AMPACT

Automatic Music Performance Analysis and Comparison Toolkit

Johanna Devaney
Ohio State University

Introduction

Introduction

AMPACT Components

Introduction

AMPACT Components

Using AMPACT

Introduction

AMPACT Components

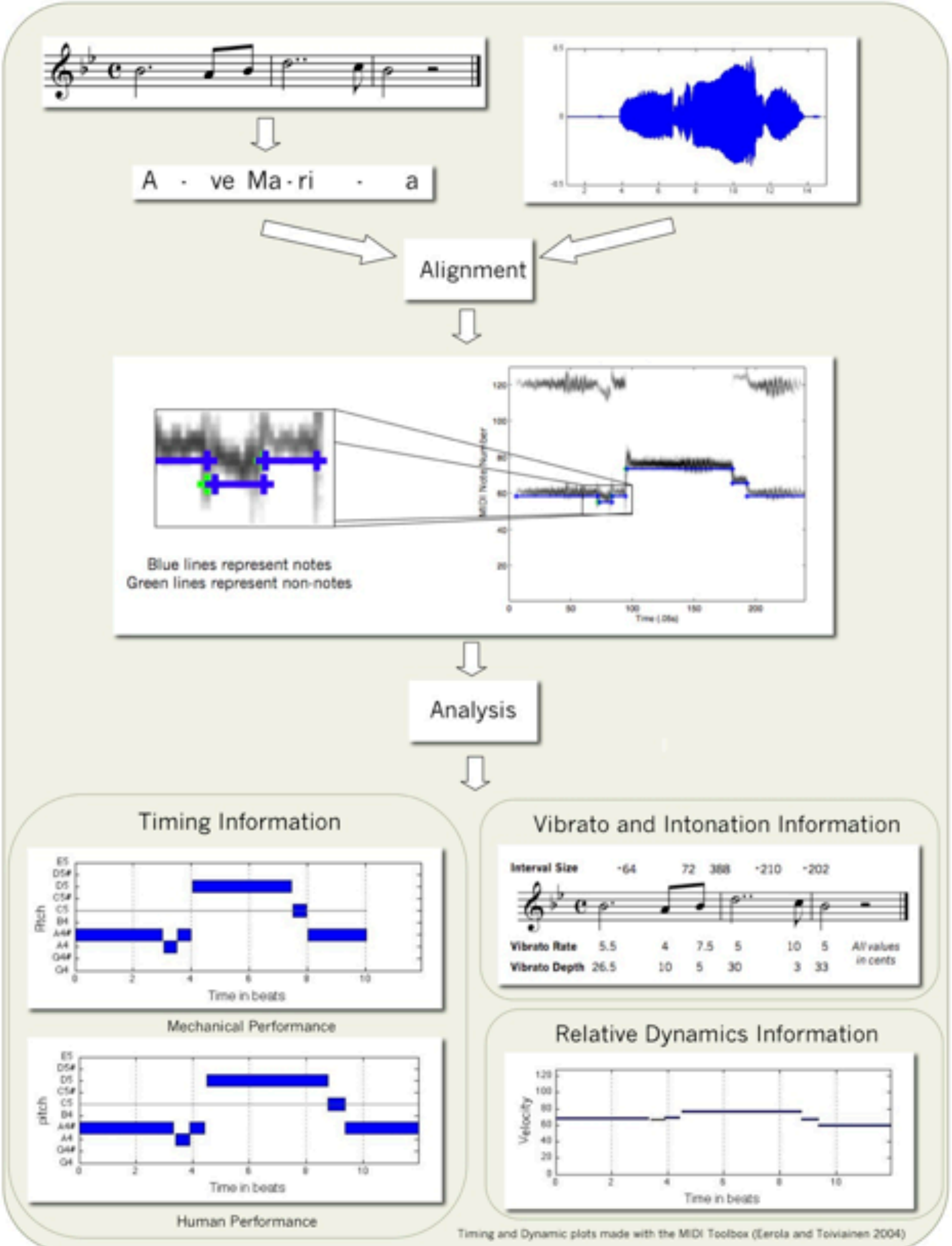
Using AMPACT

Future Directions

Introduction

- AMPACT is a MATLAB-based toolkit for extracting performance data from monophonic recordings for which a score is available
- Performance data extracted includes:
  - timing
  - dynamics
  - intonation
  - vibrato

Components


A - ve Ma - ri - a
Alignment
Blue lines represent notes
Green lines represent non-notes
Analysis
Timing Information
Time in beats
Mechanical Performance
Time in beats
Human Performance
Vibrato and Intonation Information
Interval Size -64 72 388 -210 -202
Vibrato Rate 5.5 4 7.5 5 10 5
Vibrato Depth 26.5 10 5 30 3 33
All values in cents
Relative Dynamics Information
Velocity
Time in beats
Timing and Dynamic plots made with the MIDI Toolbox (Eerola and Toiviainen 2004)

MIDI-Audio Alignment

- Provides more robust estimates of note onsets and offsets that blind detection algorithms
- Algorithm described in Devaney, Mandel, and Ellis 2009
  - An initial dynamic time warping alignment is refined by a hidden Markov model trained on the acoustic features (power, periodicity, pitch) of the instrument (or voice)
  - When aligning voice, lyric information is also used

Performance Data

- Timing (from alignment)
  - inter-onset intervals
  - tempo information
- Dynamics
  - using Glasberg and Moore's loudness model (2002)

Performance Data

- Intonation
  - fundamental frequency estimates using YIN (de Cheveigné & Kawahara 2002)
  - perceived pitch, using Gockel, Moop & Carlyon (2001)
  - interval size calculated as the difference between successive perceived pitch estimates
- Vibrato
  - calculated by finding the dominant frequency of the FFT of the pitch contour

Using AMPACT

- MATLAB code available at ampact.org and github.com/jcdevaney/AMPACT
- Documentation available at ampact.tumblr.com/documentation
- MATLAB toolbox requirements: Signal Processing Toolbox

Using AMPACT

- Required 3rd-party Toolkits
  - de Cheveigné, A. 2002. YIN MATLAB implementation http://audition.ens.fr/adc/sw/yin.zip
  - Ellis, D. P.W. 2003. Dynamic Time Warp (DTW) in Matlab. http://www.ee.columbia.edu/~dpwe/resources/matlab/dtw/
  - Ellis, D. P.W. 2008. Aligning MIDI scores to music audio. http://www.ee.columbia.edu/~dpwe/resources/matlab/alignmidiwav/
  - Murphy, K. 1998. Hidden Markov Model (HMM) Toolbox for Matlab. http://www.cs.ubc.ca/~murphyk/Software/HMM/hmm.html
  - Toiviainen, P. and T. Eerola. 2006. MIDI Toolbox. https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/miditoolbox/

Future Directions

- Adding statistical tools for making comparisons between performances in next version (0.3)
- Improving the robustness of the alignment
- Adding new descriptors
  - e.g., shape and evolution pitch contours (Devaney, Mandel, and Fujinaga 2011)
- Developing an encoding method
- Porting to python

Thank you!

References

de Cheveigné A, Kawahara H (2002) YIN, a fundamental frequency estimator for speech and music. JASA 111 (4):1917–1930

Devaney, J.C., M. I. Mandel, and D. P.W. Ellis. 2009. Improving MIDI-audio alignment with acoustic features. In Proceedings of Workshop on Applications of Signal Processing to Acoustics and Audio, 45–8.

Devaney, J.C., Mandel, M.I., and I. Fujinaga. 2011. Characterizing singing voice fundamental frequency trajectories. In Proceedings of Workshop on Applications of Signal Processing to Acoustics and Audio, 73-76.

Devaney, J. C., M. Mandel, and I. Fujinaga. 2012. A Study of Intonation in Three-Part Singing using the Automatic Music Performance Analysis and Comparison Toolkit (AMPACT). In Proceedings of International Society of Music Information Retrieval Conference, 551–6.

Gockel, H., B. C. J. Moore, and R. P. Carlyon. 2001. Influence of rate of change of frequency on the overall pitch of frequency-modulated tones. Journal of the Audio Engineering Society 109 (2): 701–12.

Glasberg, B. R. and B. C. J. Moore. 2002. A Model of Loudness Applicable to Time-Varying Sounds. Journal of the Audio Engineering Society 50(5): 331–342.