

Sound Effector Based on Speech Recognition



UNIVERSITY of
ROCHESTER

Yuhui Chen, Electrical and Computer Engineering, University of Rochester
Shengwen Yang, Electrical and Computer Engineering, University of Rochester

Abstract

- Objectives
- Introduction
 - Speech recognition
 - Sound effect
- Algorithm
 - Feature extraction
 - Codebook generation
 - Speaker recognition
- Functions
- Implementation
 - Speech recognition
 - Sound effector by voice control

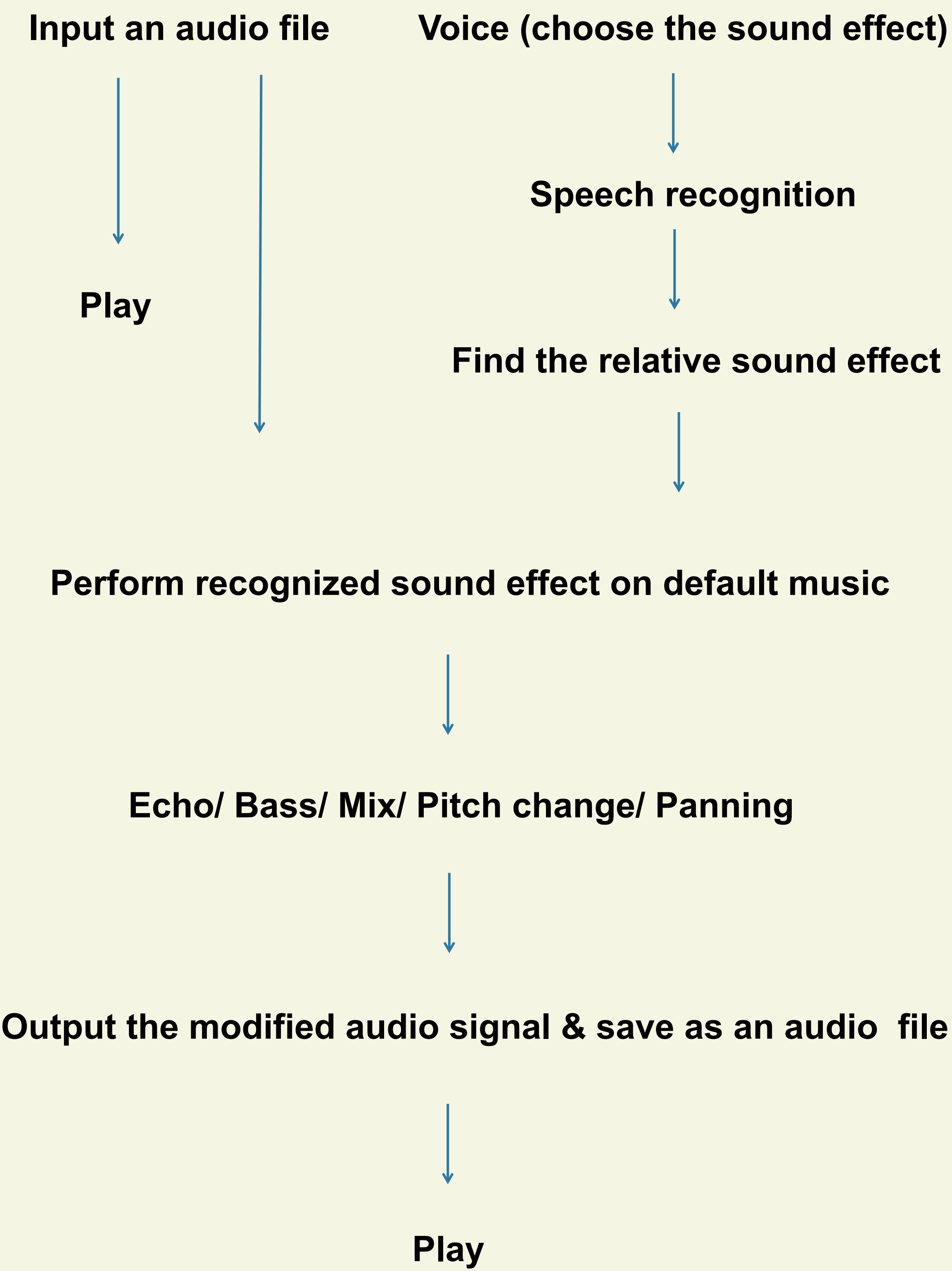
Speech Recognition

- Speech recognition (SR) is the translation of spoken words into text. It is also known as "automatic speech recognition" (ASR), "computer speech recognition", or just "speech to text" (STT). Nowadays, speech recognition is widely used in vehicle-mounted system, electronic products, such as Siri of Apple, and medical area.
- The speech signal contains both the message being spoken and the information of the speaker. Therefore we can use the speech signal for both speech and speaker recognition. The goal of these two tasks can be summarized as follows:
 - Speech recognition: Extract the underlying linguistic message in an utterance.
 - Speaker recognition: Extract the identity of the person speaking the utterance.
- Speaker recognition can be further divided into two different tasks:
 - Speaker identification: To determine which one of a group of known voices best matches the input voice sample.
 - Speaker verification: To determine from a voice sample if a person is whom claims to be.

Sound Effects

- Echo
 - adding echo effect to the audio
- Bass
 - creating bass effect
- Mix
 - mixing another audio signal with the input
- Pitch change
 - resampling
- Panning
 - creating stereo audio and panning from left channel to right channel

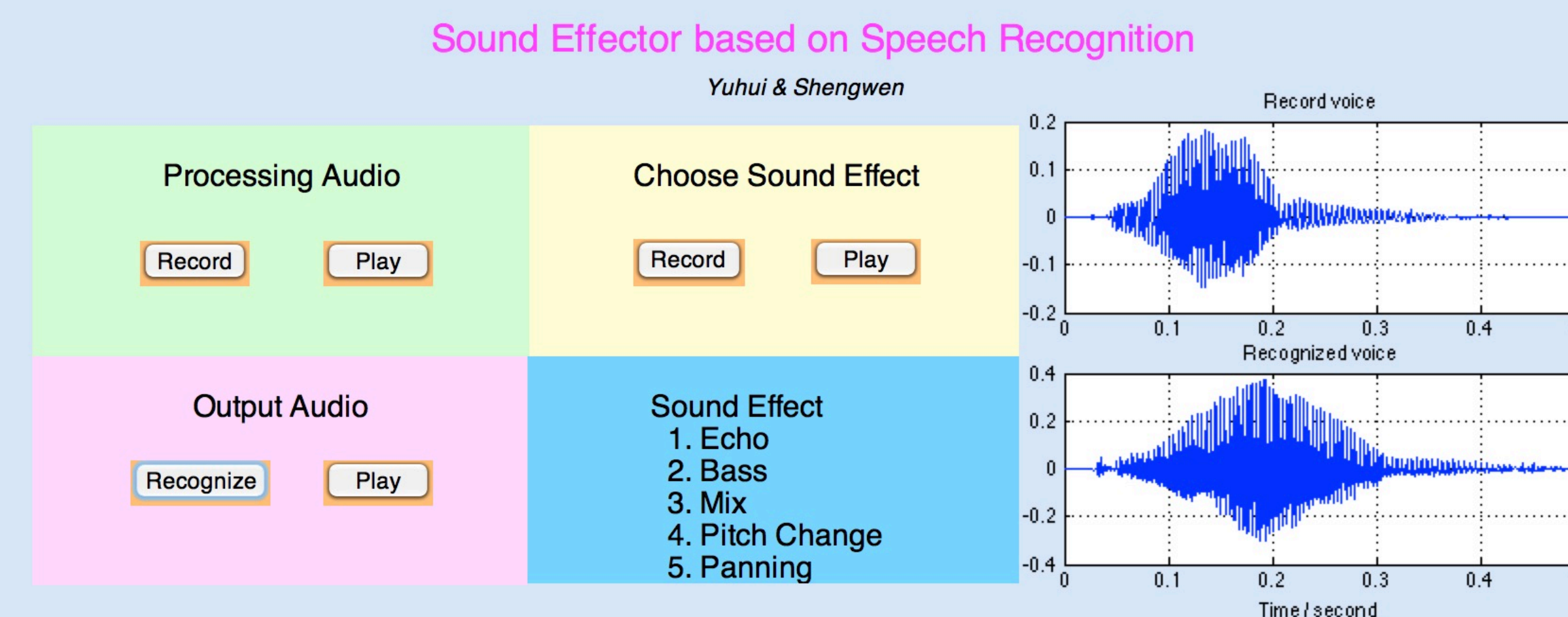
System Overview



Sound Effector

- Spectrogram of Echo
- Spectrogram of Bass
- Spectrogram of Mix
- Spectrogram of Pitch change
- Spectrogram of Panning
 - Left channel
 - Right channel

Graphical Users Interface



Feature Extraction

