Computational Auditory Scene Analysis (CASA) for Separating Monophonic Music

Bachelor's Thesis

Nikita Mortuzaiev

June 22nd, 2022

Supervisor: Ing. Mgr. Ladislava Smítková Janků, Ph.D.

Objectives

- Research the field of CASA, its applications and goals

- Study existing works

- Implement a CASA system to process monophonic piano music

- Experiment with the implemented system

Auditory Scene Analysis

- The term was coined by Albert Bregman in 1990

- ASA is described as the process of integrating "auditory objects" into meaningful streams

- In simple words, it relates to the human ability to separate sounds

- Bregman made a lot of experiments to emphasize simultaneous and sequential grouping

Computational ASA

Relates to the
"cocktail party effect"

Tries to mimic
the mechanisms
in the human ear

CASA is a study of
auditory scene analysis
by computational means

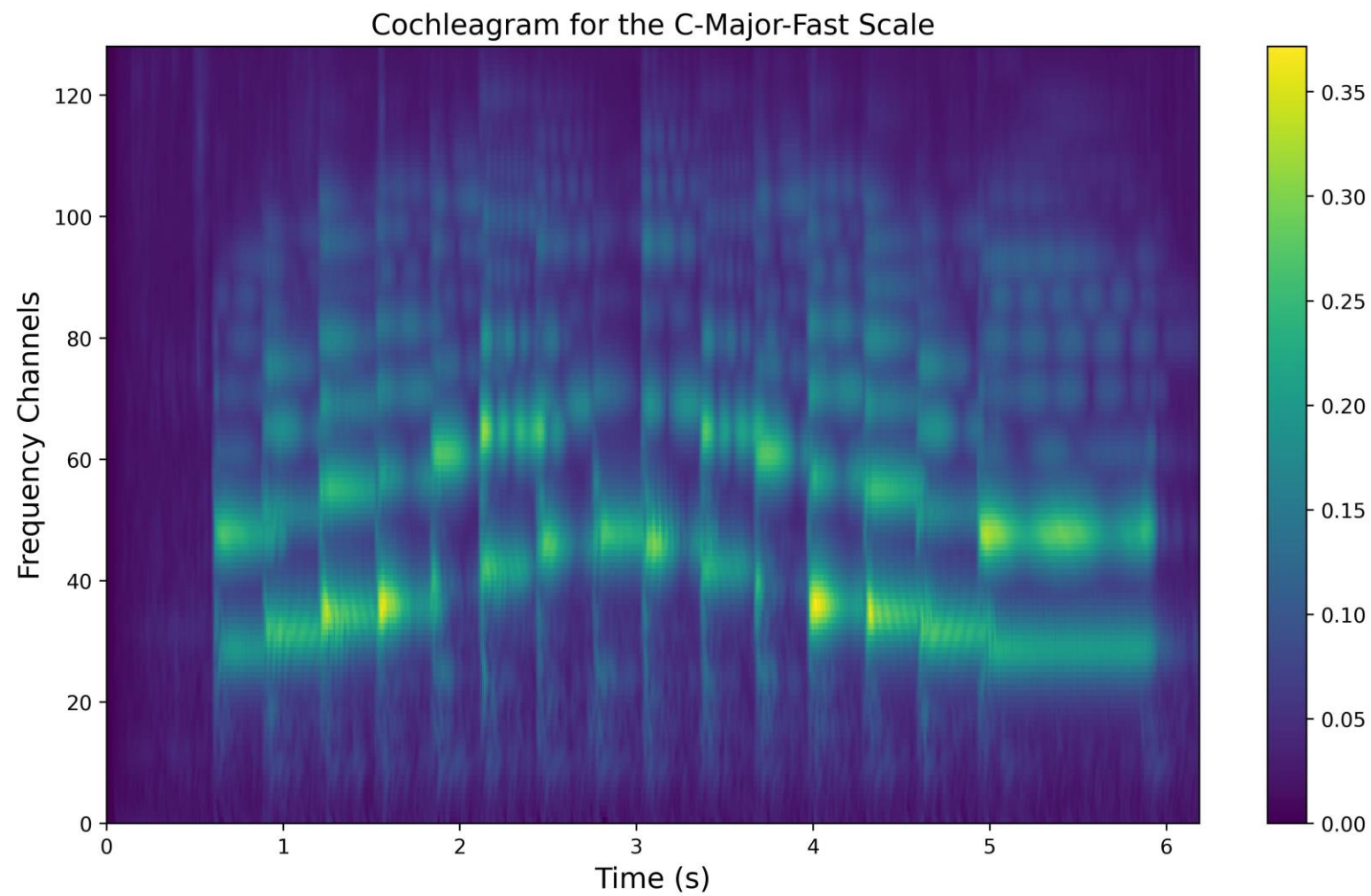Aims to separate sounds
from mixtures

The goal is to find an
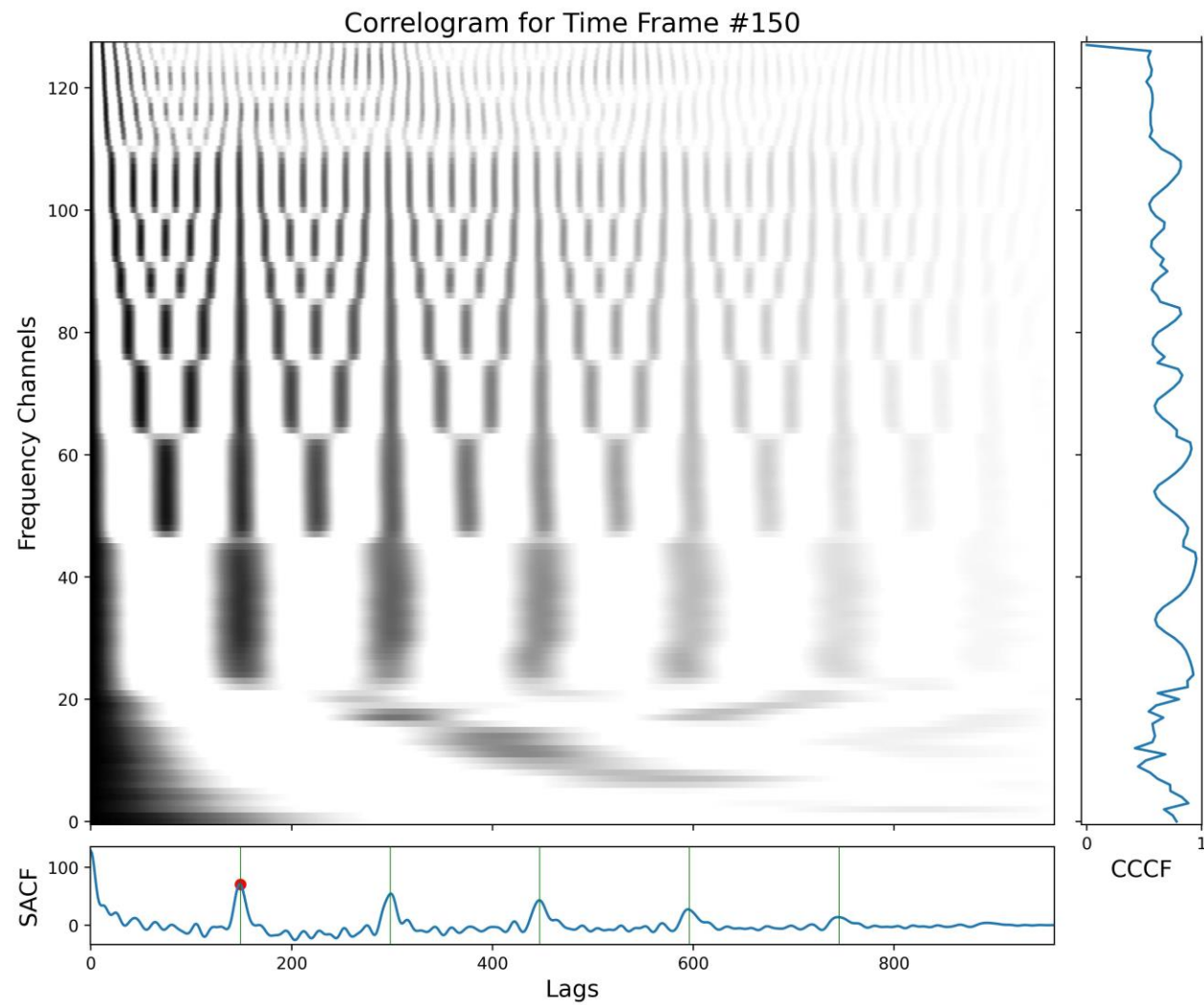"ideal binary mask"

Typical Architecture

- Peripheral Analysis

- Feature Extraction

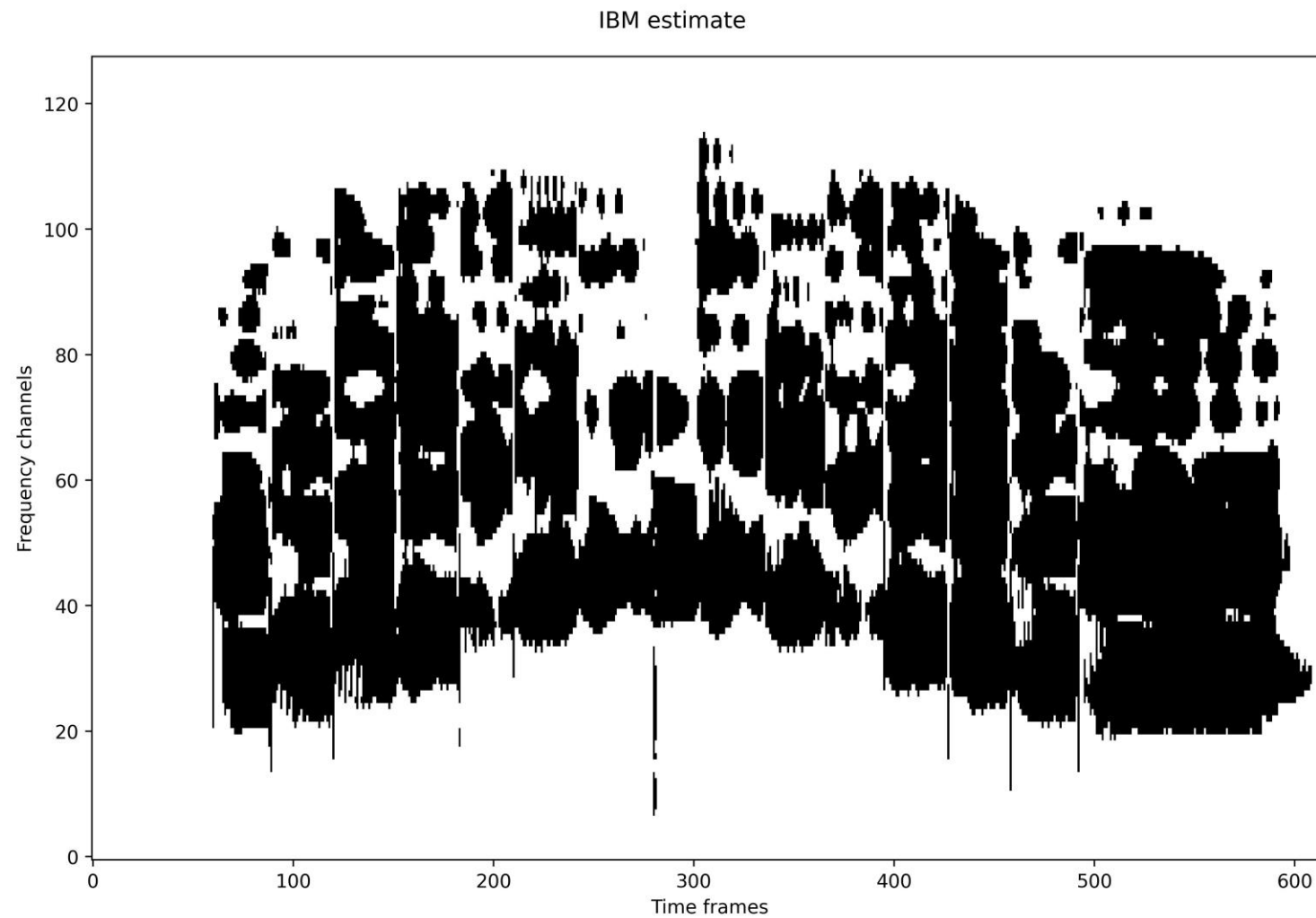- Mid-Level Representations
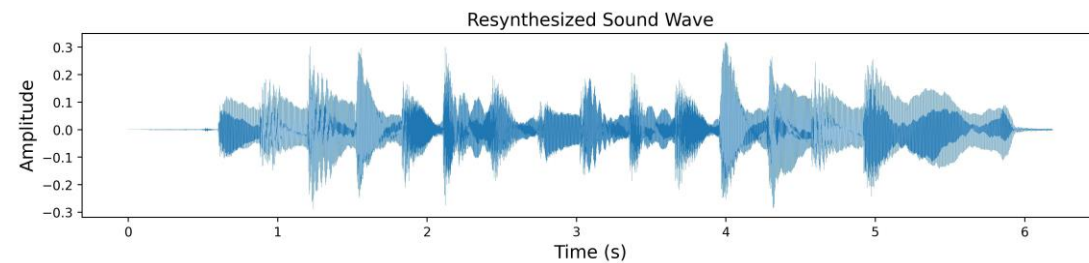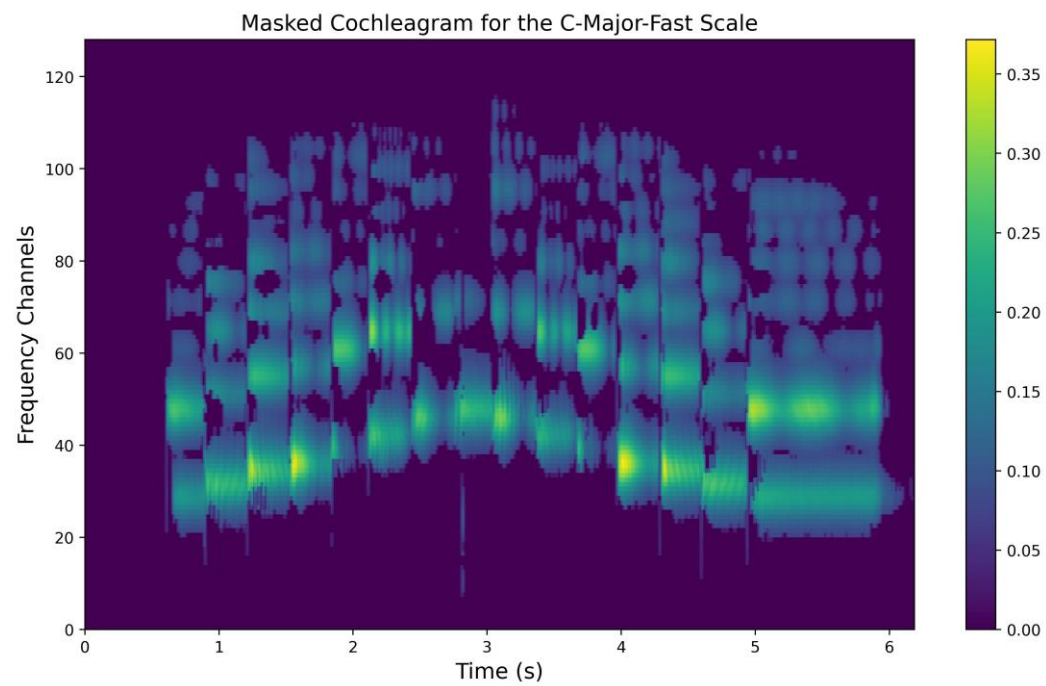
- Scene Organization

- Resynthesis

Cochleagram

Cochleagram for the C-Major-Fast Scale

Feature Extraction

Correlogram for Time Frame #150

Segmentation and Grouping

IBM estimate

Resynthesis

Masked Cochleagram for the C-Major-Fast Scale

Resynthesized Sound Wave

Experiments

- White noise backgrounds

- Other prerecorded backgrounds

- Tests in connection with a simple classifier

- Other experiments

Dataset

- 34 piano recordings stored as two channels in WAV files

- Various major and minor scales in different modes

- Perfect melodic fourths and octaves

- All piano keys one after another in different octaves

- 36 background sounds: clanging, grinding, rattling, etc.

- Rustling paper or a plastic bag, ticking lock, clatter of kitchen utensils...

Experiments with White Noise and Other Backgrounds

Experiments with a Simple Classifier
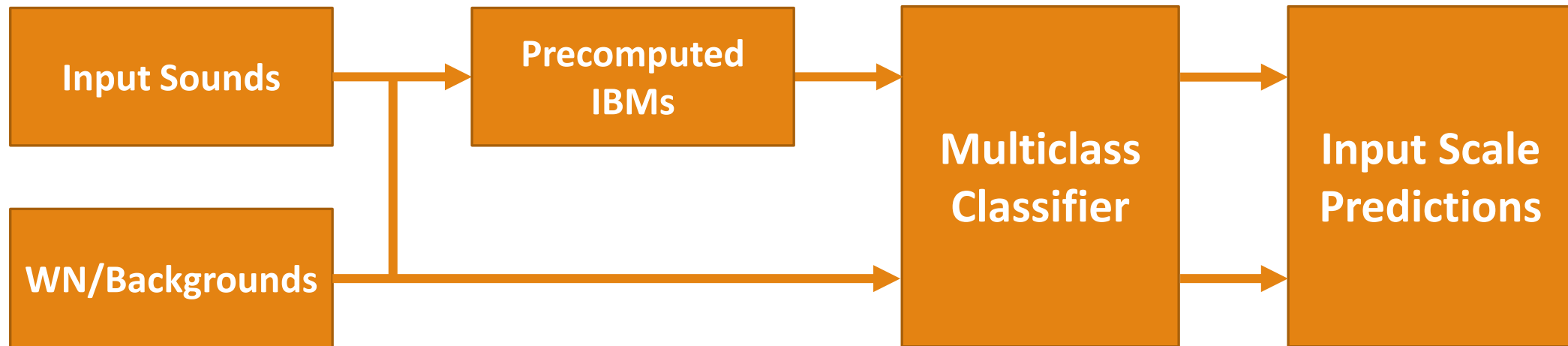
Thank You for Your Attention!

Opponent's Questions

- Which further extensions/improvements would you suggest for your model?
  ("*Jaké další možnosti rozšíření byste navrhl pro Váš model?*")

- Where do you see its further use?
  ("*Kde spatřujete jeho další využití?*")