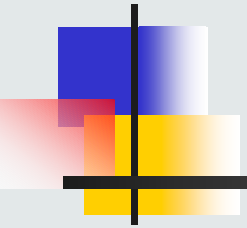
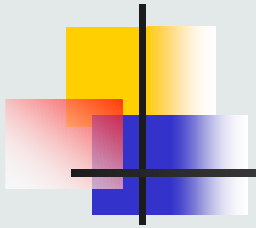


New Feature for Automatic Speech/Music Discrimination

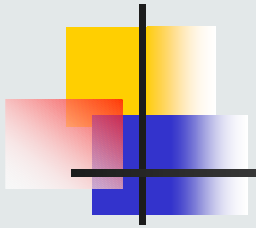
Jayme Garcia Arnal Barbedo

Amauri Lopes



Applications

- Automatic Speech Recognizers
- Automatic Music Transcriptors
- Hearing Devices
- Automatic Radio Selection

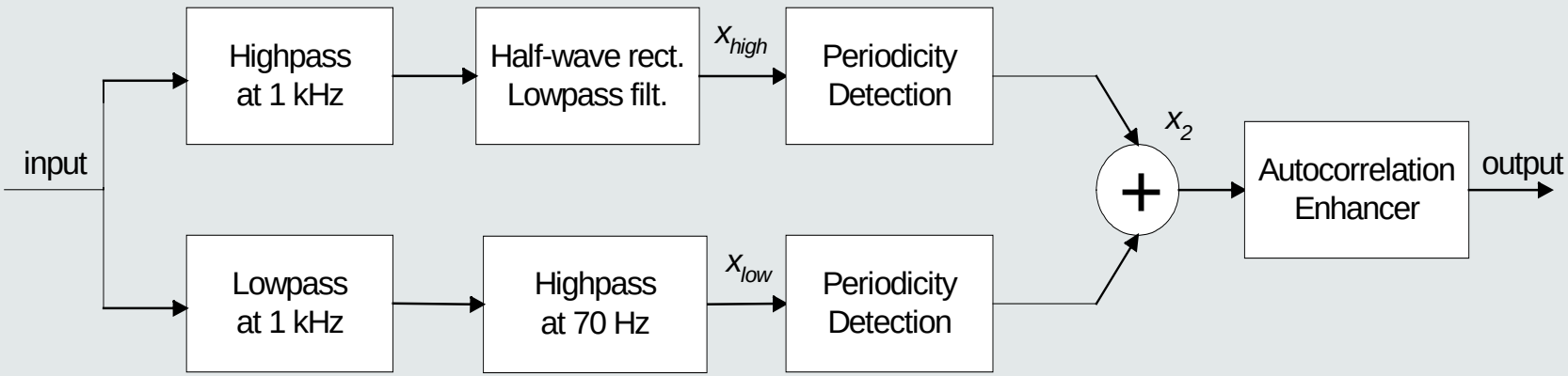


Comparison

- Previous Proposals
  - Large number of parameters
  - Use of complex mapping techniques
- This Work
  - Only one parameter
  - No mapping technique is used

Feature Extraction

- Multipitch Analysis Model



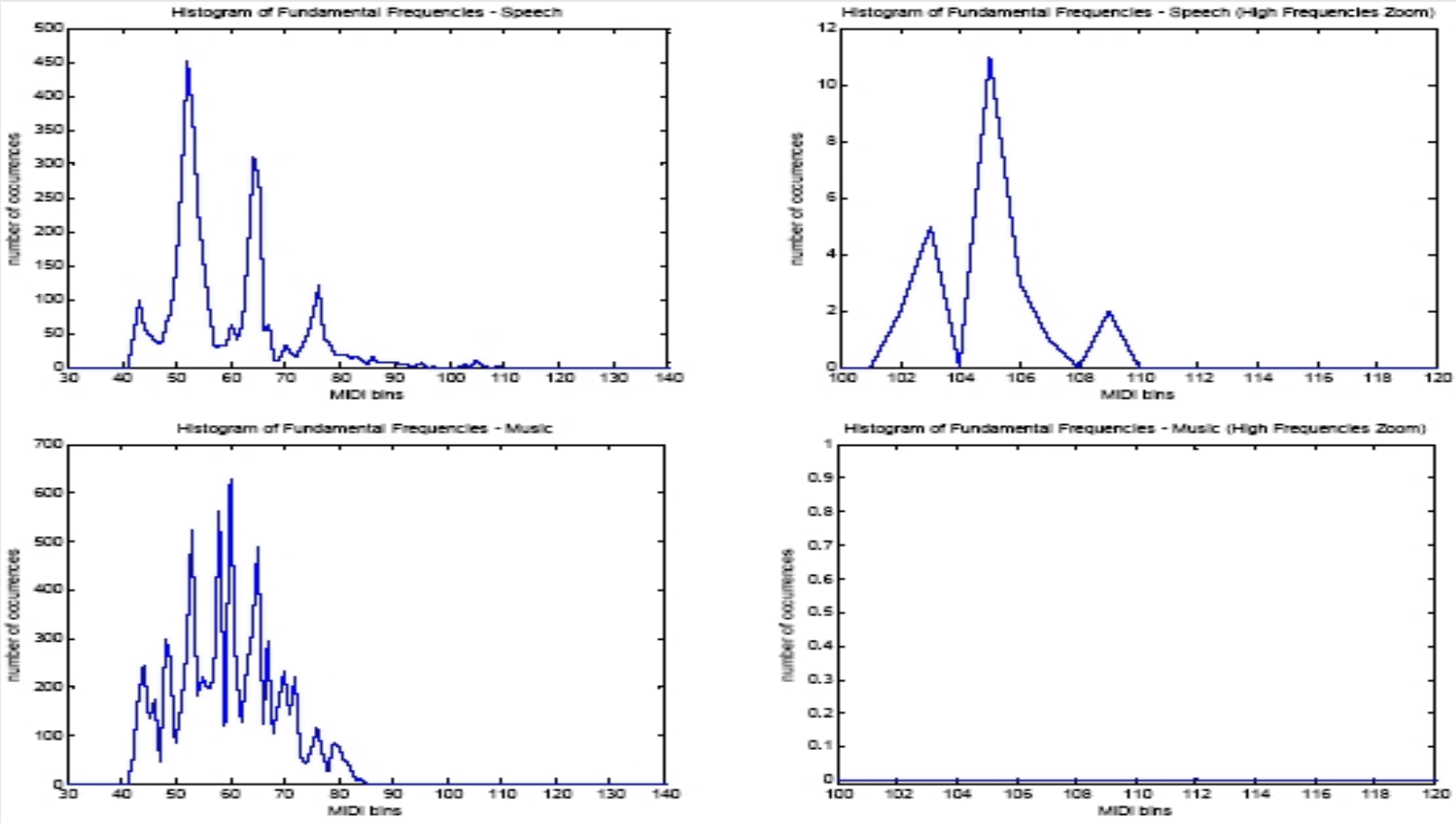
Feature Extraction

- Transformation of f0 into MIDI number:

$$m = 12 \log_2 \left(f / 440 \right) + 69$$

Feature Extraction

- Construction of histogram



Interpretation and Results

- If $n \geq 100$ for more than 0.1% of the cases, the signal is classified as speech, otherwise is classified as music.
- Results

Group	Accuracy
Speech (all files)	94.01%
Speech (only without environmental noise)	96.05%
Music (all files)	93.63%
Music (without rap files)	94.87%

Conclusions and Future Work

- Good performance
- Low computational effort
- Easy to implement
- Robust to unknown conditions and noisy signals
- To be used as part of more sophisticated programs in the future.