Human-Computer Interaction System with Artificial Neural Network Using Motion Tracker and Data Glove

Cemil Oz and Ming C. Leu

Department of Mechanical and Aerospace Engineering,
University of Missouri-Rolla, Rolla, Missouri 65409, USA
{ozc, mleu}@umr.edu

Abstract. A Human-Computer Interaction (HCI) system has been developed with an Artificial Neural Network (ANN) using a motion tracker and a data glove. The HCI system is able to recognize American Sign Language letter and number gestures. The finger joint angle data obtained from the strain gauges in the sensory glove define the hand shape while the data from the motion tracker describe the hand position and orientation. The data flow from the sensory glove is controlled by a software trigger using the data from the motion tracker during signing. Then, the glove data is processed by a recognition neural network.

1 Introduction

Using our hands is a primary way of interacting with the outside world. We perform many everyday tasks with our hands; however, we usually use constrained peripheral devices such as a mouse, keyboard or joystick to work with a computer and computer-controlled applications. Sensory glove based input devices could be used overcome this limitation [1]. Commercial devices, such as the VPL data glove and Mattel power glove, have led to an explosion of research and development projects using electronic glove interfaces to computer applications and computer controlled devices. These applications include virtual reality, video games, scientific visualization, puppetry, and gesture-based control.

There has been a significant amount of research work done in the area of gesture recognition in the last decade due to the recent advances in hardware and software for human-computer interaction (HCI). Most of these studies were mainly in the area of sign language recognition [2-6] and game control as well as some other HCI tasks [7].

In this study, an HCI system is designed and implemented with an artificial neural network for recognition of ASL letter and number gestures using Cyberglove and Flock of Birds devices. The neural network for gestures recognition is activated by a software trigger. When the software finds the velocity of the hand below a threshold value, the data from the Cyberglove™ is sent to the gestures recognition network. The data flow from the input devices to the recognition process is then turned off until a high-speed hand movement reactivates it. The software for gesture recognition is based on an ANN method. The system is able to recognize 31 hand gestures. The system is developed for human-computer interaction in a robotic workcell.

S.K. Pal et al. (Eds.): PReMI 2005, LNCS 3776, pp. 280 – 286, 2005.

2 Overall System

The objective of this project is to control a robotic workcell with commands issued through hand gestures, which are recognized by a hand gesture recognition system. The system consists of the following components: object recognition, target detection, inverse kinematics solution, path planning, and robot control [8]. Figure 1 illustrates the overall system. The dashed-line blocks are not discussed in this paper.

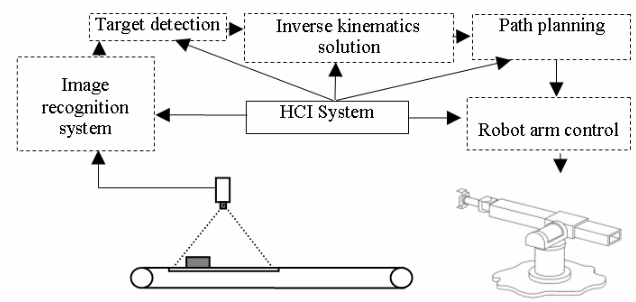

Fig. 1. Overall system block diagram

3 System Hardware

We use a right-hand Cyberglove™ (Figure 2) to retrieve the values of finger joint angles for gesture features. The glove measures the bending angles at various positions and the frequency of data collection can reach 150 Hz. The glove contains fifteen sensors: three for the thumb, two for each of the other four fingers, and four sensors located between the fingers. To track the position and orientation of the hand in 3-D space, we mount the Flock of Birds® motion tracker (Figure 3) on the wrist. The receiver is located in a DC pulsed magnetic field and its effective range is up to 8 feet around the transmitter. The measuring frequency can reach 144 Hz.

Open Inventor SDK (Software Development Kit) is used in the software development for 3-D scene rendering and interactive programming. It is a high-level

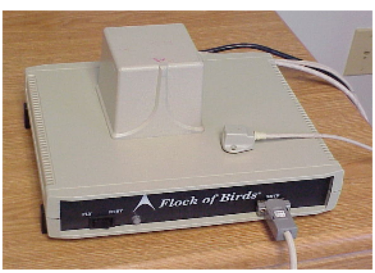

Fig. 2. The Cyberglove™ with 18 sensors **Fig. 3.** The Flock of Birds® 3-D motion tracker

tool kit developed in OpenGL for graphic rendering and user interaction. The software system is implemented using the Object Oriented Programming (OOP) technology; therefore, it is easily extendable.

4 ANN Based Gesture Recognition

We have designed an HCI system for American Sign Language (ASL) letter and number gestures. ASL has twenty-six letters and ten numbers. Although most of the ASL letter and number gestures depend on finger positions only, some of them also depend on hand orientation, and two of them are dynamic. There are great similarities between the signs of g and q, h and u, and k and p. The letters of each of these pairs have the same hand shape, but their hand orientations are different. There are also great similarities between i and j, and x and z. The letters of each of these pairs have the same hand shape, but the signs for j and z are dynamic. Figure 4 shows the hand signs for ASL letters.

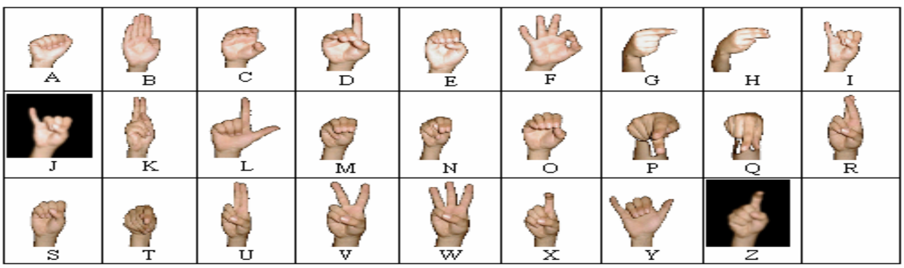

Fig. 4. ASL alphabet gestures

Our HCI system is based on finger positions only; therefore the alphabet characters g, h, p, j and z are not used. In other words, thirty-one gestures in total are used in the HCI system. The HCI system provides the user with the capability of generating many different commands using the 31 letters and numbers. In total, there are 55 commands: 10 commands for the conveyor, 30 commands for the robot arm, 10 commands for the image processing, and 5 commands for the whole system. Each hand shape either corresponds to a unique command such as start, stop, etc. or it is part of a more complex command composed of multiple hand shapes. There are complex commands which may use up to four hand gestures. For example, RT11 command means "take object number 1 and place in bin number 1". Some commands are given in Table 1.

A multi-layer ANN is designed to recognize the ASL letter and number gestures. The ANN model was detailed in previous papers [9, 10]. The input to the network consists of 15 Cyberglove data elements. The proposed ANN, a multi-layer feedforward network, consists of 15 input neurons, 20 hidden neurons, and 31 output neurons. A Levenberg-Marquardt backpropagation algorithm is used for training. The ANN has been trained and tested for two different data sets: single-user data and

Table 1. Representative commands for the HCI system

Whole system	Conveyor	Robot arm	Image processing
S- stop	CS- start conveyor	ROC- cubic path	IS- start image capturing
E- end	CE- stop conveyor	ROS- sinusoidal path	IE- stop image capturing
		RAN- add noise	ICF- extract image features
		RT11- take object 1 to bin 1	IA- add image to database

multi-user data. The output vector consists of 31 elements. The maximum of these elements corresponds to an alphabet or sign. The training set is composed of two files, input and output. The input file contains Cyberglove data which belong to 31 characters, and each character has 15 values which are provided as input to the ANN. The target vector consists of 31 values. All of these values are set to 0 except one value, which is set to 1. The position of the element with 1 defines the hand shape. The overall design of the HCI system is given in Figure 5. It consists of four parts: selection of input data, trained network, output decoding and command generation.

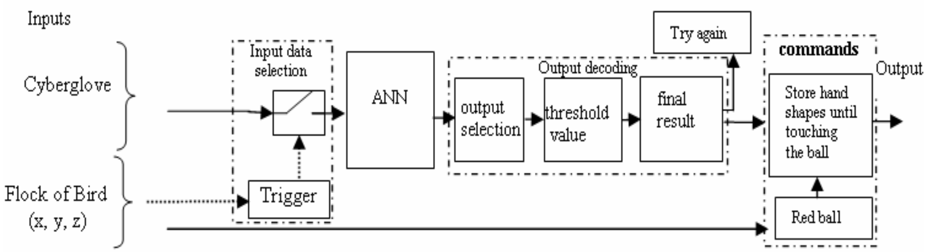

Fig. 5. The HCI system block diagram

A software trigger controls the reading of input data to the recognition network. When the trigger finds the velocity of the hand below a threshold value, in our case 0.05 (unit/second), the data from the Cyberglove and Flock of Birds is sent to the gesture recognition network. The data flow between the input devices and the recognition process is turned off until a hand velocity reactivates it. The reactivation velocity is set at 0.5. The three-layer word recognition network has 15 inputs, 20 hidden neurons and 31 output neurons. The selection block determines the maximum output of the ANN, and then the treshold block checks whether the maximum output value is acceptable or not. In the final part, the system output is stored in a variable to form a command. After issuing the command, the corresponting procedure of the command is procesed by touching a virtual red ball. The recognized letters and commands are displayed on the screen using an Open Inventor interface. Some of the ASL letters recognized by our system are illustrated in Figure 6, and some of cereated commands are showned in Figure 7.

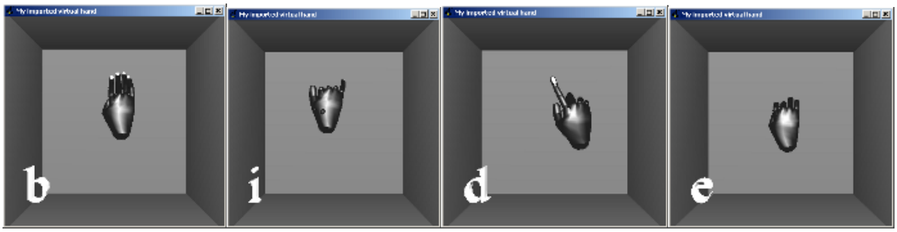

Fig. 6. Sample outputs representing four recognized hand shapes representing ASL letters

Fig. 7. Three command samples

5 System Test Results

Two ASL letter and number recognition systems have been developed, one with a single-user model and one the other with a multiple-user model. The recognition system with a single user model was first trained with data from three samples. When that was not effective, we trained the ANN with six, nine, then twelve and finally, fifteen samples. Similarly, the recognition system with a multi-user model was first trained with data from three samples, then six, nine, twelve and finally, fifteen samples. At the testing stage, real-time data from alphabet and number gestures are used. Each of these two systems was tested starting with *A* and ending with 9. The testing results are given in Table 2 and Table 3. The recognition accuracy of the single-user system is about 96% when the system is trained with 15 samples and tested in real time.

Table 2. Test result for a single user

	ASL alphabet																														
Samples	**A**	**B**	**C**	**D**	**E**	**F**	**I**	**L**	**M**	**N**	**O**	**P**	**Q**	**R**	**S**	**T**	**U**	**V**	**W**	**X**	**Y**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
3	/	-	-	/	-	-	-	/	-	/	/	-	/	-	-	/	-	/	-	/	/	-	/	-	-	-	/	/	/	/	/
6	-	-	-	-	-	-	/	-	-	/	-	-	/	-	-	-	-	-	-	/	/	-	/	/	-	-	/	/	-	-	/
9	-	-	-	-	-	-	/	-	-	/	-	/	-	-	-	-	-	-	-	/	-	-	-	-	-	-	-	-	-	/	-
12	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
15	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	(/) unrecognized, (-) recognized																														

Table 3. Test result for multiple users

	ASL alphabet																													
Samples	**A**	**B**	**C**	**D**	**E**	**F**	**I**	**L**	**M**	**N**	**O**	**P**	**Q**	**R**	**S**	**T**	**U**	**V**	**W**	**X**	**Y**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
3	/	/	-	/	-	-	-	/	/	-	/	/	/	-	/	/	-	/	-	/	/	-	/	-	-	-	/	/	/	/	/
6	/	-	-	/	-	-	/	-	/	/	-	-	/	-	-	-	/	-	-	/	/	/	/	/	-	-	/	/	-	-	-
9	-	-	-	-	-	-	-	/	-	/	-	/	-	-	-	-	-	/	-	/	-	-	-	/	-	/	-	-	-	/	-
12	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	/
15	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	(/) unrecognized, (-) recognized																														

6 Conclusion

A Human-Computer Interaction system has been developed with an artificial neural network using position/orientation sensors and sensory glove data for recognition of 31 ASL alphabet and number gestures. The system uses a Cyberglove, and a Flock of Birds 3-D motion tracker to provide the data input, and it has been trained and tested for single and multiple users for 31 hand gestures. The test results have shown that the proposed technique is capable of performing accurate real-time recognition of hand gestures.

Acknowledgments

This research was partially supported by a Ford Foundation grant, as well as by the Intelligent Systems Center at the University of Missouri–Rolla.

References

[1] D. J. Sturman and D. Zelter, "A Survey of Glove-Based Input," IEEE Computer and Applications, 1994, pp. 30-39.

[2] S. G. Wysoski, M. V. Lamar, S. Kuroyanagi, and A. Iwata, "A Rotation Invariant Approach on Static Gesture Recognition Using Boundary Histograms and Neural networks," Proceeding of the 9th International Conference on Neural Information Processing (ICONIP'02), vol. 4, 2002, pp. 2137-2141.

[3] C. Vogler and D. Metaxas, "ASL Recognition Based on a Coupling Between HMMs and 3D Motion Analysis," Proceedings of the sixth IEEE International Computer Vision Conference, January 1998, pp.363-369.

[4] G. Lalit and M Suei, "Gesture-Based Interaction and Communication: Automated Classification of Hand Gesture Contours," IEEE Transactions on Systems, Man, and Cybernetics, Part c: Application and Reviews, vol. 31, no. 1, February 2001, pp. 114-120.

[5] Y. Cui and J. Weng, "A Learning-Based Prediction and Verification Segmentation Scheme for Hand Sign Image Sequence," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, no.8, August 1999, pp. 798-804.

[6] C. Vogler and D. Metaxas, "Parallel Hidden Markov Models for American Sign Language Recognition," IEEE Proceeding of International Computer Vision Conference, 1999, pp. 116-122.

[7] L. K. Lee, S. Kim, Y. K. Choi, and M. H. Lee, "Recognition of Hand Gesture to Human-Computer Interaction," IECON 2000, 26th Annual Confjerence of the IEEE 2000, pp. 2177-2122
[8] R. Koker, C. Oz, and A. Ferikoglu," Object Recognition Based on Moment Invariants Using Artificial Neural Networks," Proceedings International Symposium on Intelligent Manufacturing Systems, August 30-31, 2001.
[9] C. Oz, N. N. Sarawate, and M. C. Leu, "American Sign Language Word Recognition with a Sensory Glove Using Artificial Neural Networks," Proceedings of ANNIE Conference, St. Louis, MO, November 7-10, 2004; also Book on Intelligent Engineering Systems through Artificial Neural Networks, Vol. 14, ISBN 0-7918-0228-0, 2004, pp. 633-638.
[10] C. Oz and M. Leu, "Recognition of Finger Spelling of American Sign Language with ANN Using Position/Orientation Sensors and Data Glove," Proceedings of 2nd International Symposium on Neural Networks, Chongqing, CHINA, May 30 - June 1, 2005; also Lecturer Notes in Computer Science 3497, Springer, ISBN 3-540-25913-9, 2005, pp. 157-164.