Piano Transcription

MUMT611

Presentation III

1 March, 2007

Outline

- Introduction
- Techniques
  - Comb Filtering & Autocorrelation
  - HMMs
  - Blackboard Systems & Fuzzy Logic
  - Neural Networks
- Examples
- Bibliography, Discussion

Introduction

"Transcription of Music is defined to be the act of listening to a piece of music and of writing down the musical notation for the sounds that constitute the piece."

Hankinson, 3/15

Introduction

Piano Transcription

- **Polyphonic instrument with significant corpus of music**
- **The Bach chorales "serve as an interesting and useful starting point because they embody a very structured domain of musical practice." (Martin 1996)**
- **A structured domain gives a system a limited amount of options to choose from when faced with ambiguity, which makes problem solving easier.**
- **Goal is to produce a machine-readable (and indexable) format from audio input. (MIDI, CSOUND)**

Introduction

Piano Transcription

- **Several Techniques have been used to attack this problem**
  - **Comb Filters and Autocorrelation (early)**
  - **Blackboard systems**
  - **Hidden Markov Models**
  - **Neural Networks**

Techniques

Piano Transcription

- **Moorer (1975) was the first to attempt polyphonic music transcription**
  - **Used comb filtering and autocorrelation techniques**
  - **Limited to two voices, differing in timbre**
  - **Limited amount of intervals (No octaves as they contained similar harmonic qualities)**
  - **Limited range of two octaves**

Techniques

Piano Transcription

- **Raphael (2002) uses HMMs**
  - **Uses Mozart Sonata 18 K.570**
  - **Restricted to C_2 to F_6, and chords with four or less notes**
  - **HMM trained on data taken from other Mozart Piano Sonatas**
  - **Note error rate of 39% out of 1360 notes. (530 'wrong' notes)**

Techniques

Piano Transcription

- Martin (1996) uses a blackboard approach
  - Non-serial method of processing
  - Uses 'expert' modules to solve problems
  - Limited to a 18th century counterpoint
  - Failure to detect octaves (a common problem)
  - Flexibility of the blackboard system is a bonus (additional 'experts' can be added to the process to assist)
  - No musical knowledge in this system (i.e. tonality) but it could be added to assist in the process.

Techniques

Piano Transcription

- **Bello & Sandler (2000) also use a blackboard approach**
  - **Employed neural network techniques in their model to adjust the note hypothesis. (a 'learning' model)**
  - **Also had problems with octave recognition**

Techniques

Piano Transcription

- **Marolt (2004, 2005) uses neural network techniques**
  - **Uses connectionist approach (interconnected simple units that change over time) rather than a computational approach (pre-set rules applied)**
  - **Networks trained to recognize notes that are passed to them from other networks**
  - **Also has octave errors & problems with repeated notes**
  - **Produced a system called SONIC that uses the neural network techniques**

Methods

- **Three piano pieces selected from my own music collection: Pop, Jazz & Baroque. Started as 160Kbps AAC or MP3 files.**
- **Chopped to 12s or 30s clips in Quicktime Pro, and then exported to mono wave, 44Khz 16bit.**
- **Processed through SONIC on an Intel Macintosh running Darwine (for Windows compatibility)**
- **No extra parameters were added for tuning**

SONIC in Practice

	Pop	Jazz	‘Baroque’
Original	Fake Plastic Trees	Scattin' at the Kit Kat	English Suite
Transcribed	Fake Plastic Trees	Scattin' at the Kit Kat	English Suite

Hankinson, 12/15

SONIC in Practice

	Pop	Jazz	‘Baroque’
Original	Fake Plastic Trees	Scattin' at the Kit Kat	English Suite
Transcribed	Fake Plastic Trees	Scattin' at the Kit Kat	English Suite

Hankinson, 12/15

SONIC in Practice

	Pop	Jazz	‘Baroque’
Original	Fake Plastic Trees	Scattin' at the Kit Kat	English Suite
Transcribed	Fake Plastic Trees	Scattin' at the Kit Kat	English Suite

Hankinson, 12/15

SONIC in Practice

	Pop	Jazz	‘Baroque’
Original	Fake Plastic Trees	Scattin' at the Kit Kat	English Suite
Transcribed	Fake Plastic Trees	Scattin' at the Kit Kat	English Suite

Hankinson, 12/15

SONIC in Practice

	Pop	Jazz	‘Baroque’
Original	Fake Plastic Trees	Scattin' at the Kit Kat	English Suite
Transcribed	Fake Plastic Trees	Scattin' at the Kit Kat	English Suite

Hankinson, 12/15

SONIC in Practice

	Pop	Jazz	‘Baroque’
Original	Fake Plastic Trees	Scattin' at the Kit Kat	English Suite
Transcribed	Fake Plastic Trees	Scattin' at the Kit Kat	English Suite

Hankinson, 12/15

SONIC in Practice

	Pop	Jazz	‘Baroque’
Original	Fake Plastic Trees	Scattin' at the Kit Kat	English Suite
Transcribed	Fake Plastic Trees	Scattin' at the Kit Kat	English Suite

Hankinson, 12/15

SONIC in Practice

	Pop	Jazz	‘Baroque’
Original	Fake Plastic Trees	Scattin' at the Kit Kat	English Suite
Transcribed	Fake Plastic Trees	Scattin' at the Kit Kat	English Suite

Hankinson, 12/15

Thoughts

- The Radiohead piece had many notes below a dynamic threshold, and the system did not recognize it
- Octave errors are apparent in all pieces
- Handled Harpsichord with no problem (not just for piano)

Conclusion

- **Monotonic is solved; polyphonic is much harder**
- **Systems are good, and getting better**
- **Movement towards a more 'humanized' approach to machine learning (Computers learn the way we learn)**
- **Still a long way to go (dynamics, ornaments)**

Bibliography

Bello, Juan Pablo, Guiliano Monti, and Mark Sandler. 2000. Techniques for Automatic Music Transcription. In *Proceedings of the First International Conference on Music Information Retrieval (ISMIR)*, Plymouth, Massachusetts.

Bello, Juan Pablo, and Mark Sandler. 2000. Blackboard System and Top-Down Processing for the Transcription of Simple Polyphonic Music. In *Proceedings of the COST G-6 Conference on Digital Audio Effects (DAFX-00)*, Verona, Italy.

Dixon, Simon. 2000. On the Computer Recognition of Solo Piano Music. In *Proceedings of the Australasian Computer Music Association Conference*, Brisbane, Australia.

Klapuri, Anssi, Tuomas Virtanen, Antti Eronen, and Jarno Seppanen. 2001. Automatic Transcription of Musical Recordings. In *Proceedings of the Consistent & Reliable Cues Workshop, CRAC-01*, Aalborg, Denmark.

Marolt, Matija. 2005. A Connectionist Model of Finding Partial Groups in Music Recordings With Application to Music Transcription. In *Proceedings of the Seventh International Conference on Adaptive and Natural Computing Algorithms*, Coimbra, Portuagal.

Marolt, Matija. 2004. A Connectionist Approach to Automatic Transcription of Polyphonic Piano Music. *IEEE Transactions on Multimedia* 6 (3): 494-95.

Martin, Keith D. 1996. *A Blackboard System for Automatic Transcription of Simple Polyphonic Music*. *Mit Media Laboratory Perceptual Computing Section Technical Report No. 385.*

Monti, Guiliano, and Mark Sandler. 2002. Automatic Polyphonic Piano Note Extraction Using Fuzzy Logic in a Blackboard System. In *Proceedings of the 5th International Conference on Digital Audio Effects (DAFx-02)*, Hamburg, Germany.

Poliner, Graham E., and Daniel P. W. Ellis. 2007. A Discriminative Model for Polyphonic Piano Transcription. *EURASIP Journal on Advances in Signal Processing* 2007

Poliner, Graham E., and Daniel P. W. Ellis. 2005. A Classification Approach to Melody Transcription. In *Proceedings of the Sixth Annual Conference on Music Information Retrieval (ISMIR)*, London, UK.

Raphael, Christopher. 2002. Automatic Transcription of Piano Music. In *Proceedings of the Third International Conference on Music Information Retrieval*, Paris, France.