BARNHOUSE KARAOKE: RECREATING POPULAR MELODIES WITH ANIMAL SOUNDS

TECHNICAL REPORT

Patricio Ovalle
Universitat Pompeu Fabra
Barcelona, Spain
patricio.ovalle01@estudiant.upf.edu

ABSTRACT

1 Introduction

Mosaicing

Everyday sounds and music. That's why picked animal sounds and popular melodies. Intent: something everyone can relate to, fun and lighthearted

I took inspiration from which replaces singing voice with cats meowing. See their demo of Nirvana's XXX.

2 Methodology

2.1 Target sound selection

I tried to choose target sounds that are musical and have either simple or well-known melodies that are easy to sing along to.

Another important consideration was choosing a sound with not too much complexity in the mixture. To keep the scope of the prototype small I am not implementing a source separation preprocessing step and am not using a very general pitch tracking algorithm. I selected only target sounds in which the melodies were predominant and in a distinct pitch range from other instruments in the arrangement.

The target sounds I chose are shown in Table 2.

MBID, title, timestamp xyz, some acapella, :40-:51 abc, Somewhere Over the Rainbow, :20-:30 def, Old Macdonald Had A Farm, :30-:40 ghi, Merry Christmas Mr Lawrence, :10-:30

The first two have a vocal melody that I hope to capture

The third has a piano melody

2.2 Source sound collection

Source sound collection

I chose a set of animal sounds that tend to be short and clearly recognizable. I made sure to choose a set of animals with distinct timbres from one another in hopes of adding a bit of dynamism to the result.

The parameters I used to query Freesound for the sound collection are shown in Table 1.

count	keyword	duration cutoff (seconds)
80	meow	3
30	dog bark	3
30	growl	3
40	moo	5
5	horse whinny	5
5	horse neighing	5

Table 1: Distribution of valence arousal quadrants

I chose a large selection of sounds to increase the chance of having high coverage in pitch range. For the same reason I also chose animals whose vocal cords tend to produce sounds in different registers, focusing on those that overlap with the range of the human voice, XXX hz to XXX hz.

With the exception of the 'moos', each collection was filtered to be less than X seconds to increase the chances that the query would return single-shot sounds, which I speculate are typically higher quality than longer field recordings and more suitable for mosaicing popular melodies.

The 'moos' were filtered to be less than Y seconds because I think cows take longer to vocalize their thoughts.

2.3 Frame selection strategy

I experimented with two sets of features for frame selection.

duration, [pitch feature(s)], MFCCs

duration, [pitch feature(s)]

MFCCs capture timbre I tried with and without MFCCs because I expected timbre to be less important to a convincing result than pitch. I wanted to see if omitting the MFCCs would improve the matching of pitch.

2.4 Recombination strategy

I used the XXX pitch tracking algorithm from the Essentia audio analysis toolkit.

This algorithm uses XXX and XXX, which makes it well-suited for this application because the target sounds I am using are XXX.

3 Further Experiments

Out of curiosity I also went on to perform audio mosaicing on some non-melodic target sounds. This is probably out of scope for this project, but I wanted to share my results.

In particular I was interested in speech and prosody I was reminded of XXX (youtube video of non-english language that really sounds like english)

This is a future direction that I am keen to explore. Using a source sound collection of speech and a target sound of speech to create disorientingly realistic non-speech speech.

I wonder if you would end up with phonemes and truncated syllables from the source sounds, or maybe the same sentence but each word from a different speaker, or a string of English words that together don't make sense (The running black hi ball joust your back home is back to the left left in left of your pocket bag and left)

I also included a couple of target sounds without melodies.

4 Discussion

References