ASSESSING THE SUITABILITY OF PANAKO FOR MUSIC IDENTIFICATION IN MOVIES

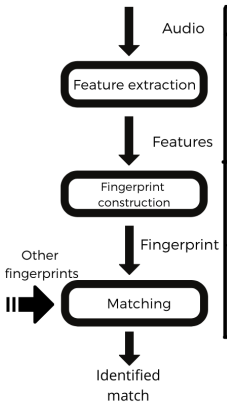
BY RUBEN NAIR
SUPERVISORS: DR. CYNTHIA LIEM, DR. JAEHUN KIM

TU Delft

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 recognizing music in movies, based on the benchmark?
2. What is the influence of configurable parameters on the performance of Panako?

2 PANAKO



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

Figure 1: High level flowchart of functionality Panako

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.** Original soundtracks (using soundtracks of movies provided by Muziekweb) were layered with different types of noise over it, at four different signal-to-noise ratios (SNR). 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 again tested on the real movie clips from the first step.

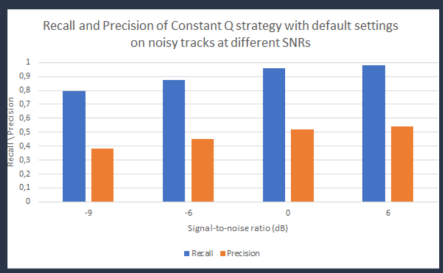
5 CONCLUSION

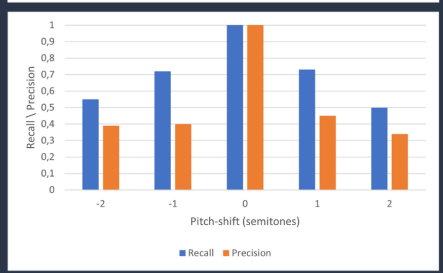
-

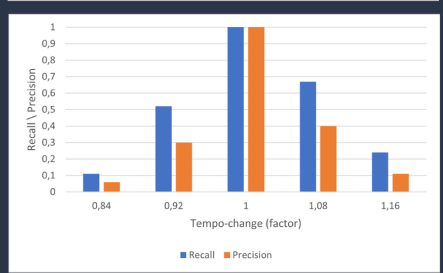
4 RESULTS

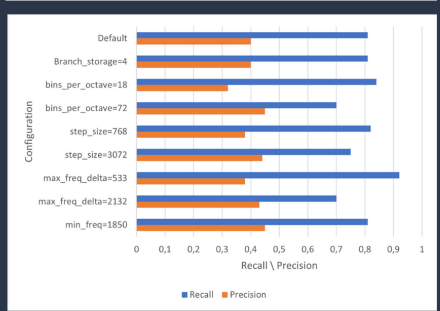
Constant Q

Default: 8,3% correct matches





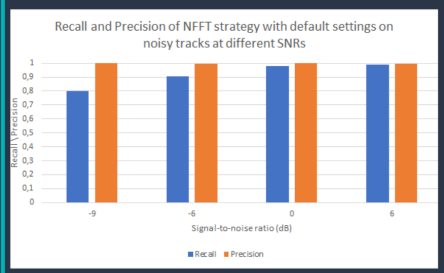


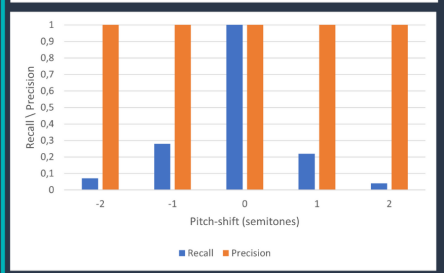


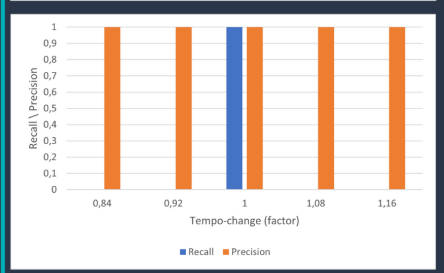
Configuration	Correct matches
bins_per_octave=72	3,9%
step_size=3072	6,9%
max_freq_delta=533	5,4%
min_freq=1850	9,8%
all combined	5,4%

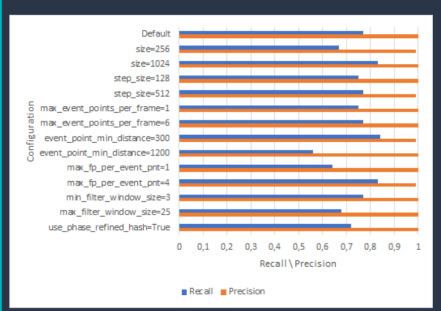
NFFT

Default: 0% correct matches









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