Programming by playing and approaches for expressive robot performances
Angelica Lim, Takeshi Mizumoto, Toru Takahashi, Tetsuya Ogata, and Hiroshi G. Okuno (Kyoto University)

Goal
A robot accompanist with your musical expression.

Background
Our ensemble thereminist robot\(^1\) can respond to a flutist’s cues to change tempo.

But continuously indicating subtle changes and variations is impossible to cue continuously. These minute changes in volume, tempo, pitch, etc. are known as “expression”.

Who would be the perfect musical partner? A player with the same expression as the other player would be in perfect sync.

Challenges
Expressiveness should take into account:
- note length
- intra- and inter-note volume changes
- vibrato
- articulation
- pitch bends
- timbre (if applicable)

Related work
**Expressive flutist robot, Salis et al. 2007**
- Neural network to reproduce a human flutist’s note length and vibrato (flute → flute)

**VocalListener, Nakano and Goto, 2008**
- Synthesized voice with expressive pitch and volume of a human singer (singer → singer)

Programming by playing
It is difficult to describe expression, but easy to demonstrate. We want to allow the human to program the accompanist by playing in their desired style. A continuous pitch and volume representation captures almost all aspects of expression in a piece.

Extraction from recording
1. Robot noise filtering
2. Pitch extraction
3. Power estimation

The robot then generates a (flute → theremin) performance in one of two ways:

1. **Inter-instrument performance transfer** The extracted pitch and power scores are constrained based on player and instrument differences, then played by thereminist robot.

2. **Expression from a score** We use (a) prosody rules such as “the last note in a phrase is longer”, and (b) note volume envelopes based on a human’s articulation example, to suggest emotion

A model for expression
How can we make robots sound more human?
**Based on music psychologist Juslin’s GERMS model\(^2\).**
1. Prosody controller
2. Emotion controller
3. Humanness controller
4. Motor smoothness controller
5. Originality controller

Future Work
Implementing full suite of controllers in GERMS model, and learning parameters for prediction

---

\(^{1}\) Lim et al., “Robot Musical Accompaniment: Integrating Audio and Visual Cues For Real-time Synchronization with a Human Flutist,” IROS, 2010