

Audio Event Detection for Automatic Scene Recognition

Xun Xu

Department of Computer Science and Engineering
Shanghai Jiao Tong University

June 19, 2015

- 1 Introduction
- 2 Audio Event Detection
- 3 Scene Recognition
- 4 Evaluation
- 5 Demo

- 1 Introduction
 - Problem Description
 - Approach
- 2 Audio Event Detection
- 3 Scene Recognition
- 4 Evaluation
- 5 Demo

Problem Description

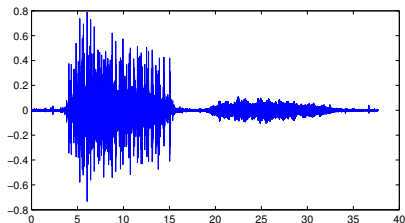
In this project, our problem is to recognize a scene where an audio is recorded. Sound example:

Play Sound

Problem Description

In this project, our problem is to recognize a scene where an audio is recorded. Sound example:

Play Sound



⇒

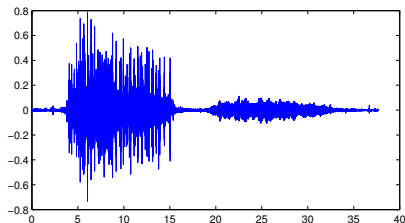
concert

Our Approach

Our approach is to detect the audible events in a clip.
Then infer the scene from the detected events.

Our Approach

Our approach is to detect the audible events in a clip.
Then infer the scene from the detected events.



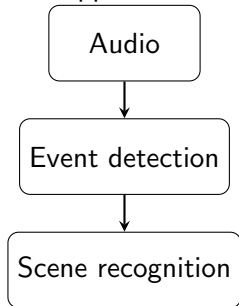
*applause,
instrument*



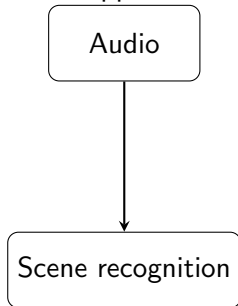
concert

Our Approach vs. Other Approaches

Our approach:



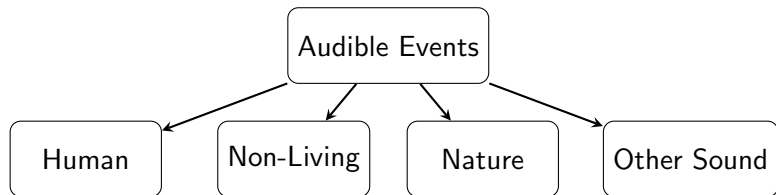
Other approaches:



- 1 Introduction
- 2 Audio Event Detection
 - Audible Event Taxonomy
 - Audio Data
 - Feature Extraction
 - Event Model
- 3 Scene Recognition
- 4 Evaluation
- 5 Demo

Audible Event Taxonomy

We labelled common audible events into 4 classes, there are 120 events in total.



We download the audio data for events from Sound Search Engines (SSEs). (List SSEs here)

Feature Extraction

The features we used for training model for audible events are MFCCs.

We use Gaussian Mixture Models to train the features. (Explain why using GMMs)

- 1 Introduction
- 2 Audio Event Detection
- 3 Scene Recognition
 - Scene-Event Map
 - Audio Segmentation
 - Scene Inference
- 4 Evaluation
- 5 Demo

We use the scripts for movies, plays and TV series to extract the scenes.

We calculate the TF-IDF scores for scenes and events.

Audio Segmentation

In testing, we segment the audio into smaller parts for event detection.

- 1 Introduction
- 2 Audio Event Detection
- 3 Scene Recognition
- 4 Evaluation**
 - Event Detection Evaluation
 - Scene Recognition Evaluation
- 5 Demo

Scene Recognition Evaluation

- 1 Introduction
- 2 Audio Event Detection
- 3 Scene Recognition
- 4 Evaluation
- 5 Demo

Live demo for our system.

Thank you!

Any Question?