Conducting Digitally Stored Music by Computer Vision Tracking

Reinhold Behringer, Leeds Metropolitan University, r.behringer@leedsmet.ac.uk

The Problem
Human musicians in an ensemble synchronise their playing to acoustic (other musicians) and/or visual (conductor) clues.
To enable automatically playing synthesisers in an ensemble with human musicians, a system is required that allows precise synchronisation of the synthesiser’s acoustic output with the human ensemble players.

The Proposed Approach
Use computer vision methods for automatic tracking of conductor motion.
A tempo map is created, based on recognition of beat and measures.
The latency of the processing needs to be considered by extrapolation of the timeline along the pre-stored score data.
As an additional option, an acoustic score following system (e.g. by IRCAM) could be used to achieve precise acoustic synchronisation.

1st International Conference on Automated Production of Cross Media Content for Multi-Channel Distribution, 30.11. – 2.12. 2005, Florence
Preliminary Results

Hand and baton motion has been recorded from a series of manual conducting.

To improve recognition of the baton, a IR-LED can be mounted on the baton tip.

Point tracking of the baton tip results in temporal pattern of positions.

IR LED switch

Battery – could be mounted inside the handle.

Simple prototype of baton with IR-LED tip.

Baton motion during conducting.

Vertical baton motion over time.

Temporal derivative (speed) of vertical baton motion provides 0-transitions and pace of conducting.

Horizontal baton location can provide additional clues.

Related Work (References)


Conclusions

Implementation needs to be done:
- real-time visual tracking in realistic scenario,
- score following using adaptive and recursive estimation,
- extrapolation by “smooth” temporal prediction.