A Gesture Interpretation and Voice Recognition Multi-modal Human Machine Interface

Francis K.H. Quek
Vision Interfaces and Systems Laboratory (VISLab)
Electrical Engineering and Computer Science Department
The University of Illinois at Chicago
Chicago, IL 60607
email: quek@eecs.uic.edu
WWW: http://vislab.eecs.uic.edu

Co-PI: Ramesh Jain, ECE Department, University of California at San Diego. jain@ece.ucsd.edu.
Program Area: Other Communication Modalities
Keywords: Gesture Interfaces, Deixis, Computer Vision, Dynamic Vision, Hidden Markov Models.

1 Project Summary

We are investigating various aspects of non-formal two and three-dimensional gesture interaction. In two-dimensional gesture interaction, we have developed a vision-based system: *FingerMouse* which replaces mouse operation by pointing motions above the keyboard. We have tested the modality with promising results. In three-dimensional gesture interaction, we base our research on the foundation of psycholinguistic and semiotic knowledge. Since spatial description and structuring is dominant in the natural hand gesturing of humans, we concentrate our research in this area. We believe that a significant class of gestures may be recognized if one can 1. determine and locate the extrema the gestural stroke, 2. determine the dynamics of hand motion during and around the gesture stroke 3. recognize the hand poses at the extrema of the stroke. We approach vision-based gesture interpretation by 1. segmenting gesture sequences by a hidden markov's model, 2. analyzing the dynamics of hand motion and using dynamic features in our HMM, 3. applying a rule-based learning scheme to learn hand pose models. We have applied our approach to the control of a robot.

2 Project References

1. Quek, F., Zhao, M., and Mysliwiec, T., "Dynamic Hand Gesture Analysis," Submitted, *Computer-Human Interaction '96*, CHI'96.

2. Quek, F., Zhao, M., and Mysliwiec, T., "Free-hand Pointing with FingerMouse," Submitted, *Computer-Human Interaction '96*, CHI'96.

3. Quek, F., "Non-Verbal Vision-Based Interfaces," Keynote Speech, *International Workshop in Human Interface Technology '95*, IWHIT'95, October 12-13, Aizuwakamatsu, Fukushima, Japan.

4. Quek, F., "Eyes in the Interface," *International Journal of Image and Vision Computing*, Vol.13, Number 6, August, 1995, pp. 511-525.

5. Quek, F., Mysliwiec, T., and Zhao, M., "Finger-Mouse: A Freehand Pointing Interface," in *Proceedings of the International Workshop on Automatic Face- and Gesture-Recognition*, Zurich, Switzerland, June 26-28, 1995, pp. 372-377.

6. Hunter, E., Schlenzig, J. and Jain, R. "Posture Estimation in Reduced-Model Gesture Input Systems" in *Proceedings of the International Workshop on Automatic Face- and Gesture-Recognition*, Zurich, Switzerland, June 26-28, pp 290-295, 1995.

7. Zhao, M., and Quek, F., "RBI: A rule-based induction approach and its application to hand pose recognition," *Technical Report VISLab-95-001*, Vision Interfaces and Systems Laboratory, The University of Illinois at Chicago, June, 1995.

8. Zhao, M, and Quek, F., "RBI: A Rule-Based Induction Approach and Its Application in Hand Pose Recognition," VISLab. Technical Report, VISLab-95-001, June 1995.

9. Schlenzig, J., Hunter, E., and Jain, R., "Recursive identification of gesture inputs using hidden Markov models," *Proceedings of the Second IEEE Workshop on Applications of Computer Vision*, pp 187-194. IEEE Computer Society Press, 5-7 December 1994.

10. Katkere, A., Hunter, E., Kuramura, D., Schlenzig, J., Moezzi, S., and Jain, R., "ROBOGEST:

Telepresence using Hand Gestures," *Technical Report VCL-94-104*, Visual Computing Laboratory, University of California, San Diego, December 1994.

11. Schlenzig, J., Hunter, E., and Jain, R., "Vision-based Human Hand Gesture Interpretation using Recursive Estimation," In *Twenty-Eighth Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, California, November 1994.

12. Mysliwiec, T., "FingerMouse," *Masters Project Report*, Department of Electrical Engineering and Computer Science, The University of Illinois at Chicago. Technical Report VISLab-94-001, Vision Interfaces and Systems Laboratory, The University of Illinois at Chicago, October 1994.

13. Quek, F., "Toward a Vision-Based Hand Gesture Interface," *Proceedings of the Virtual Reality System Technology Conference*, Singapore, August 23-26, 1994, pp. 17-29.

14. Quek, F., "Vision-Based Gesture Interpretation," *Virtual Reality Fall 1993 Conference*, New York, September 1993.

15. Quek, F., "Hand gesture interface for human-machine interaction," in *Proceedings of the Virtual Reality'93 Conference*, New York, 15–17, March 1993, pp. 13-19.

3 Area Background

My primary area of interest is the role computer vision plays in the human-computer interface. Since one may view HCI as a kind of dialog between the human and computer, and since vision plays a major role in inter-human interaction, we believe there is a significant role for machine vision in the interface. There are three ways in which machine vision reseacrh may be applied to HCI.

First, computer vision can serve as a non-contact means by which the machine receives human input (e.g. hand gesture interfaces, sign language interpretation, body posture interaction, face recognition, facial expression interpretation, and lip reading). We are particularly interested in hand gesture interaction and facial expression interpretation.

Second, computer vision techniques may be used to model and present information for user visualization. We have an on-going project in face animation to present the user with a face with a range of expressions and lip-phoneme synchronizations.

Third, the *perception-space* may be the ideal venue of interaction between humans and machines for a variety of problems which involve significant perceptual complexity. We are developing a model for such interaction which we call *Human-Machine Perceptual Cooperation*. The idea is to exploit the facility by which humans communicate at the level of directing each other's attentional resources to solve complex visual interpretation problems. Our model employs two channels of interaction: *context* (what to look for – conceptual focus) and *focus of attention* (where to look – spatial focus). We have applied this paradigm to the task of interpretating complex hand-drawn maps.

4 Area References

1. Quek, F., Jain, R., and Weymouth, T.E., "An abstraction-based approach to 3D Pose Determination from Range Images," *IEEE Transactions on Pattern Recognition and Machine Intelligence*, vol. PAMI-15, no. 7, pp. 722-736, July 1993.

2. Goshtasby, A. and Quek, F., "Using Gaussians in image filtering and data approximation," *IMACS World Congress*, Atlanta, Georgia, July 4-8, Vol. 1, pp. 193-195, 1993.

3. Quek, F., and Petro, M., "Human-Machine Perceptual Cooperation," *Proceedings of the International Conference on Computer-Human Interaction INTERCHI'93: Human Factors in Computing Systems*, pp. 123-130, Amsterdam, The Netherlands, 24–29, April 1993.

4. Quek, F., and Petro, M., "Interactive map conversion: Combining machine vision and human input," in *Proceedings of the IEEE Workshop on Applications of Computer Vision*, Palm Springs, CA., pp. 255-264, Nov. 30–Dec. 2 1992.

5 Related Program Areas

Virtual Environments, Speech and Natural Language Understanding, Intelligent Interactive Systems for Persons with Disabilities.

6 Potential Related Projects

Don't know. I hope to find out at this workshop.