KEYSTROKE RECOGNITION FOR VIRTUAL KEYBOARD

Jani Mäntyjärvi[1], Jussi Koivumäki[2], Petri Vuori[3]

[1] Nokia Research Center, P.O.Box 407, FIN-00045 NOKIA GROUP, Helsinki, Finland
[2] Nokia Networks, P.O.Box 319, FIN-90651 OULU, Finland
[3] Nokia Mobile Phones, P.O.Box 86, FIN-24101 SALO, Finland

ABSTRACT

The progress in the field of human-computer interaction with hand held electronic devices, such as, personal digital assistants (PDAs) and mobile phones searches for new interaction techniques. Proximity sensing extends the concept of computer-human interaction beyond actual physical contact with a device. In this paper, a virtual keyboard implementation is presented and keystroke recognition experiments with the keyboard utilizing proximity measurements are described. An infrared (IR) transceiver array is used for detecting the proximity of a finger. Keystroke recognition accuracy is examined with k-nearest neighbor (k-NN) classifier while a multilayer perceptron (MLP) classifier is designed for online implementation. Experiments and results of keystroke classification are presented for both classifiers. The recognition accuracy, which is between 78% and 99% for k-NN classifier and between 69% and 96% for MLP classifier, depends mainly on the location of a specific key on the keyboard area.

1. INTRODUCTION

Recent development of human-machine interaction seeks for novel interaction and perception methods. For instance, the user interface for wearable devices developed by Starner et al. [13] exploits IR proximity sensing and pattern recognition algorithms for gesture recognition.

The challenge of fusing the sensor data is commonly met in several application areas with soft computing techniques like fuzzy logic, and neural networks [1, 3, 4, 5]. Proximity sensing has been successfully used in diverse fields of application. One example is the field of industrial robotics, where IR proximity sensing has been utilized for providing information for obstacle impact avoidance [12].

The sensing methods discussed in this study are suitable for developing novel applications to improve the user interaction of mobile terminals. Hand held devices are becoming a part of everyday life. Furthermore, the diversification of applications and usage environments creates needs for developing novel ways of interaction [11]. Despite recent developments in this area, the user interfaces of mobile terminals still resemble conventional computer control. Due to the small size of these devices there are some unavoidable compromises related to keyboard usage. Different solutions to these problems have been suggested; including voice control, touch screen, unconventional keyboard [8, 9]. Similarly, the use of proximity sensing in touch screen and virtual keyboard applications has been studied and commercial products are available [2, 14].

In this work, a virtual keyboard implementation is presented. The proximity sensing system enabling virtual keyboard realization consists of four pairs of IR transmitters and receivers. Transceivers are designed to operate in cycles for generating four dimensional signal vectors that are further processed into feature vectors, which are then used for locating a finger on the virtual keyboard area. Finger positions on the virtual keyboard area are mapped into keystrokes. Keystroke recognition accuracy is examined offline with k-NN classifier, while a MLP classifier is designed for online keystroke recognition.

2. VIRTUAL KEYBOARD

The approach of this study abandons the concept of the physical keyboard, at least in the traditional way of thinking. Instead, four pairs of IR sensors were used to create a virtual keyboard area beside the input device.

The proximity sensing system including the IR transceiver control and the arrangement of the virtual keyboard is presented in Fig. 1. The actual device is enclosed in a plastic box. The test keyboard area is drawn on a white paper sheet and the device is situated so that each IR transceiver has one row of keys in front of it. For the real implementation, there are several ways of realizing a keyboard area, for example, by projecting the borders of the keyboard with light emitting diodes (LEDs).

0-7803-7304-9/02/$17.00 C2002 IEEE

In the virtual keyboard, four IR transceivers are placed approximately 2 cm apart in the device. Transmitters are mounted on top of and the receivers under a printed circuit board (marked with arrows in Fig. 1). The "silvery" surface inside the box generates suitable apertures for delimiting the space angle of each receiver.

The transceivers of the device operate in cycles controlled by the clock frequency of the proximity-sensing device. The signal received by one transceiver at a given time is a sum of signals reflected from a finger. By processing these proximity signals it is possible to recognize different keystrokes on a certain 'sensitive' area in normal usage situations.

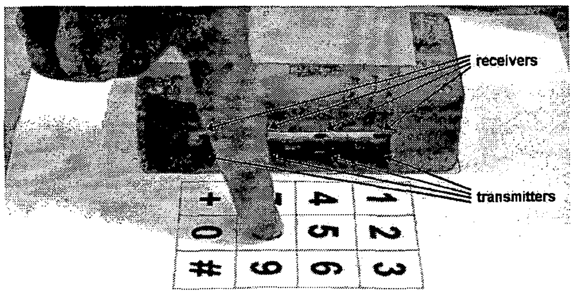

Figure 1: The virtual keyboard.

3. METHODS

In this work, two common classification methods are applied. MLP and k-NN are widely used classification methods in data analysis [1, 3, 4, 7, 10, 15]. The k-nearest neighbor classifier is used to recognize different keystrokes for the virtual keyboard. The MLP classifier is designed for online keystroke recognition.

3.1. Multilayer Perceptron

Typically, a MLP network consists of an input layer, various numbers of hidden layers and an output layer, (Eq. 1). Each layer consists of a predefined number of computation nodes. In training, a pattern memory is distributed into the values of the weight factors of the network by using training material and the back propagation training algorithm. The trained network forms a set of outputs that are linked to the arbitrary set of inputs. [3]

$$y_i = f(\sum_j \omega_{ij} x_j) \qquad (1),$$

where x_j is an input node j, y_i is an output node i, ω_{ij} is the weight associated between nodes x_j and y_i, and $f(\cdot)$ is the activation function.

3.2. K-nearest neighbor

In k-NN classification the similarity between a test feature vector and prototype vectors (training set) are computed. The similarity measure is typically Euclidean distance. K most similar prototype vectors are selected. These are called k-nearest neighbors. The class of the feature vector is determined by selecting the class that has majority among the k-nearest neighbors. [7]

3.3. Generating Feature Vectors for Classification

The keystroke recognition system includes the preprocessing and classification tasks presented in Fig. 2. The input to the preprocessing system is a matrix $\boldsymbol{X} = [\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{x_3}, \boldsymbol{x_4}]^T$, where rows of $\boldsymbol{X}$ are sampled signals from each of the four IR receivers. The following phases are performed separately for each row of the input matrix:

1. A low-pass filtering of the vector component with a sixth-order IIR type Butterworth filter with a cut-off frequency of 75 Hz,
2. Decimation of the filtered vector component with a factor of 5.
3. Normalization of the decimated vector component to a range of [0,1].

The individual rows are recombined to a feature matrix $\boldsymbol{X'}$, which is used for classification experiments. In online classification a feature vector at instant time t is used.

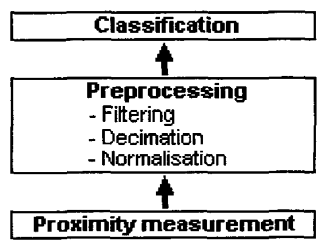

Figure 2: Illustration of the keystroke recognition system.

3.4. Classification

The offline classification experiments with k-nn are performed in the following manner. First the data set $\boldsymbol{X'}$ is divided into 10 data sets $\boldsymbol{X'_1}, \ldots, \boldsymbol{X'_{10}}$ in a random manner by using stratified ten-fold cross-validation [6]. Data sets $\boldsymbol{X'_1}, \ldots, \boldsymbol{X'_{10}}$ contain the same number of samples and approximately the same frequency of classes. K-nn classifiers are built ten times and they are used with different data sets. Each time one data subset in turn is excluded from the training and is used for testing. The cross-validation classification performance is the average of the performances of the ten test sets.

In order to implement the online keystroke recognition system, several MLP neural network architectures with differing training parameters and scaling functions are tested for achieving the best classification performance. The architecture of MLP classifiers selected for classification experiments consists of an input layer with four input neurons, one hidden layer with 10 weights, and an output layer with 12 output neurons. The number of output neurons is determined by the number of classes to be recognized. MLPs exploit the back propagation algorithm in training and all neurons use logistic activation functions.

4. EXPERIMENTS AND RESULTS

Data acquisition and data storing is carried out by using the LabView® graphical programming environment from National Instruments. Signals from four IR receivers are sampled with National Instruments DaqCard 1200-measurement board connected to a PC-CARD slot in a laptop PC. The sampling procedure with a sampling rate of 1.5 kHz is controlled by the clock frequency of the proximity-sensing device. Preprocessing of the signals and offline keystroke recognition is performed with MATLAB® software from The MathWorks Inc. The MLP classifier for online classification is implemented into a LabView® program.

The test keyboard is rectangular; the sides of the area perpendicular to the device are 6 cm long and the parallel sides are 8 cm long. There are a total of 12 keys, which all have the same area of 4 square centimeters, 2cm×2cm (Fig. 1). The data used in experiments is recorded in normal keyboard usage situations by using seven testees. During recording, the right hand index finger is held vertically and moved within each key area, as presented in Fig. 1.

In order to get an equal amount of data from each class, data is recorded for approximately the same period of time within each key area. Sample sequences of error conditions are also recorded consisting of recordings outside the keyboard area and on the edges of the buttons.

4.1. Keystroke signals

Responses of the four IR transceivers as a function of distance are examined and they are presented in Fig. 3. The finger is moved in one-cm intervals away from transceiver 1 and signal levels from all transceivers are recorded.

As Fig. 3 depicts, transceiver 1 receives most of the reflected IR signal from close by (0-3 cm). The received intensity of the IR reflections of the other transceivers (2, 3 and 4) is at its highest, when the finger is about 2-4 cm away from the array plane. The signal schemes of the other three transceivers are quite similar to this graph, i.e., when the finger is moved in perpendicular direction in front of these three transceivers. The response of the IR receiver located directly in front of the finger decreases very rapidly as a function of distance limiting the size of the virtual keyboard.

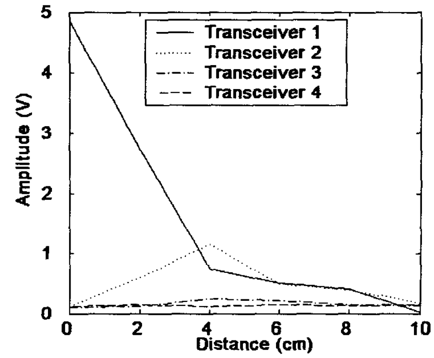

Figure 3: Responses of IR transceivers.

4.2. Classification Experiments

Classification results of the keystroke recognition are presented in Fig. 4. Black bars represent the recognition accuracy of a k-NN classifier with cross-validated experiments while white bars illustrate the accuracy of a MLP classifier designed for online implementation.

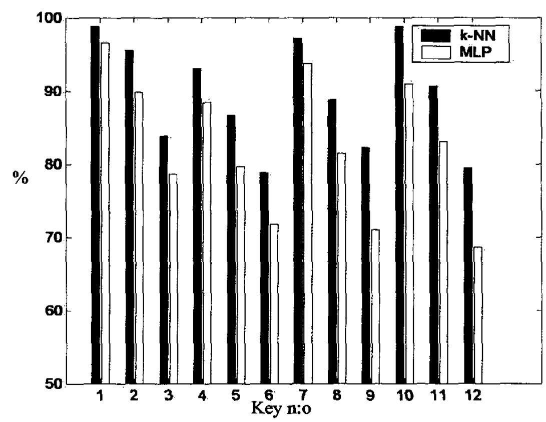

Figure 4: Classification recognition results assigned to a corresponding key.

The recognition accuracy decreases as the distance from the sensor array increases. Keys located farther away from the transceiver array (3,6,9,12) have notably lower recognition accuracies. This is due to the restrictions discussed in the previous chapter. A more detailed presentation for the k-nn classification experiments is in Table 1, representing the confusion matrix for keystrokes, with lowest recognition accuracy. Results are presented as the perceptual amount of data classified to each class. The matrix depicts that the classes with lowest recognition accuracy located farther away from the transceiver array, and next to each other are confused in classification. For example, 14% of data from the class key 6 is classified to the class key 3. The misclassifications that are not shown in the table are negligible.

5. CONCLUSION AND FUTURE WORK

This work demonstrates the realization of virtual keyboard utilizing IR proximity sensing system and keystroke recognition with k-NN and MLP neural networks. Keystroke classification results are between 78% and 99% for k-nn classifier and between 69% and 96% for MLP classifier. The classification accuracy depends mainly on the location of a specific key on the keyboard area.

In the experiments, the finger had to be held perpendicular to the surface of the keyboard. This is obviously an unwanted feature, but it can be made less significant when the hardware and recognition software are developed further. In order to direct the IR radiation better into a narrow horizontal plane, the design of aperture optics will be further investigated. Further research includes also testing in different usage environments with a wide variety of persons including different skin colors.

6. REFERENCES

[1] V. Cherkassky, F. Mulier, *Learning from Data: Concepts Theory and Methods*, John Wiley & Sons. 1998.

[2] *CarrollTouch Infrared Touchscreens*, www.elotouch.com/products/cttec/default.asp, 2002.

[3] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Prentice Hall, 2nd edition, 1999.

[4] S. Iyengar, R. Brooks, *Multi-Sensor Fusion: Fundamentals and Applications with Software*, Prentice Hall, 1997.

[5] G.J. Kim, B. Yuan, *Fuzzy Sets and Fuzzy Theory and Applications*, Prentice Hall, 1995.

[6] R. Kohavi, "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection," *Proceedings of the International Conference on Artificial Intelligence*, Morgan Kaufman, pp. 1137-1143, 1995.

[7] S. R. Kulkarni, G. Lugosi, S. S. Venkatesh "Learnin Pattern Classification - A Survey," IEEE Transactions on Information Theory, Vol.44, No. 6, pp. 2178- 2206, October 1998.

[8] K. Murai, K. Hashimoto, *Information Inputting and Processing Apparatus*, (U.S Pat. No. 5,635,958), 1997.

[9] B.A. Myers, "A Brief History of Human Computer Interaction Technology," *ACM interactions, Vol. 5, no. 2*, pp. 44-54, March 1998.

[10] S. Osowski, D.D. Nghia, "Neural networks for classification of 2-D patterns Signal Processing Proceedings," *5th International Conference on Signal Processing, Vol.3*, pp. 1568-1571, 2000.

[11] A. Schmidt, M. Beigl, H-W. Gellersen, "Sensor-Based Adaptive Mobile User Interfaces", *Proceedings on the 8th International Conference on Human-Computer Interaction*, pp. 251-255, 1999.

[12] H.R. Singh, S. Chauhan, A. Mobin, S.S. Agrawal, "Design and development of voice/tele operated intelligent mobile robot," *Proceedings of IEEE Region 10 Annual Conference on Speech and Image Technologies for Computing and Telecommunications, Vol.1*, pp. 177-180, 1997.

[13] T. Starner, J. Auxier, D. Ashbrook, M. Gandy, "The Gesture Pendant: A Self-illuminating, Wearable, Infrared Computer Vision System for Home Automation Control and Medical Monitoring," *Proceedings of the Fourth International Symposium on Wearable Computing*, pp. 87-94, 2000.

[14] *Virtual Keyboard Unit,* www.vkb.co.il, 2002.

[15] C.H. Wang, B.H. Xiao, R.W. Dai, "A New Integration Scheme with Multilayer Perceptron Networks for Handwritten Chinese Character Recognition", *Proceedings of 15th International Conference on Pattern Recognition, Vol.2*, pp. 961-964, 2000.

Table 1: A confusion matrix of keystroke classification with k-NN. Perceptual amount of data classified to each class. Actual classes are presented on the first row, predicted classes are presented on the first column.

KEY	1	2	3	4	5	6	7	8	9	10	11	12	Error
3	0%	0%	84%	0%	0%	4%	0%	0%	0%	0%	0%	0%	12%
6	0%	0%	14%	0%	1%	79%	0%	0%	3%	0%	0%	0%	3%
9	0%	0%	0%	0%	0%	10%	0%	1%	82%	0%	0%	5%	2%
12	0%	0%	0%	0%	0%	0%	0%	0%	11%	0%	1%	80%	8%