Automatic Piano Transcription

The piano is seen by some to be a good instrument for starting to tackle the problem of polyphonic music transcription. Martin (1996) chose piano music because it serves

> as an interesting and useful starting point because [it embodies] a very structured domain of musical practice...The importance of a structured domain is that it allows the transcribing agent to exploit the structure, thereby reducing the difficulty of the task.

Polyphonic transcription seeks to take an audio input as a source, and produce a MIDI, CSOUND or similar score that is able to be manipulated by humans and computers. Given this problem, there have been a number of different approaches to solving it with varying results.

Moorer (1975) is widely recognized to have been the first person to attempt polyphonic music transcription through a combination of comb filtering and autocorrelation. His earliest experiments worked with a highly limited number of possibilities (limited to two octaves, no simultaneous octaves sounding, two voices of differing timbre).

Martin (1996) implemented a 'blackboard' system of polyphonic music transcription. His approach used a series of 'expert' modules that are presented with a problem, and each of those experts then tackles the problem independently. For example, in his system, the 'Chord_MissingInterval' uses its musical knowledge to scan the blackboard for potential chord hypothesis and if it detects that a chord hypothesis is missing a required interval (m3, M3, P5) it adds it in. This system worked, but was 'static' - that is, it did not adapt its behaviour depending on given input.

A similar approach was implemented by Bello, Monti and Sandler (2000a, 2000b) with the exception of the inclusion of a neural network chord analyser. This blackboard system differed from Martin's in that instead of using a static 'expert' to recognize chords, the neural network could adjust its hypothesis network so that multiple note hypotheses could survive, therefore building a chord from the input.

Raphael (2002) proposes an approach based on Hidden Markov Models. These HMMs are trained with a signal-score pair, meaning that it can match probable chord hypotheses with actual score data. Using this model, they analysed the 3rd movement of Mozart's Sonata 18, K. 570. They achieved a 'note error rate' of 39% (530 out of 1360 notes were 'wrong.')

Marolt (2004, 2005) brings a neural network approach to the problem of recognition. In his approach, networks of simple oscillators track partials within a signal. In his approach, he uses an auditory model that emulates biological sound reception. The outputs of the 'sound receptors are then passed to networks of adaptive oscillators. Each of these oscillators can track a partial in a given signal, and 76 networks (A1 to C8) of up to ten oscillators are employed to track the complete number of partials in a signal. The outputs of these networks then form a hypothesis for a fundamental note.

Marolt's work has produced a piece of software called SONIC. This software is downloadable from his website. The results of running a series of input samples through this software is available on my website, http://www.transientstudent.net/academics/mumt-611/transcription/experiment.

With each new generation of methods, the ability to automatically detect and acurately transcribe polyphonic music gets better. There is a movement towards creating systems that more accurately model the human method of processing musical information and creating machines that learn and adapt from exposure to many different sources, rather than running through a series of pre-determined steps.

Piano Transcription 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 systems 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* v. 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 (ISMIR)*, Paris, France.