ISBN 978-93-84422-79-0
9th International Conference on Recent Trends in Science Engineering, Computers and Technology (RTSECT-2017)
Singapore Aug. 10-11, 2017

Vision Based Classification of Hand Gestures towards Human Robot Interaction by Using Neural Network

Mehmet DEMİR[1] Raşit KÖKER[1], Mehmed KANTARDZIC[2]

[1]Sakarya University, Faculty of Technology Electrical and Electronics Engineering Department
54187 Sakarya – TURKEYE-mail: mehmetdemir125@hotmail.com rkoker@sakarya.edu.tr

[2]University of Louisville J. B. Speed School of Engineering Computer Engineering and Computer Science Department Louisville, KY 40292, USA E-mail: mehmed.kantardzic@louisville.edu

Abstract: *In this research study, a computer vision based system has been designed for the purpose of the operation of a robotic system by using hand gesture. Human robot interaction is one of the most important issue especially in robotics. During the studies hand gesture images have been collected by using a camera. Six tasks have been defined for the purpose of the operations on a robotic system. Moment invariants have been used as a feature to classify the commands given by the hand gesture. Then, a neural network has been designed based on the classification of the hand gestures, which refer to defined tasks. At the end by using obtained weights the designed neural network has been integrated to the prepared computer vision software towards the operation of a robotic system. The obtained results and implementations of the algorithms have been presented in the paper.*

Keywords: *Human robot interaction, image processing, hand gesture recognition.*

1. Introduction

Hand gestures are one of the most important issues in the human robot interaction. On the same way hand gestures are also used as significant techniques in human computer interaction. It can be used in many applications, such as human computer interaction, tele-operated robots vision based control of robots and etc. Hand gestures can be collected based on using a special glove or based on computer vision techniques to develope these applications [5]. Specific gestures of interest may be identified from the corpus of gestures and specific commands can be obtained for the execution of some actions belonging a robotic system [6]. The designed hand gesture recognitions systems can be used in human computer interaction or human robot interaction by using computers or any of embedded systems for the real-time implementation.

Recently, hand gesture recognition applications are used by the researchers in many robotic applications. Mobile robots have the ability to move around in their environment and they are not limited in one physical location [8]. In the human-interaction hand gesture recognition takes a significant role because of the fact that hand gestures are naturally a powerful way of communication, and they can be used in the remote control of robots [9]. In the interpretation of the hand gestures for human robot interaction two approaches have been commonly used that are gloved-based and vision based methods. Glove based methods requires wearing a cumbersome contact devices and there is generally carrying a load of cables which connect the device to computer [10]. Furthermore, visual servoing may be more concerned to gesture and interpretation of the environment [11, 12]. There are many ways concentrating on the different parts of body to recognise gestures by using computer vision techniques. One of these ways is the evaluation of the usage of pointing gestures for the location of the objects. This approach is based on different feature maps such as intensity, edge, motion and disparity which are valuable information to track the head or hands [13, 14]. As an example to another way

Davis et al. presented a study based on the two dimensional locations of the fingertips and palms [15]. Ghosh et al. presented a study about static hand gesture recognition system based on using mixture of features and support verctor machine (SVM) classifier. They extracted localized contour sequences and block based features to propose a novel mixture of features [7].

In this paper, a vision based system has been designed for the control of a robotic system. Some tasks have been determined at the beginning and a hand gesture recognition system has been designed by using a neural network classifier. Moment invariant have been used as a feature extraction technique. Six different tasks have been described to be classified. All used image processing algorithms and classification design have been explained in the paper. Image Processing

2. Image Processing

The related image processing algorithms have been given in this section. The software has been developed by using C++ programming language. OpenCv library has also been used during the development of codes.

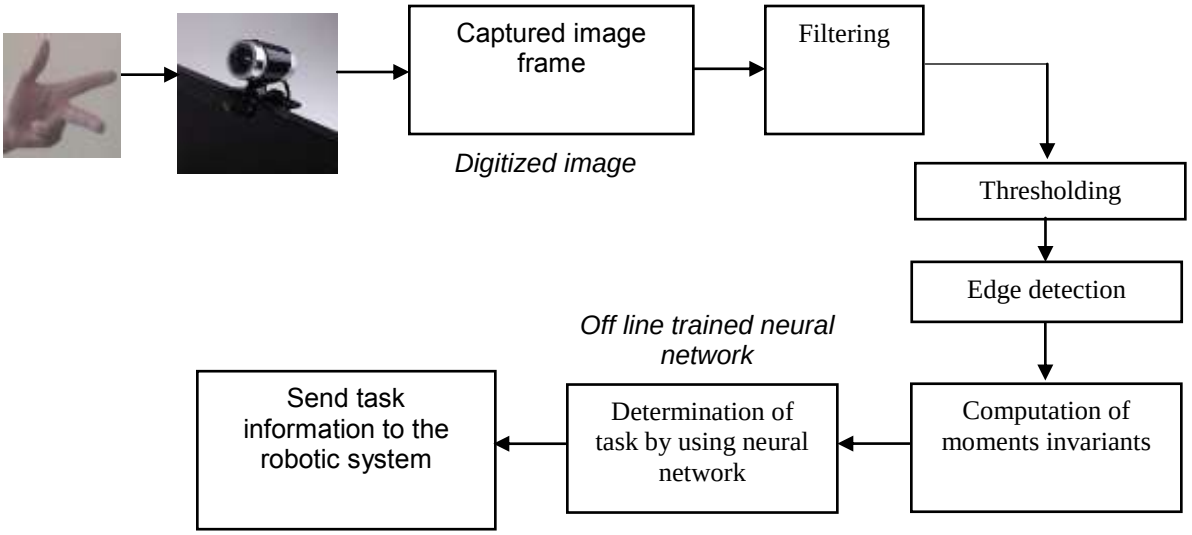

Fig.1 The block scheme of the designed vision based control of a robotic system.

2.1. Low-Level Image Processing Algorithms

An experimental captured image was a 256 colored gray image. In the developed software initial image was reduced down to 250x250 pixels to decrease the process time. To remove the various types of noise in digital images spatial filters can be applied. These spatials are apllied on images based on using some matrixes ranging from 3x3 to 11x11. A median filter was used in this study to remove undesirable effects due to the noise and some other effects. Noise reduction is significant in this study since moment invariants are very sensitive to the noise. The block scheme of the designed vision based system has been presented in Fig. 1.

2.2. Intermediate Level Image Processing Algorithm

In intermediate level processing, edge detection can be categorized. The aim of using edge detection is to simplify the hand gesture image but to have more meaning full image since the edges are known as basic image features carrying useful information regarding the object boundaries. Edge detection algorithms are also used to find complicated boundaries belonging objects based on marking the potential edge points, which correspond to the points in the image where there are changes in brightness [16]. Among the large number of edge detection algorithms, Sobel algorithm working based gradient is used due to its popularity on computational simplicity [3]. The gradient based edge detection methods detects the edge points based on searching for the maximum and minimum points in the first derivative of the image.

2.3. High Level Image Processing Algorithms

Classifications of hand gestures have been mainly aimed in this section. Moment invariants have been used here as a descriptor to classify hand gestures. For each hand gesture image class many images have been collected to set up a database for the preparation of training and test for the neural network. The backpropagation error learning algorithm has been used to train the neural network. After the training of neural network the obtained weights have been used in the main program to get the hand gesture class information. Sample captured images with their edge images belonging six hand gesture classes have been shown in fig.2.

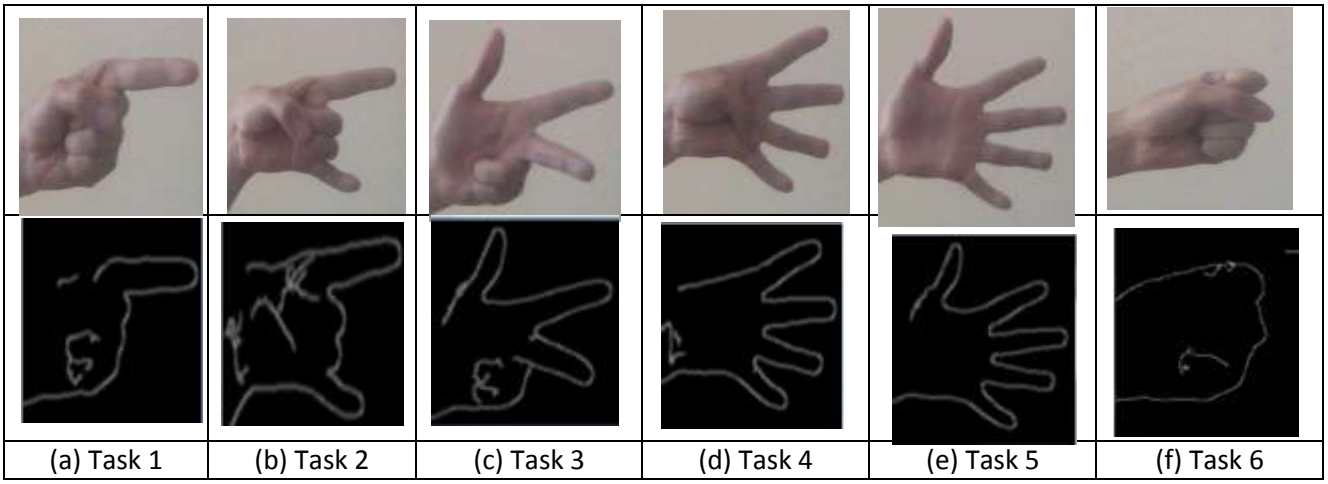

Fig. 2 Sample captured images with their edge images belonging six hand gesture classes

3. Neural Network Design

A neural network is a technique, which seeks to build an intelligent computer program based on using models simulating working neurons in the human brain. The neural network are used to perform computations on computers that includes different neurons connected to the other neurons with synaptic weights for the simulation of the human brain. Researchers have been dealing with that these neurons are responsible for the human ability to learn and it in this sense that a neural network to achieve machine learning is simulating physical structure. The duty of each of these computational units are computing some function of its inputs and transmits the results to the joined process elements in the network [17, 18, 19].

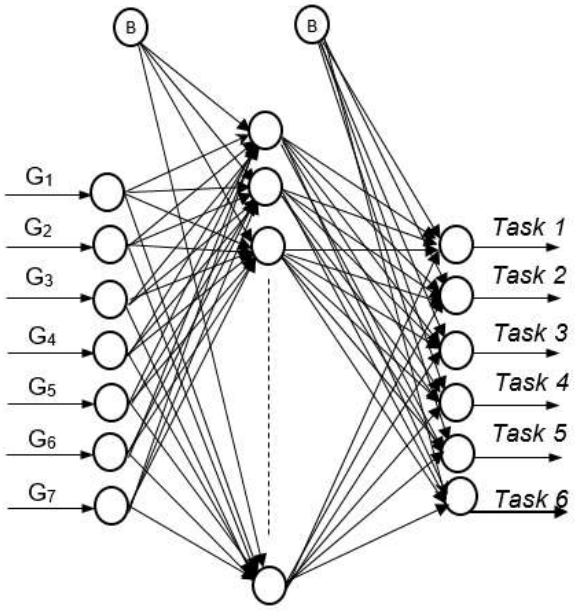

Fig. 3. The topology of the designed neural networks

The neural networks were trained using the MATLAB neural-network toolbox. A conventional backpropagation algorithm using a threshold with a sigmoidal activation function and a gradient-descent learning algorithm was employed [21]. The momentum coefficient, learning rate, and number of neurons in the hidden layers were determined experimentally. The topology of the designed neural networks has been given in fig. 3.

During the training in Matlab 'trainbr' function has been used. Trainbr is a neural network training function based on updating the weights and bias values according to Levenberg-Marquardt optimization. It is minimizing a combination of squared errors and weights, and then determining the correct combination, therefore, it produces a network which is generalizing well. This process is called as Bayesian regularization [20]. For training in this study, 420 data has been prepared regarding six different hand gesture images. Seventy data for each hand gesture image has been put into the learning set. These values were recorded in a file to form the learning set for the network. On the other hand, 180 data in other words 30 data for each hand gesture image has been prepared for the test set of the neural network. The test dataset is a part of the input dataset used to test how well the neural network will perform on new data. The error is computed as the mean squared error (MSE). The training parameters and test results are presented in Table 1.

Table 1 Training parameters and test results from the neural networks.

Number of neurons in input	7
Number of neurons in input	6
Activation function	Tangent sigmoid
Number of neurons in hidden layer	50
Range used for number of neurons in hidden layer	[20-50]
Number of epochs	12
Sample size in training set	420
Sample size in validation set	20
Sample size in test set	180
MSE values for training set	0.00851449
MSE values for validation test	0.0219872
MSE values for test dataset	0.0154130

4. Results and Discussions

The experimental works have been implemented by using a web camera on a labtop. Six different hand gesture images have been collected firstly offline. 110 images for each selected hand gesture classes have been collected in different rotation, orientation and scale. 660 images have been used totally and moment invariant values have been computer by using these images to set up the dataset. 420 of them have been used to set up training set data. 180 images of each hand gesture have been used for the test set of the neural network and remaining 60 images have been included in test and validation set equally. The training and testing performances of the neural network have been given in Table 2. In the training set each task have been recognized with the performance of greater than %99. On the other hand, for the test set the recognition rate is also recognized with the success rate greated than % 99.

Table 2. Training and testing performances o the neural network

	Training performance	Test performance
Task 1	99.9888	99.98985
Task 2	99.9689	99.9153
Task 3	99.9881	99.98986
Task 4	99.9871	96,64568
Task 5	99.9882	99,995
Task 6	99.9694	99.14305

The moment invariants are sensitive to the noise and other effects since they are calculated based on the geometric centroid. In Fig. 4. the hand gesture images have been given with arm and without arm including centroid point. It is seen in the figures the centroid point is affected because of the arm. That is why to increase the success of the recognition rate the images have been taken as much as hand gesture by cutting the arm in the image.

(a) (b)

Fig. 4. The centroids on the images with arm (a) and without arm (b)

The neural network predictions are obtained around one since digital coding has been coded to define the classes. The outpus of the neural network have been accepted as based on the selected threshold value, which is 0.70. It means any output higher than 0.70 has been accepted as 1. To say something about the process times, image acquisition has been done based on 27 frames per second. Capturing hand image and processing to find the hand gesture class information has been implemented in this period. The new image capture has been done when the new hand image is in the work area. As seen above the results have been satisfactory and the response time is quick to be used in human robot interaction.

5. Conclusion

In this study, a vision based hand gesture classification system has been designed to command a robotic system by using neural networks as a human robot interaction application. Based on data mining techniques moment invariants have been selected and used as a feature extraction technique. 99 percent classification rate has been observed for the test set for any given hand gesture image. As much as possible getting the hand gesture area segmented from the arm is increasing the classification rate since the arm image is affacting the centroid point, which is critical in the computation of moment invariants. The response of the designed system is very quick in microseconds, that is why the proposed solution system may be used in any robotic system for the operation, where the speed is required. As a future study, some other feature extration techniques or classifiers can be used to obtain hundred percent successful classification rate.

6. References

[1] Huang Z., and J. Leng, "Analysis of Hu's Moment Invariants on Imace Scaling and Rotation", Proceedings of 2010 *2nd International Conference on Comp. Engineering and Technology (ICCET)*, pp. 476-180, Chengdu, China. IEEE.

[2] Koker R., Oz C., and A. Ferikoglu, "Development of Vision Based Object Classification System for an Industrial Robotic Manipulator", *The 8th IEEE Intern. Conf. on Electronics, Circuits and Syst. (ICECS*), pp. 1281-1284, Vol. 3, 2001, Malta.

[3] Koker R., Intelligent Predictive Control of a Six-Joint Robotic Manipulator Towards a Vision Based Application, Intelligent Information Systems Conference, pages 23-35, 8-10 June 2010, Siedlce, Poland.

[4] Ming-Kuei H., "Visual Pattern Recognition by moment invariants", *Information Theory, IRE Transactions,* Vol.8, pp. 179-187, 1962.

[5] Yu C., Wang X., Huang H., Shen J., and K. Wu, "Vision-based hand gesture recognition using combinational features", *IEEE Sixth Int. Conf..on Intelligent Information Hiding Multimedia Signal Processing,* 15-17 Oct 2010, Darmstadt, Germany.

[6] Raheja J. L., Shyam R., Kumar U. and P.B. Prasad, "Real-Time Robotic Hand Control Using Hand Gestures", *IEEE Second Int. Conf. On Machine Learning and Computing (ICMCL),* 9-11 Feb 2010, Bangalore, India.

[7] Ghosh D.K. and S. Ari, "Static Hand Gesture Recognition Using Mixture of Features and SVM Classifier", *IEEE Fifth Int. Conf. on Communication Systems and Network Tech.*, 4-6 April 2015, Gwalior, MP, India.

[8] Lee C.H., Chiu M.H., "Recurrent neuro fuzzy control design for tracking of mobile robots via hybrid algorithm", *Expert Systems With Applications,* Vol. 36 (5), pp. 8993-8999, 2009.

[9] Faudzi A.A., Ali M.H.K., Azman M.A., and Ismail Z.H., "Real-Time Hand Gestures System for Mobile Robots Control", *Procedia Engineering* 41, pp.798-804, 2012.

[10] Hasanuzzaman M.V., Ampornaramveth et. al., "Real-time Vision Based Gesture Recognition for Human Robot Interaction", *IEEE International Conference on Robotics and Biomimetics, ROBIO 2004,* pp. 413-418, Shenyang, China.

[11] Lee S.H., Cheong S.N., Ooi C.P., and W.H. Siew, "Real Time FPGA Implementation of Hand Gesture Recognizer System", *Proceedings of the 15th WSEAS intern. Conf. on Computers,* pp. 217-222, Wisconsin, USA, 2011.

[12] Malima A.E., Ozgur E., et al., "A fast Algorithm For Vision-Based Hand Gesture Recognition For Robot Control", *IEEE 14th Signal Processing and Communication Applications,* 2006.

[13] Nickel K., and R., Stiefelhagen, "Visual recognition of pointing gestures for human-robot interaction", *Image and Vision Computing,* Vol. 25 (12), pp. 1875-1884, 2007.

[14] Jojic N., Brumitt B., Meyers B., Harris S.,T. Huang, "Detection and Estimation of Pointing Gestures in Dense Disparity Maps", *IEEE Intern. Conf. on Automatic Face and Gesture Recognition*, Grenoble, France, 2000.

[15] Davis J., and M. Shah, "Recognizing Hand Gestures", *Proc. of The European Conf. on Computer Vision,* ECCV, pp. 1875-1884, 1994.

[16] Haitham Badi, Sameem Abdul Kareem and H. Sabah, "Feature Extraction Technique for Static Hand Gesture Recognition", Recent Trends In Hand Gesture Recognition GCSR Vol. 3 Edited by Ankit Chaudhary ISBN 978-618-81418-4-1 (e-book).

[17] Haykin S., "Neural networks", Macmillian College Publishing Company; 1994. p. 138.

[18] Ross Timothy J. Fuzzy logic with engineering applications. New York: McGraw-Hill; 1995.

[19] Wasserman PD. Neural computing. New York: Van Nostrand Reinhold; 1989.

[20] Beale M.H., Hagan M.T., H.B. Demuth, "Neural Network Toolbox™ User's Guide", *Mathworks,* R2016a.

[21] Luengo J., and F. Herrera, "Shared domains of competence of approximate learning models using measures of separability of classes", *Information Sciences* 185 (2012) 43–65.