CHORD DETECTION USING DEEP LEARNING

Xinquan Zhou, Alexander Lerch

Georgia Institute of Technology Center for Music Technology

OVERVIEW

- Chord detection can contribute to applications such as music similarity measures, sematic analysis and etc.
- Several studies indicate that deep learning methods can be very successful when applied to Music Information Retrieval (MIR) tasks, especially when used for feature learning
- Exploration of Deep Learning for chord detection

PROPOSED METHOD

- Investigated Deep Learning for learning high-level features for chord detection
- Employed time splicing and convolution methods for pre-processing
- Applied Hidden Markov Models (HMMs) to features that Deep Learning extracts for decoding the final chord sequence

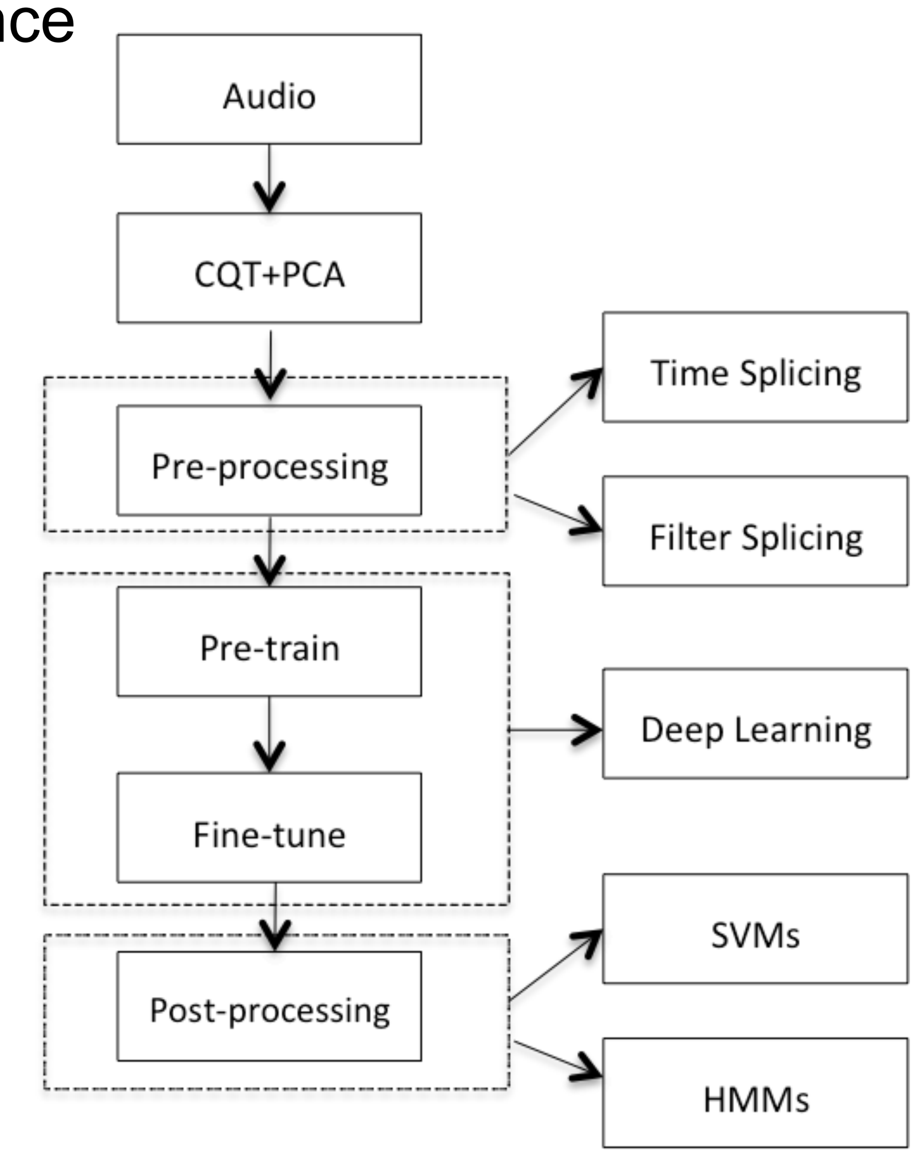

Fig 2. Illustration of the proposed method

WHAT IS CHORD DETECTION?

- Extraction of chord annotation of music from audio signals

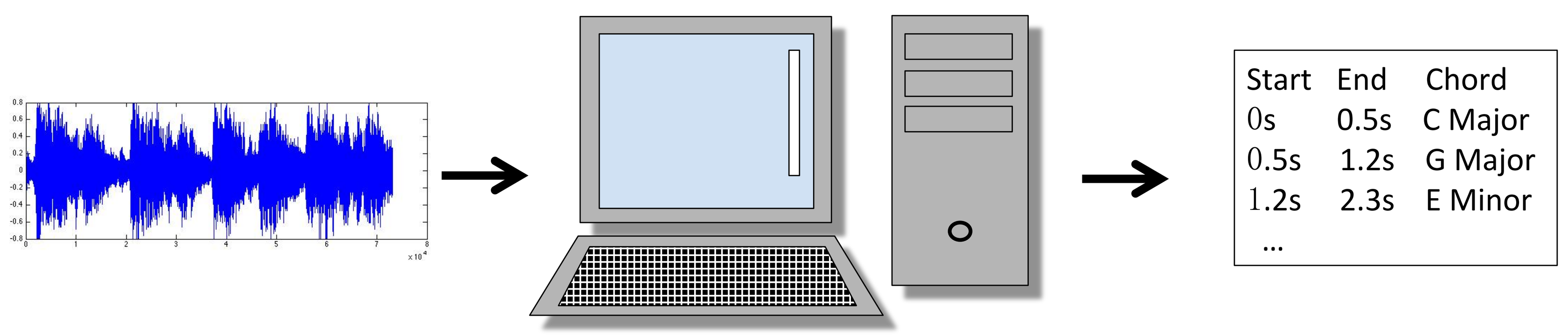

Fig 1. Overview of a general chord detection

CONCLUSION

- Our model is able to learn high-level probabilistic representations for chords detection
- The choice of appropriate input filtering can significantly affect the system performance.
- HMMs or Viterbi decoding algorithm fits the task better than the static classifiers
  - The use of a bottleneck architecture is advantageous.

Future Work

- Learning a pitch class vector instead of chord likelihood by incorporating multi-label
- Finding a way to design the preprocessing filters systematically.

Reference

[1] Philippe Hamel and Douglas Eck. "Learning features from music audio with deep belief networks." In ISMIR, pages 339–344. Utrecht, The Netherlands, 2010.

[2] Frantisek Grezl, Martin Karafi´at, Stanislav Kont´ar, and J Cernocky. "Probabilistic and bottleneck features for lvcsr of meetings." In Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP) , volume 4, pages IV–757. IEEE, 2007.

[3] Eric J Humphrey and Juan Pablo Bello. "Rethinking automatic chord recognition with convolutional neural networks." In Proceedings of the International Conference on Machine Learning and Applications (ICMLA) , volume 2, pages 357–362. IEEE, 2012.

RESULTS

- First, three different post classifiers are compared: the maximum of the softmax output (Argmax), an SVM, and an HMM.

Classifier	Argmax	SVMs	HMMs
WCSR	0.648	0.645	0.755

Table 1. Evaluation results of Weighted Chord Symbol Recall (WCSR)

- Selecting HMMs as post classifiers, the performance of both common architecture and bottleneck one is evaluated.

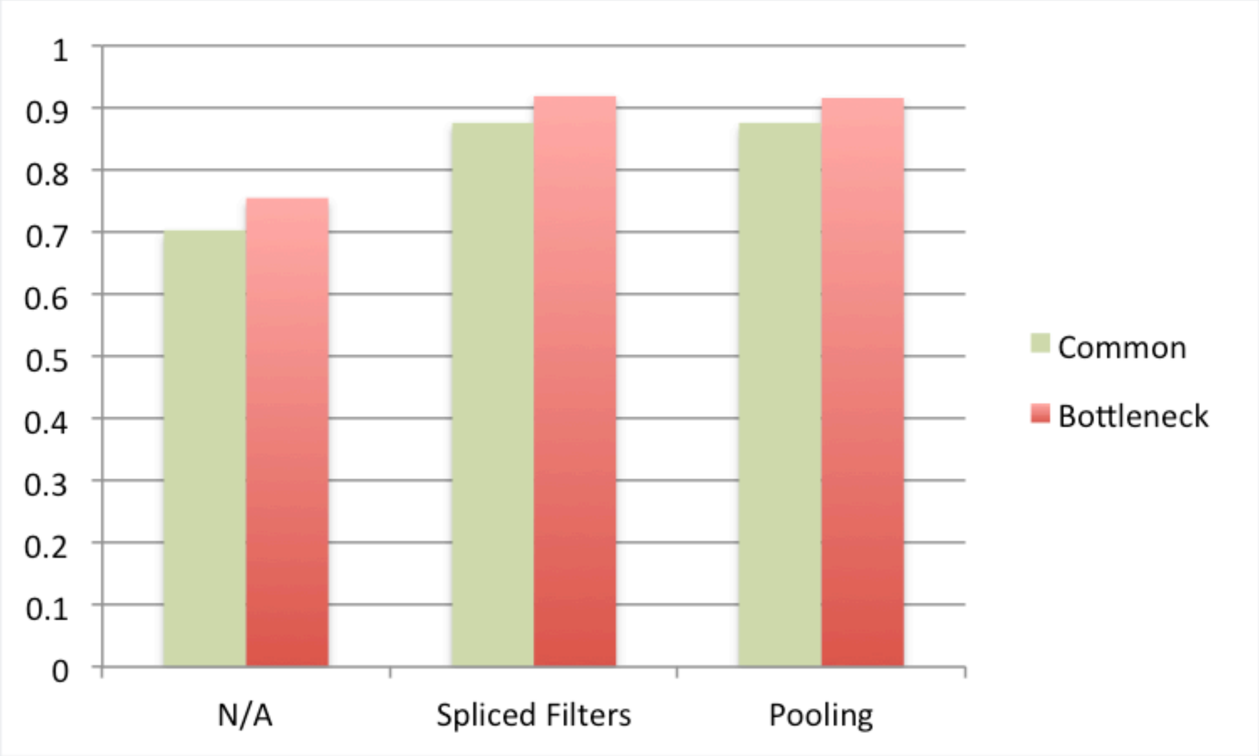

Fig 3. Results of different architectures

- Finally, we present the results of Chordino with the default settings, computed on our dataset and compare it with our system.

Method	WCSR
Chordino	0.625
Proposed	0.919

Table 2. Final results

- Our dataset is a 317-piece collection composed of Beatles, RWC pop, Zweieck and Queens.

Georgia Tech | Center for Music Technology
College of Architecture

For more information, please contact royzxq@gmail.com

CHALLENGES

- Chord is changing along with time, such time information is significant
- First exploration using deep neural networks on chord detection
- Trade-off between accuracy and flexibility