Project Report on

TALKING TOM

Akhilesh Anant Gonabal (16EE203)

Under the Guidance of,

Dr. Krishnan CMC

Department of Electrical and Electronics Engineering, NITK Surathkal

Date of Submission: 03-17-2019

in partial fulfillment for the award of the degree

of

Bachelor of Technology

In

Electrical and Electronics Engineering

At

Department of Electrical and Electronics Engineering

National Institute of Technology Karnataka, Surathkal

National Institute of Technology Karnataka

Department of Electrical and Electronics Engineering

CERTIFICATE

This is certify that the project entitled **'TALKING TOM',** submitted by **Akhilesh Anant Gonabal (16EE203)** is a record of bonafide work carried out by them, in the partial fulfilment of the requirement for the award of Degree of B.Tech in Electrical and Electronics Engineering at National Institute of Technology Karnataka, Surathkal.

Dr. Krishnan CMC

Assistant Professor
Department of EEE
NITK, Surathkal

Contents

List of Figures

1 Introduction

Talking Tom is a virtual pet app in which a virtual animal squeakily repeats anything that the user says into the device's microphone, while interacting with the animal by tapping and swiping on the screen. The app records the voice of the user through microphone, in digital format, which is then modulated to a certain pitch/tempo and played through the microphone, creating a Chipmunk-like effect.

In this project, various possible techniques for voice conversion have been explored. These techniques can be categorized as : Time Domain and Frequency Domain based Pitch Shifting.

2 Implemented Methods

In the standard time-domain pitch-shifting technique, for pitch-shifting the signal by a factor of α, the input signal is first re-sampled by a re-sampling factor equal to $1/\alpha$. Since re-sampling changes the length of the signal, a timescale modification method is used to preserve the time duration of the original signal. The time duration of the re-sampled signal should be scaled by a factor equal to α.

The Delay based method works in the time domain and crossfades between two channels with different varying delays and gains to produce a smoothly transitioned pitch shifted signal. The tracks are then re-synthesized at a new time scale. This method can yield good results on both polyphonic and percussive material, especially when the signal is separated into sub-bands. In Frame Based Approach, the first step is to split the signal into short analysis frames of fixed length.

The analysis frames are spaced by a fixed number of samples, called the analysis hopsize. To achieve the actual time-scale modification, the analysis frames are then temporally relocated to have a synthesis hopsize. This frame relocation results in a modification of the signal's duration by a stretching factor of α.

3 Delay Based Pitch Shifting

The delay based method works in the time domain and crossfades between two channels with different varying delays and gains to produce a smoothly transitioned pitch shifted signal.

To create an upward pitch change, we used a 30-millisecond delay and steadily decrease it at a rate that yields the desired pitch change. As the delay approaches 0, a second delay channel is started at 30 milliseconds and sweeps in a similar manner. A quick crossfade from the first to the second channel is applied, making sure the first channel is completely faded out before its delay reaches 0. This process is repeated, going back and forth between the delay channels.

A downward pitch change is achieved in a similar manner, only the delay channels are started with a near zero initial delay, and the delay is increased out to around 30 milliseconds, at which time the alternate channel is started and the cross-fade performed. The changing delays and cross fading pattern is shown below:

The desired output pitch may be controlled by varying the rate of change of the channel delays. Cross-fading reduces the audible glitches that occur during the transition between channels.

Pros

- Fast to process
- No filtering
- Better quality than SOLA

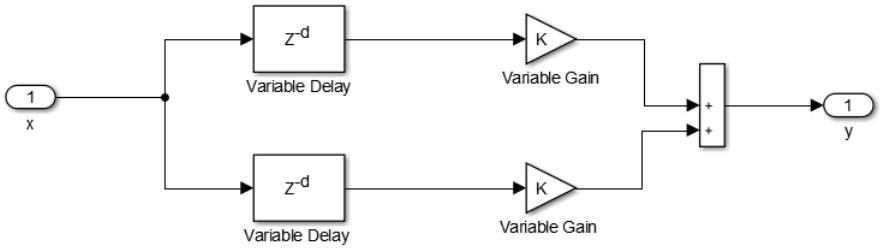

Figure 3.1: The figure illustrates the pitch shifting algorithm.

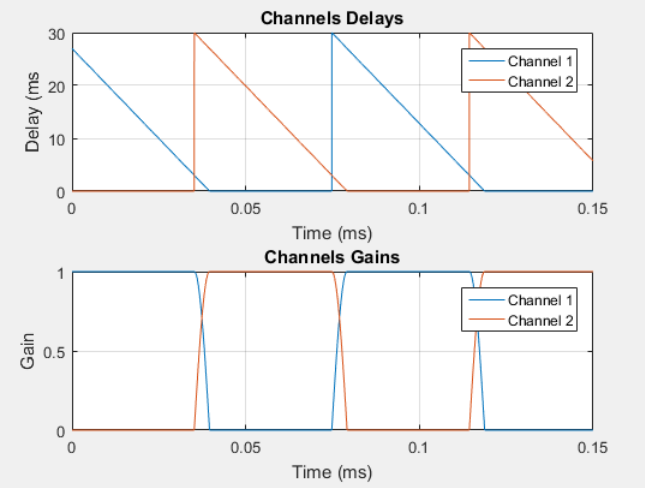

Figure 3.2: The changing delays and cross fading pattern

Cons

- Unwanted smeared frequencies
- Unintuitive

4 Conclusions

Delay based pitch shifting algorithm was found to give clearer results than other time domain techniques. Hence, in general, time-domain techniques are simple and fast, and work fine for periodic and quasi-periodic signals. However, their quality is not good for signals which contain a lot of non-harmonic components. On the other hand, frequency- domain algorithms are more suitable for complex signals, but the price of the high- quality is the computational complexity.

References

[1] Delay-Based Pitch Shifter - MATLAB and Simulink - MathWorks India, in.mathworks.com/help/audio/examples/delay-based-pitch-shifter.html

[2] https://in.mathworks.com