Dynamic Gesture Recognition Using Neural Networks; A Fundament for Advanced Interaction Construction

Klaus Boehm, Wolfgang Broll, Michael Sokolewicz

Presented by Ross Byers

Interaction in Virtual Spaces

- Gestures can provide clear interaction between humans, despite inherent flexibility.
- Gestures provide opportunity for use in Virtual Spaces: Interacting with objects, as well as providing instructions
- Static gestures (postures) established.

Interaction in Virtual Spaces

- Static Gestures are limited.
- Requirement to hold pose tiring
- Exact manipulation is difficult
- No force feedback

Dynamic Gestures

- Natural Interaction is movement based, thus dynamic.
- Time and movement introduce complications.

Recognition of Dynamic Gestures

- The authors used a Kohonen Feature Map (KFM), a type of Neural Network.
- Two layers – Input and output
- Unsupervised training
- Output is a two-dimensional grid of neurons, where spatial proximity on the grid is correlated with similarity.

Preprocessing

- Because of high dimensionality (30+ in this example), the data must be preprocessed.
- 'Vertical' preprocessing collects information for each time step
- 'Horizontal' preprocessing filters and derives data.
- Recording of training data was best assisted by a second person.

First Recognition Approach

- Direct Mapping
- Finds match for the best gesture
- Presents issues with longer and shorter gestures
- Introduces lag
- Requires differing buffer sizes

Second Recognition Approach

- Gesture Parts
- Instead of Requiring ‘all at once’ recognition, recognize a library of sub-gestures.
- Simplest example would use equidistant time-slices. This does not correctly model real behavior.
- KFM used for part recognition only.
- Second, specialized NN used for full gesture recognition.

Results

- Unfortunately, the authors did not share their accuracy data.
- They do however, reflect on processing time.
- The first approach required a reduction of input data to run in real time, and was considered ‘suitable’ for 10 gestures.
- The second approach allows for more preprocessing, thus improving performance even with second neural network.