Methods for Polyphonic Music Transcription

Jeremy Nash[a], Mark Liu[a], Paul Schroeder[a]

[a]*Electrical Engineering Department, University of Michigan, Ann Arbor, MI 48109*

Abstract

This paper presents a review of modern methods for polyphonic music transcription. NNN

Keywords: Transcription, sparse coding, non-negative matrix factorization, Bayesian non-parametrics, music information retrieval

1. Motivation

2. Problem Statement

3. Related Work

4. Methodologies

The two main methodologies in modern polyphonic music transcription reflect the supervised/unsupervised learning dichotomy in machine learning. The first approach learns separate discriminative models for the presence of a note over spectral features.

4.1. Spectrogram
4.2. Constant Q transform
4.3. Semitone filter bank
4.4. Discriminative models
4.5. SVM
4.6. Feed forward neural network
4.7. Non-negative matrix factorization
4.8. Bayesian nonparametric models
4.9. Smoothing
4.9.1. HMM smoothing
4.9.2. Probabilistic spectral smoothness
4.10. Recurrent neural network

5. Evaluation

6. Conclusion

7. Individual Effort

- **Jeremy** wrote and evaluted the SVM method in [1], the Bayesian nonparametric method in [2], and the sparse non-negative matrix factorization method in [3]. Jeremy also wrote the paper.

- **Mark** implemented the hidden Markov model in [1] and the LSTM network in [4].

- **Paul** implemented the LSTM network in [4].

References

[1] G. E. Poliner, D. P. Ellis, A discriminative model for polyphonic piano transcription, EURASIP Journal on Advances in Signal Processing 2007 (2006).

[2] D. M. Blei, P. R. Cook, M. Hoffman, Bayesian nonparametric matrix factorization for recorded music, in: Proceedings of the 27th International Conference on Machine Learning (ICML-10), pp. 439–446.

[3] S. A. Abdallah, M. D. Plumbley, Polyphonic music transcription by non-negative sparse coding of power spectra, in: Proc. 5th Intl Conf. on Music Information Retrieval (ISMIR), pp. 10–14.

[4] S. Bock, M. Schedl, Polyphonic piano note transcription with recurrent neural networks, in: Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on, IEEE, pp. 121–124.