

Evaluation of Real-Time Audio-to-Score Alignment



EC Information Society Technologies
6th PCRD Project IST-026883
<http://www.i-maestro.org>

Arshia Cont
Ircam UMR CNRS 9912
CRCA, UCSD San Diego
cont@ircam.fr

Diemo Schwarz,
Ircam-Centre Pompidou
Paris, UMR CNRS 9912
schwarz@ircam.fr

Norbert Schnell
Ircam-Centre Pompidou
Paris, UMR CNRS 9912
schnell@ircam.fr

Christopher Raphael
Indiana University
Bloomington, IN
craphael@indiana.edu



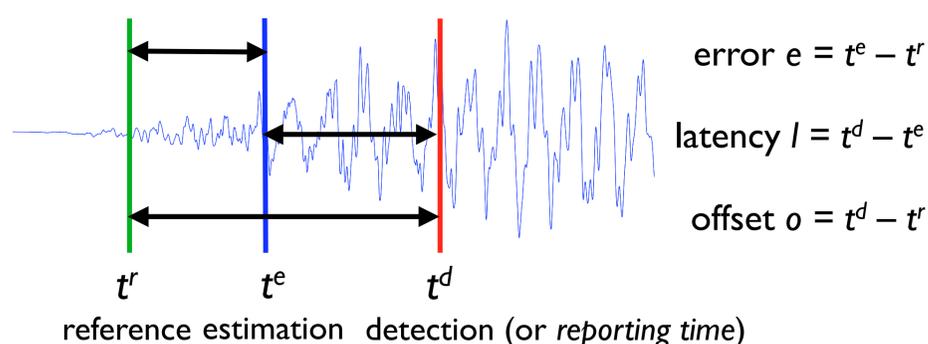
Real-Time Music Interaction
team <http://imtr.ircam.fr>

Score Following

- ▶ real-time alignment of a known musical score to the audio signal produced by a musician playing this score
- ▶ to synchronise an electronic accompaniment to the performer (not the other way round), leaving him with all possibilities of expressive performance.

Measures...

- ▶ taken for each score event (note, trill) on candidate score following system's result output



...and Metrics

- ▶ miss rate p_m , missalign rate p_e ($|e| > \theta$)
- ▶ piece completion p_c (follower doesn't hang)
- ▶ average latency μ_l , average absolute offset μ_o , variance of error σ_e , average imprecision (absolute error) μ_e (of events with $|e| < \theta$)

File Formats

- ▶ see MIREX Wiki http://www.music-ir.org/mirex2007/index.php/Score_File_Format
- ▶ **Score:** text-based (MIDI to be avoided, XML not yet)
- ▶ Example:

event-id	onset-position	onset-ms	type	pitch	interval	duration-beat	duration-ms	cue-num	stream-id
1	1	0	note	72	0	2	4000	1	0
2	3+1/4	4500	note	60	0	0+1/4	500	2	0
3	3+3/4	5000	note	58	0	0+1/2	1000	3	0
4	3+3/4	5000	note	48	0	0+1/2	1000	0	0

- ▶ **Reference:** score file with performance times t^r in onset-time column

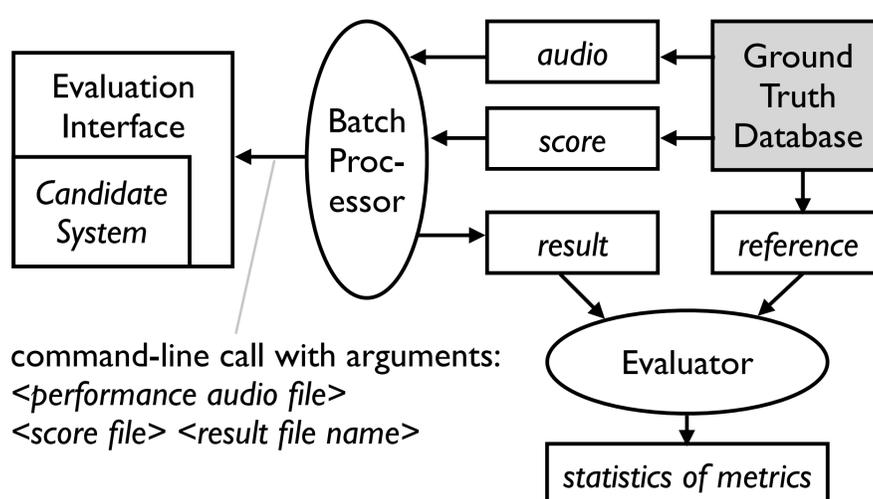
- ▶ **Result:** text matrix
- ▶ Example:

event-id	estimation-time t^e	reporting time t^d	pitch
1	112	152	72
2	5613	5623	60
3	6134	6233	58
4	6144	6233	48

Evaluation

- ▶ gives an indication of the quality of an alignment algorithm, allows comparison of:
 - ▶ different methods, implementations, parameters
 - ▶ improvements gained by training
- ▶ only few attempts in 20 years history of score following
- ▶ lack of unified common framework up to MIREX 2006

Evaluation Framework



Ground-Truth Database

- ▶ should cover wide range of different performances, styles, instruments, ...
- ▶ should be as precise as possible (10s of ms)
- ▶ MIREX 2006 Database:
 - ▶ example at <http://crca.ucsd.edu/arshia/mirex06-scofo>

Piece name	Composer	Instrument	Files	Duration	Events
Explosante-Fixe	Boulez	Flute	47	17:10	2022
Violin Sonatas	Bach	Violin	3	13:50	3996
K. 370	Mozart	Clarinet	4	14:44	2710
Dorabella	Mozart	Voice	4	01:44	229
Total			58	47:38	8957

- ▶ created by fusion of various sources of information:
 - ▶ Basis: off-line alignment system:
 - ▶ good but not perfect
 - ▶ Transient Marks:
 - ▶ many false positives
 - ▶ Spectrogram:
 - ▶ rich information for humans
 - ▶ Audio feedback:
 - ▶ disambiguate hard cases
 - ▶ **All this integrated manually in AudioSculpt**

