Visual feedback in a laptop instrument

Michael Quigley

Psychology of Music

Spring 2013

Background

- **Cooperative Interaction** - each sensory mode strengthens the impression of the other (Iwamiya, 2004)
- Sound can influence the temporal perception of motion in a visual field (Bruns and Getzmann 2008)
- Information about a physical property can be represented in an auditory fashion (Effenberg 2005)
- Humans appear to posses some level of perceptual abstraction that allows sensory information to be understood and represented amodally, under certain conditions (Grassi 2005)

Motivation

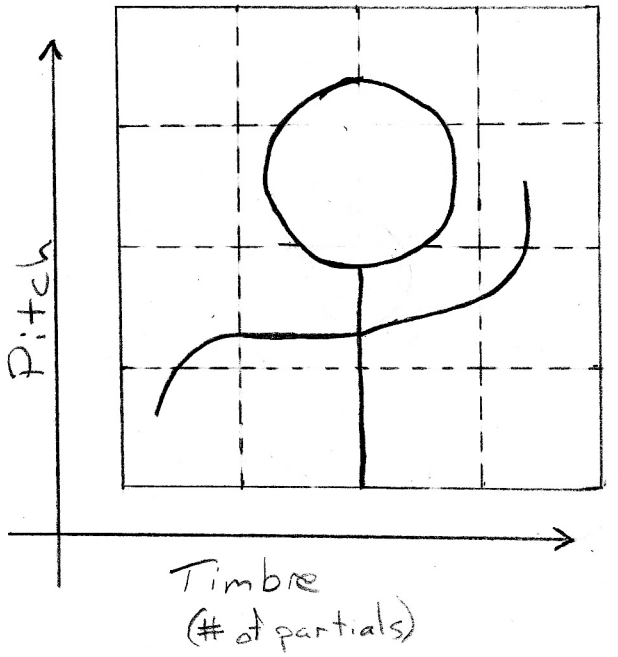
- Non-vocal instruments have been defined by their acoustic and tactile limitations (e.g. trombone, violin, marimba)
- The video camera allows for new forms of interaction with digital music systems!
- But motion-activated music systems often lack the multimodal feedback inherent in traditional instruments

Hypothesis

In a video camera controlled software instrument, the presentation of visual feedback will aid in the understanding of musical mappings to visual space.

Experimental Design

- Max/MSP/Jitter patch that processes the video input of a laptop camera to produce a monophonic tone that changes in pitch and timbre based on motion

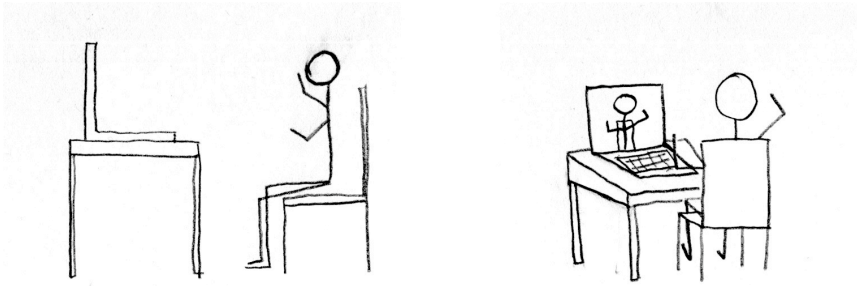


Experimental Design

- A pre-test survey was given to gather background information, used to generate an Ollen Musical Sophistication Index
- Participants were then familiarized with the system
- They were then asked to listen to a recording created using the system and recreate it to the best of their ability, for three specific recordings
- The researcher recorded the time it took to produce the matching changes in pitch or timbre
- A post-test survey was given with 7-point rating questions and short answer questions to analyze the participants' interpretations of the system

Experimental Design

- Those in the experimental group were presented with a mirror image of the video input that was manipulated to help convey changes in musical features

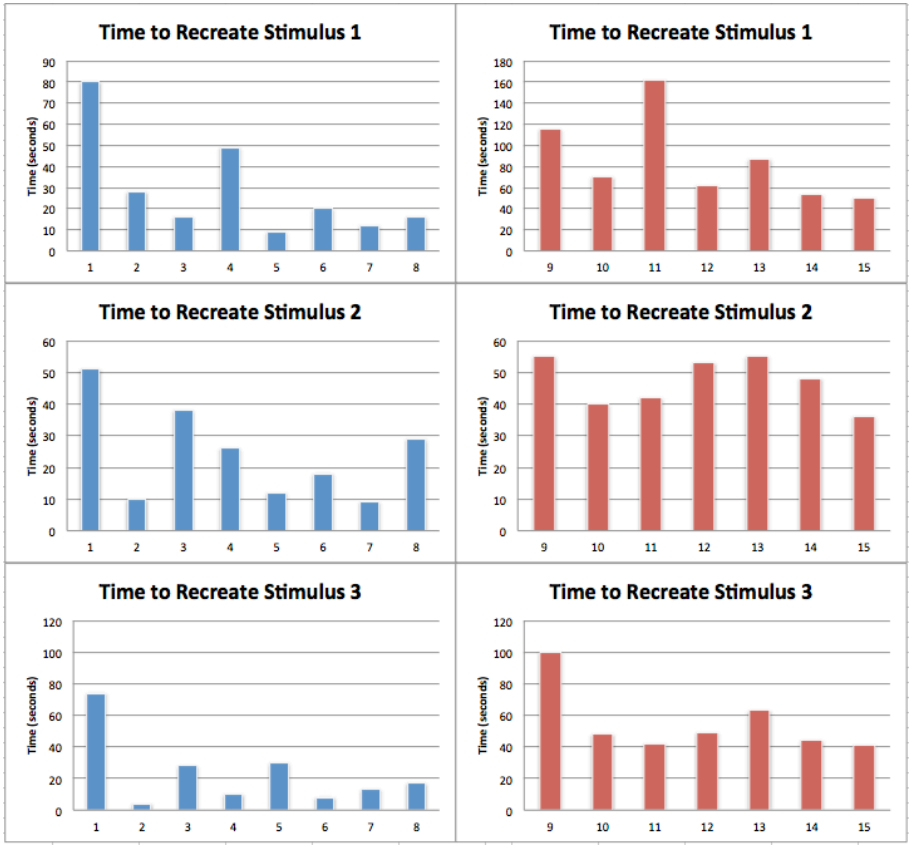


- Increase in pitch → increase in color balance of red
- Increase in timbral complexity → increase in brightness

Results

- *Experimental*
- Age $M = 29.8$, $SD = 8.2$
- OMSI $M = 876.1$, $SD = 139.4$

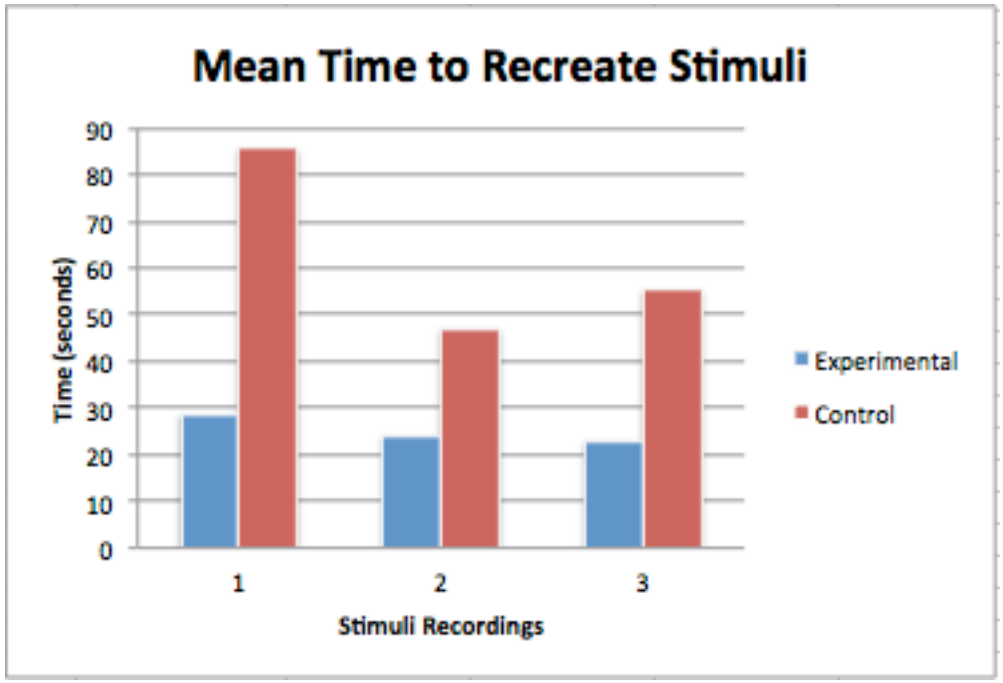
- *Control*
- Age $M = 27.1$, $SD = 5.3$
- OMSI $M = 470.7$, $SD = 419.4$



Stimulus 1: (A 440Hz, sine wave → square wave)
Experimental Recreation Time - $M = 29$, $SD = 24.2$
Control Recreation Time - $M = 86$, $SD = 40.5$

Stimulus 2: (sine wave, A 440Hz → C# 554Hz → E 659Hz → A 880Hz)
Experimental Recreation Time - $M = 24$, $SD = 14.9$
Control Recreation Time - $M = 47$, $SD = 7.7$

Stimulus 3: (square wave, A 440Hz → C# 554Hz → E 659Hz → A 880Hz)
Experimental Recreation Time - $M = 23$, $SD = 22.6$
Control Recreation Time - $M = 55$, $SD = 21.1$



Results: Control and Experimental

- Spatial axes along which changes in pitch and timbre occurred √

- One a 7-point rating scale…
  - Experimental group had higher understanding of the relationship between pitch and its spatial equivalent (M = 6.1, *SD* = 0.8) than those in the control group (M = 3.7, *SD* = 1.4)
  - Also had higher understanding of the relationship between timbre and its spatial equivalent (M = 5.4, *SD* = 1.9) than those in the control group (M = 2.9, *SD* = 0.9)
  - Interestingly, control group rated the appropriateness of the audio-spatial relationships for a musical context higher (M = 5.6, *SD* = 1.3) than those of the experimental group (M = 5, *SD* = 1.2)

Results: Experimental Only…

- 62.5% of participants correctly identified the visual feedback component that corresponded to change in pitch (red tint) and 62.5% for timbre (brightness)

- 50% of participants correctly identified both visual feedback components, while 25% identified one but not the other, and 25% did not identify either visual feedback component

- On a 7-point rating scale…
  - How related the visual display was to the audio output (M = 4.4, *SD* = 1.4)
  - How well he or she understood the pitch-visual relationship (M = 5.1, *SD* = 1.4)
  - Understanding of the timbre-visual relationship (M = 5.1, *SD* = 1.1)
  - Helpfulness of the visual display in understanding how audio parameters were mapped to participants' interaction (M = 3.8, *SD* = 2.3)
  - Appropriateness of the relationship between audio output and visual feedback for a musical setting (M = 4.4, *SD* = 1.5)

Conclusions

- On average, those participants given access of the visual display performed better than those that were not.
- Additionally, they gave higher ratings for their understandings of audio-spatial relationships.
- Results support the hypothesis that the presentation of visual feedback in a video camera controlled software instrument does aid in the users' understanding of musical mappings to visual space.

Conclusions

- However, it is important to note the differences in Ollen Musical Sophistication Indices across groups.

Experimental Median (916.5) and Mean (876.1)

vs. Control Median (222) and Mean (470.7)

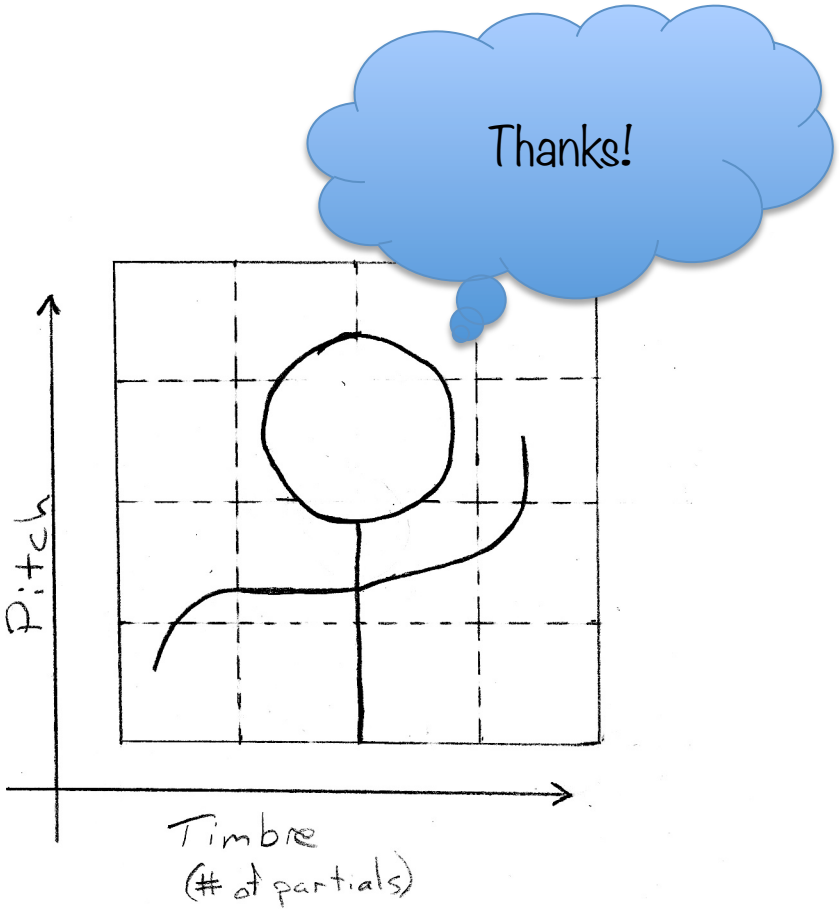
- It is possible that higher OMSI accounts, to some degree, for the faster recreation times of the experimental group

Conclusions

- Still, experimental participants provided additional ratings of the visual components with 50% accurately describing the visual equivalents of pitch and timbre, and 62.5% getting at least one component
- Understanding of audio-visual relationships and appropriateness for musical context were right-leaning in their 7-point ratings

Conclusions

- Further research:
  - usefulness of the particular mappings utilized
    - suggest more intuitive and/or musical mappings
  - An improved study with more evenly distributed OMSIs of participants across the control and experimental groups
  - Several studies could be conducted with only participants above or below a certain index


Thanks!
Pitch
Timbre
(# of partials)