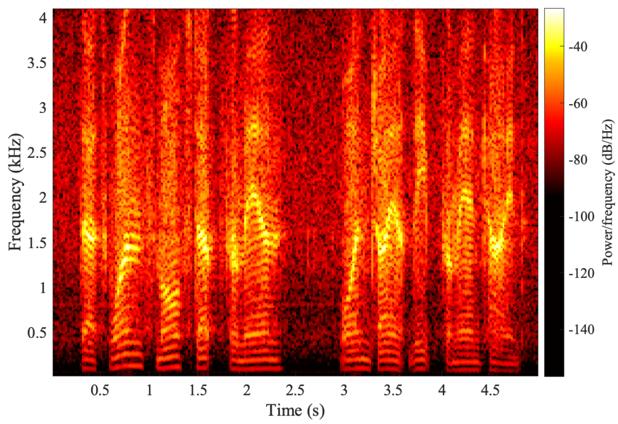

Figure 10. Spectrogram analysis of male speech processed with the Mercury-era astronaut voice distortion model.

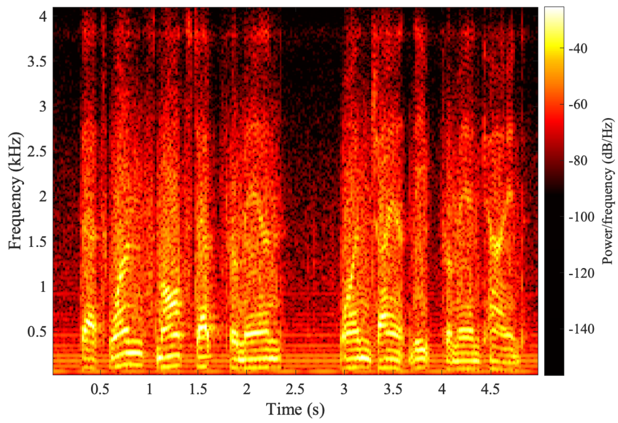

Figure 11. Spectrogram analysis of male speech processed with the Apollo-era astronaut voice distortion model.