FROG Recognizer of Gestures

Josh Alvord, Alex Grosso, Jose Marquez Hernandez,
Sneha Popley, Phillip Stromberg, and Ford Wesner
Dr. Donnell Payne (Advisor)
Department of Computer Science

Motivation

- Human-computer interaction is moving away from the traditional mouse and keyboard.
- Speech, gesture, and image recognition are more recent forms of user interaction that are gaining popularity.
- Motion-based gesture recognition is a versatile and intuitive form of interaction.
- **Since many mobile devices have acceleration-based motion sensing capabilities, could a gesture recognition system be developed that allows ANY of these devices to be utilized as the input mechanism?**

Task

- Create an independent gesture recognition framework intended for use in motion-based recognition research
- Allow for training, recognition, evaluation, and demonstration modes of the system (see adjoining image)
- Provide an intuitive, research-oriented GUI that will allow the user to evaluate parameters and algorithm effectiveness
- Implement training and recognition using statistical Hidden Markov Models (HMMs)
- Launch a plug-in for Sun SPOTs while allowing plug-ins for new devices to be created and integrated easily

Dataflow

- Utilizes an established process of manipulating data to perform **training** and **recognition**
- Training: After initial filtering, user input is used to construct a quantizer and HMM for each training set
- Recognition:
  - The filtered input gesture passes through the components constructed in the training pipeline.
  - The gesture is compared against each HMM (trained gesture) after being translated using its corresponding quantizer (K-means or K-means++, in our case).
  - It is then evaluated for recognition using a Bayesian classifier.

Training

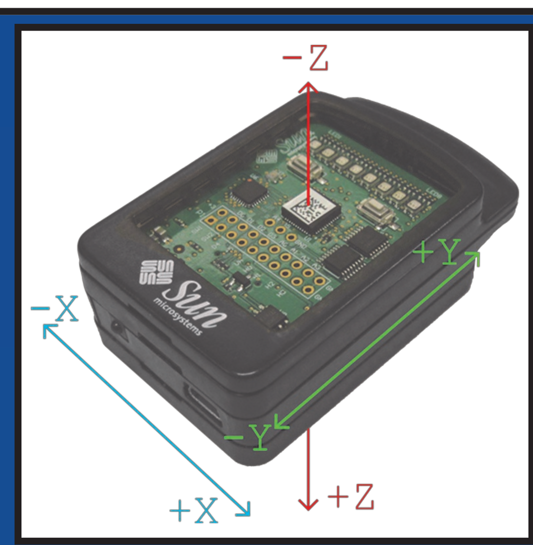
Filtering → Quantizer → Model

Recognition

Filtering → Quantizer → Model → Classifier

Plug-ins

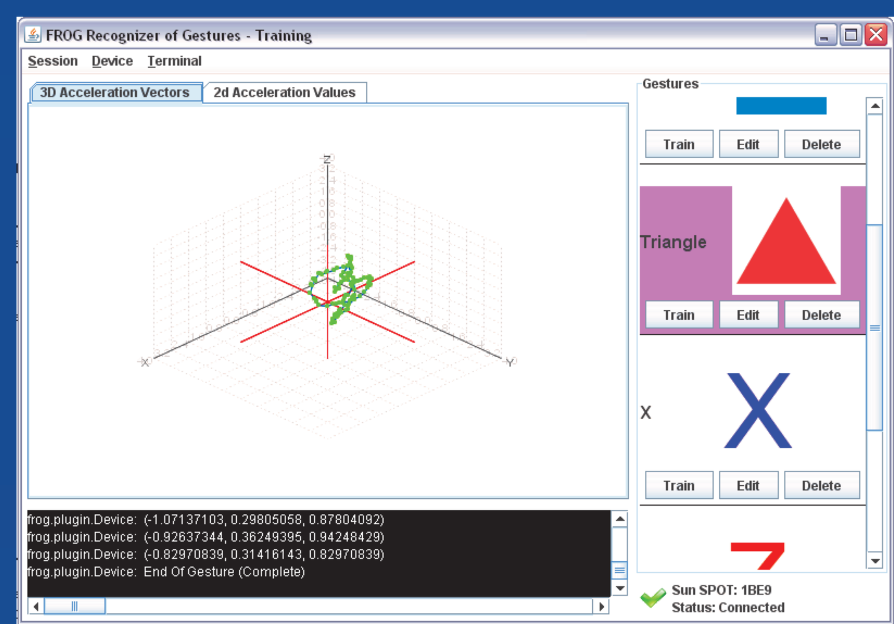
- Provides a plug-in framework that allows multiple, heterogeneous mobile devices to be used as input devices
- Launches a plug-in for Sun SPOTs, but plug-ins for other devices can be easily integrated
- Development of new plug-ins requires implementation of the FROG plug-in interface



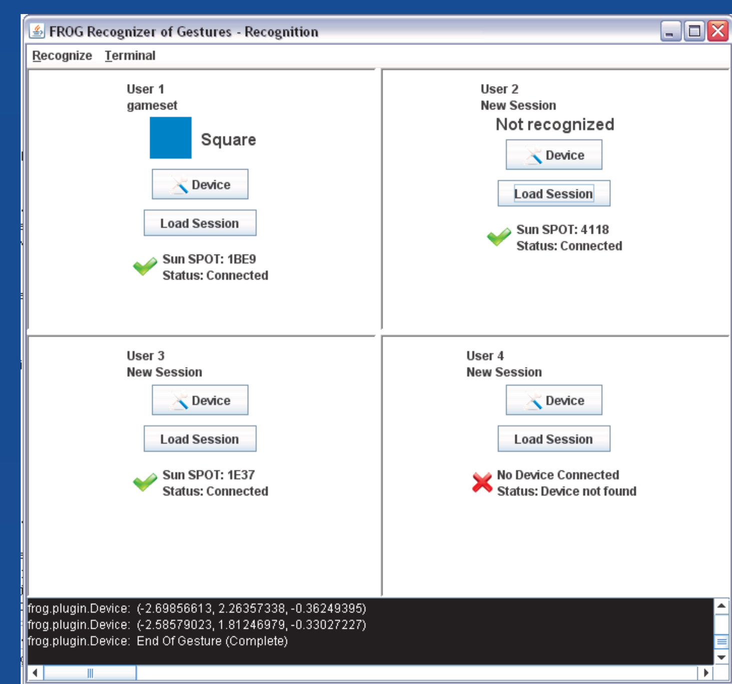
A Sun SPOT with its three axes

The Four Modes in FROG

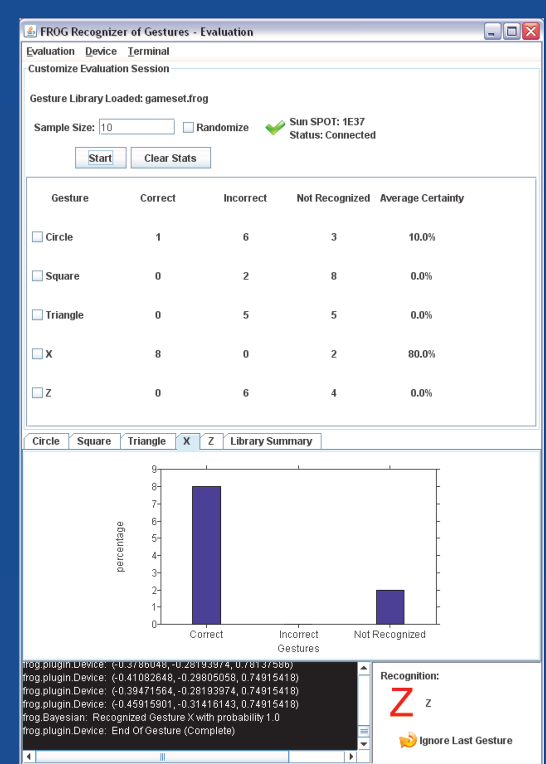
Training



Recognition



Evaluation



Demo

Theory

- An HMM is a doubly stochastic math model that we use to represent a gesture.
- In training, quantized acceleration vectors are used to form a left-to-right HMM that represents the gesture.
- The HMM is initialized by default, only to be iteratively optimized (using the Baum-Welch algorithm) according to each input training instance to produce a model for the gesture.
- In recognition, the classifier matches a performed gesture to the most likely trained gesture model.

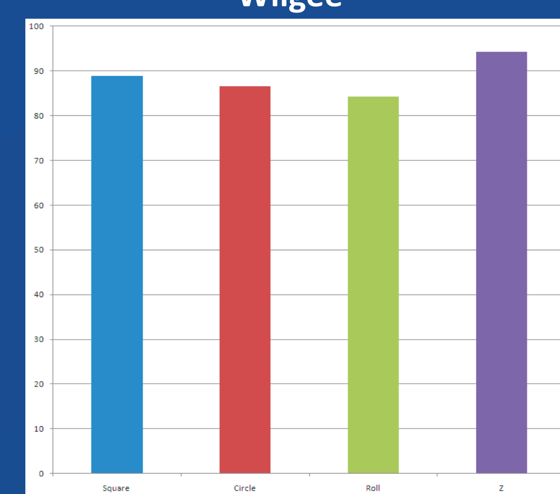
Applications

- Efficient enough to use as input for video game or multimedia presentation
- Could be used in a public environment
  - Gesture-controlled information kiosks
  - Games in movie theatres and malls
  - Potential for compatibility with nearly any modern phone

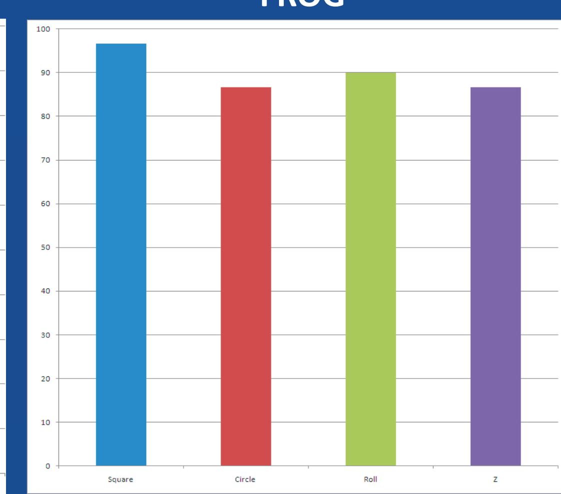
Conclusions

- The FROG project demonstrates that 3D acceleration-based gesture recognition provides a viable form of user interaction.
- Parameter modification in data sampling, filtering, quantizing, and modeling allows the system to be "tuned".
- Framework testing as well as other related research has shown that the quality of training plays a major role in accuracy of recognition.
- FROG recognition results utilizing Sun SPOTs as input devices has been shown to compare favorably to results obtained from the Wiigee project (see graphs below).
- Emerging technologies incorporated into today's mobile devices, such as 3D acceleration sensors, greatly expand the potential uses of these devices.

Wiigee



FROG



Average recognition rate in Wiigee, an open source gesture recognition project, (on the left) and FROG (on the right) with a similar set of gestures. Wiigee had 15 samples per gesture per user with four users participating in the trial, while FROG had 15 samples per gesture per user with two users participating in the trial. The average recognition rate was comparable.

Works Cited

Hobermann, Rose. Durand, Dannie. HMM Lecture Notes. 2006. Carnegie Mellon School of Computer Science. 10 September 2009. http://www.cs.cmu.edu/~durand/03-711/2006/Lectures/hmm-bw.pdf.

Rabiner, L. R. "A tutorial on hidden Markov models and selected applications in speech recognition." *Proceedings of the IEEE* 77 (Feb 1989): 257-286.

Schlömer, Thomas. Poppinga, Benjamin. Henze, Niels. Boll, Susanne. Gesture Recognition with a Wii Controller. 2008. http://wiigee.org/. 9 April 2009.

Acknowledgments

We would like to thank Dr. Payne for advising us throughout the course of the project. We also thank the Computer Science Department as well as the professors in the department for giving us the opportunity to work on this project. Finally, we thank the Mobile HCI Lab for being involved in the development process, and implementing a plug-in of their own.