ASSESSING THE SUITABILITY OF PANAKO FOR MUSIC IDENTIFICATION IN MOVIES

BY RUBEN NAIR
SUPERVISORS: DR. JAEHUN KIM, DR. CYNTHIA LIEM
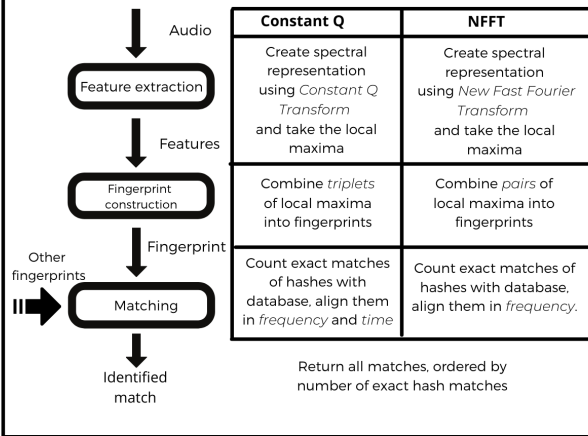
TUDelft

1 INTRODUCTION

This poster presents the research in assessing the suitability of Panako for movie music identification. This was done by answering the research question: How does Panako perform in music identification in movies?
This question was split into two sub-questions:
1. How well does Panako perform in recognising music in movies, based on the benchmark?
2. What is the influence of configurable parameters on the performance of Panako?

2 PANAKO

Audio
↓
Feature extraction
↓ Features
Fingerprint construction
↓ Fingerprint
Other fingerprints ▶ Matching
↓
Identified match

	Constant Q	NFFT
	Create spectral representation using *Constant Q Transform* and take the local maxima	Create spectral representation using *New Fast Fourier Transform* and take the local maxima
	Combine *triplets* of local maxima into fingerprints	Combine *pairs* of local maxima into fingerprints
	Count exact matches of hashes with database, align them in *frequency* and *time*	Count exact matches of hashes with database, align them in *frequency*.
	Return all matches, ordered by number of exact hash matches	

3 METHOD

Analysis was performed in three steps and repeated for both strategies (Constant Q and NFFT):
1. **Preliminary study**. Six movies were manually labeled. Clips containing music were extracted and queried using Panako's default settings.
2. **Analysis using the synthesised data from the benchmark.** This data was first tested with the default configuration of Panako. Then, all configurable parameters were changed. They were all doubled and halved where possible,
3. **Testing better configurations**. The configurations that performed better on the synthesised data than the default configuration were tested on the real movie clips from the first step.
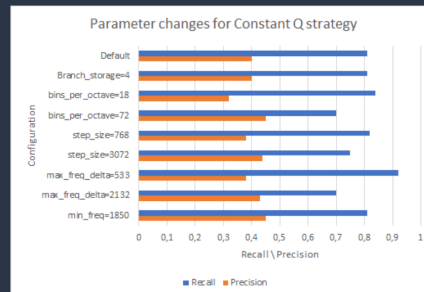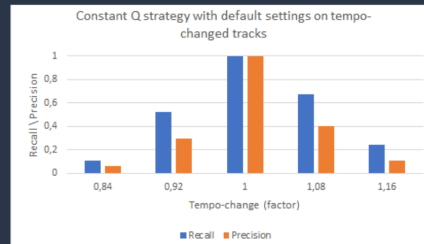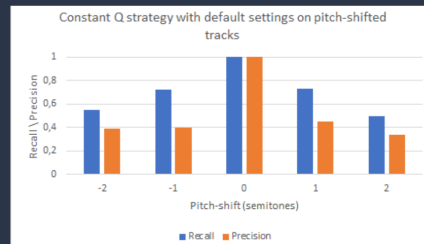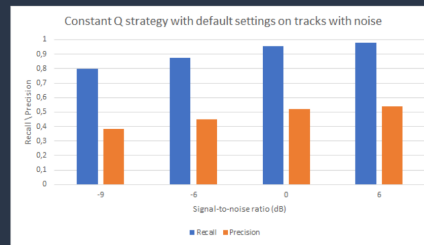
5 CONCLUSION

*   Panako performs well on the soundtracks layered with noise, but not for pitch shifts and tempo changes.
*   The configurations that performed best on the benchmark find less than 10% correct matches on real movie data, so within the scope of this research, Panako is not suitable for movie music identification.
*   This research can be extended by evaluating specific combinations of parameters and more values per parameter.
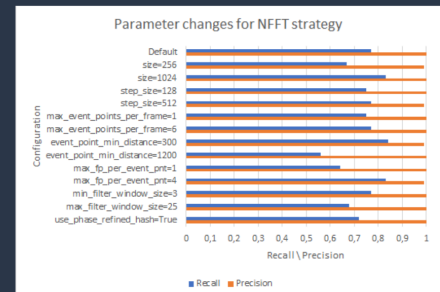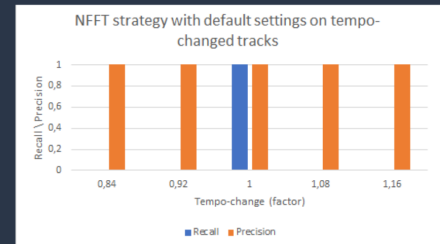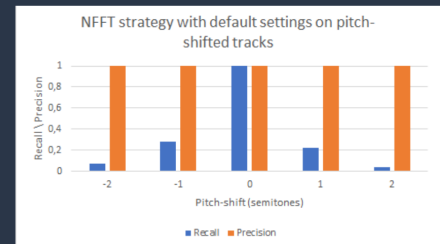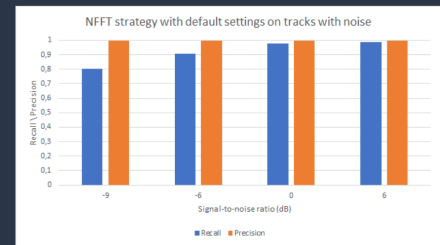
4 RESULTS

Constant Q

Default: **8,3%** correct matches

Constant Q strategy with default settings on tracks with noise

Constant Q strategy with default settings on pitch-shifted tracks

Constant Q strategy with default settings on tempo-changed tracks

Parameter changes for Constant Q strategy

Configuration	Correct matches
bins_per_octave=72	3,9%
step_size=3072	6,9%
max_freq_delta=533	5,4%
min_freq=1850	9,8%
all combined	5,4%

NFFT

Default: **0%** correct matches

NFFT strategy with default settings on tracks with noise

NFFT strategy with default settings on pitch-shifted tracks

NFFT strategy with default settings on tempo-changed tracks

Parameter changes for NFFT strategy

Configuration	Correct matches
size=1024	0%
event_point_min_distance=300	0%
max_fp_per_event_pnt=4	1,0%
all combined	2,0%

email: R.K.Nair@student.tudelft.nl