Hand Sign Recognition System and study of RBF classifier

[1]Mrs. More Madhura Mahesh, [2]Prof Dattatray S Bade

[1]*M.E. [Second Year] Electronics and Telecommunications Engineering, Alamuri Ratnamala Institute of Engineering and Technology (ARIET), Thane - Maharashtra, India*
[2]*Assistant professor, Vidyalankar Institute of Technology, wadala, mumbai.*

Abstract- *The prehistoric or primitive cave men's where communication first existed or started were using a means of sign language or gesture for communicating with each other. With the dawn of time, the man started working tirelessly to make his life better and better. The inventions of new technologies and the computer brought a huge change in the life of people. The Computer and advent of the internet made communication and transfer of information easier. But this is always an interaction with machines. The future era is conceiving of human-computer interaction (HCI) that can be used for implementing 3D applications in which user can move or rotate his hands without any input devices like keyboard or mouse. The communication involving hand gestures or signs for communication forms a basic platform for the proposed system. Many researchers have performed different work by using different algorithms. Some of them even have revised their previous algorithm. The current work and related literature review have given a way of developing a new system that can overcome all the system and should provide the best accuracy. The proposed system involves capture of images and then processing it, then filtering the unwanted part. The further processing of the image is carried out using FPGA, i.e. Field Programmable Gate Array. The system desires to give an excellent simulation, good accuracy, minimizes run time and efficiently reduces the power consumption. It will be forming a great alternative with low cost and best computational algorithm. The proposed system is desired to give better results and desired characteristics after implementation.*

Keywords- *Human Computer Interaction (HCI), FPGA (Field Programmable Gate Array), Discrete Fourier transform (DFT), Radial Basis Function (RBF), Very Large Scale Integration (VLSI).*

I. INTRODUCTION

Information from one person or group to another using words, signs or semiotic rules which are understandable by both the persons. Usually it can be verbal or non-verbal means of conversation. Nonverbal communication is defined as "Communication without words". It basically includes behaviors like expressions on face, eye contact, touching, voice tone or sometimes messages such as dress, posture or spatial distance between each other.

Gesture is a non-verbal as well as non-vocal communication in which messages are communicated using body actions visible to each other, or either replacement or conjuncting with a speech. Gesture is moving of hands, face or parts of the body. Hand gesture is a strong aspect of communication, for both speaker and listener. It provides alternative to inefficient interface devices that are used for Human Computer Interaction (HCI). Real time processing is the important feature of using hand signs for HCI. The applications of a gesture recognition system, depending upon its growing importance in life are HCI (Human computer interaction) also known as Man Machine Interaction (MMI), robot control, games, surveillance etc. using different algorithms and techniques.

Present world of technological development which is advancing every day with new researches, inventions & developments in the field of computer science, medical, image processing, VLSI etc. with advances in software and hardware. But still there is very slow and gradual growth of technology that makes an interaction between humans and machines. Therefore, the researchers had started giving more attention towards HCI rather than programming language and computer screen.

Hand recognitions are classified as hand postures and gesture recognition. Hand posture is static hand poses without movements. While, hand gestures are dynamic with various sequences of hand positions or postures. In this system hand posture recognition is used which implements a new real time system that integrates a single FPGA which incorporates all processing tasks and also image capturing on board camera. Integration of all the system modules is done using Very Large Scale Integration (VLSI) design. It combines thousands of transistors on a single chip. VLSI design provides Programmable Logic Device (PLD) and Field Programmable Gate Array (FPGA).

FPGA contains ten thousand to million logic gates with programmable interconnections. This interconnection is available for designers to perform various functions. It consist of I/O blocks designed and numbered as per functions and for each module of logic level, there are CLB (Configurable Logic Blocks). It performs logic operations. Now-a-days FPGA's are improved drastically, which can drive a huge increase in space, weight and power constraints. The FPGA will soon make it possible to develop a hand gesture recognition portable systems without a PC to decrease in response time and computing power.

The hand gestures are understood by the computer in two different approaches "Data-Glove Based" and second is "Vision-Based". The first approach uses hand configuration and the motions of fingers which are tracked by sensors and the data is digitized. The main drawback of this method is devices used are not user-friendly and are also expensive. While on the other side the "Vision-Based" approach uses a single camera that realizes interaction taking place between user and computer. Vision- Based approach is widely used in almost all the applications because of its easy and efficient way of working that yields accurate result and reduces time.

II.   LITERATURE SURVEY

Stefan Oniga, Alin Tisan *et al.*'s [1] implemented artificial neural network (ANN) using FPGA for various applications such as static hand gesture recognition. The system is a two level structure of ANN used for recognition. The first level is trained using a Hebbian algorithm for feed forward ANN which is acting as preprocessing data input. The second one is competitive ANN used for classification of data. Use of ANN for preprocessing data provides ease of implementing this method. A combination of two ANN gives a recognition rate 100% for two test sets and training set.

Yuuki Taki, Hirooma Hikawa et al.'s [2] put forward a hand posture recognition system using hybrid network classifier. It consists of SOM (Self Organizing Map) to which hand posture images are applied as input. These input images are processed and extract the feature vectors those and applied to the Hebbian network. The image of hand postures is taken as input from which feature vectors are extracted those are mapped to a lower dimensional map using SOM. Further, the Hebbian network performs naming and acquisition system feasibility is verified, by computer simulations. The results show that the system performance is better if, the number of neurons is sufficient in SOM. The Hybrid classifier is most efficient in embedded learning capability. The advancement can be done by implementing dynamic gesture with feedback SOM.

G. Tofighi et al.'s [3] trained a real-time vision-based hand posture recognition using appearance-based features of hand is developed. It involves three important steps: hand segmentation that introduced using "Adaptive histogram Template of skin" that extracts histogram of subjected hand by sampling its texture and color. The skin color areas in an image can be projected using a template. In the next feature extraction step, hand's edge contour and convex hull are used to extract global hand's features. It recognizes total ten hand posture classes. The last step is posture recognition where reference template is compared with subject hand features for identification. This system is fast and simple for hand recognition in real time.

[4] Varkonyi-koczy and Tusor made use of smart environments in order to make lifestyle luxurious, comfortable and also enhances the human quality of living. A hand posture recognition and gesture modeling system are developed. The proposed system can interface smart environments or intelligent space using simple hand gesture making communication very simple. The system uses a fuzzy neural network that can transform processed statistics of detected hand into a feature fuzzy model of hand posture. Depending on which an actual hand posture can be determined using fuzzy inference. Finally, the user's hand is detected from all the detected hand postures.

Dardas and Georganas [5] implemented a real-time system which provides interaction using hand gestures with applications or video games. The system works in three stages. The first stage is detecting hand and then tracking its features by performing face subtraction, detection of skin and contour comparison with efficient algorithms. The bag-of-features and multiclass support vector machine (SVM) are used for posture recognition. At this stage, the training image keypoints are extracted using SIFT algorithms, such as hand movement direction, posture transitions, and its scale helps to generate a database of gesture commands. The second is training stage in which vector quantization (VQ) using k-means clustering of keypoints for every training image is performed and then mapped into a bag-of-words vector. The final stage is testing stage. A multiclass SVM classifier classifies and detects accurate hand posture from keypoints of words vector of training images. The system under various conditions provides 96.23% accuracy.

 [6] the system in the paper extracts features and localizes objects in an image at 2000 frames/Sec for 512 x 512 images using high-frame-rate (HFR) vision system. A hardware logic for FPGA (Field Programmable Gate Array) based on high-speed vision uses a cell-based multi-object feature extraction algorithm. In an image of 1024 objects 25 higher order, local autocorrelation features can be extracted simultaneously for multi-object recognition. The image is fragmented into 8 x 8 cells with first order moments & zeroth calculation to obtain the locations & sizes of multiple objects. The system was verified by performing repeated experiments; for multiple objects rotating at 16 rotations/Sec, recognition of 1000 frames/search for multiple patterns projected & hand gesture with finger motion.

Summary of above studied works is as follows,

- The performances of the systems implemented above are based on software approach which requires high speed and efficient PC as a platform to execute the program for recognition.

- The use of FPGA, i.e. a silicon based module which can implement whole system is adapted to reduce the physical size and power consumption of the devices.

- Artificial Neural networks [1] are found to be hard to debug and also inefficient in computations i.e. it needs lots of chips & long run-time for large datasets.

- SVM [5] has a high algorithmic complexity and requires more memory requirements for the quadratic programming.

- After analyzing the existing SOM Hebb algorithm the main drawback is that neuron weights for clustering input are required in sufficient quantity. The SOM algorithm is provided with required data or information in the form of weights, the grouping done in map is not accurate. Thus, finding the relevant data becomes cumulative or sometimes impossible task.

III. PROBLEM DEFINITION

Human Computer Interaction (HCI) is a very important aspect of communication between computer and humans. Traditional methods of interacting are keyboard, mouse, sensors or some mechanical devices. But this way of communicating are not beneficial for elderly persons or a person with disabilities. This issue can be resolved by using gesture recognition system. Today, many researches are made to use body gesture as an interaction platform.

One of them is Hand Gesture Recognition system. The existing systems are detecting hand by processing various parameters such as extraction, segmentation, detection, etc. the hand gestures are static and dynamic. Systems are capable of recognizing both the means of input images, but the main drawback of every system lies in its performance, speed and power constraints. VLSI is an emerging technology which is used for implementing such systems. A Field Programmable Gate Array(FPGA) is used for developing this approach. According to survey many new systems were developed using FPGA, but the technique of detecting the gesture, i.e. the algorithms used in every paper is different each with different levels of accuracy.

Our aim is to design and implement a system that should compete with all the existing systems in terms of speed, performance and power and it should prove best in all the existing system. The system which we are implementing will be efficient and faster and the algorithm will be simple and easy for implementing.

IV. PROPOSED SYSTEM

As studied and compared the real-time hand recognition systems in the literature survey section the limitations were highlighted. Observation of this system was on the basis of performance of the system, the speed of the system and potential results of those and it was found that a new improved system can be implemented which can replace an existing one and which will give a better result than previous systems. The modification or change of the existing system is possible by using Radial Basis Function (RBF) classifier. It is a real- valued function having values of those vectors which depends on the distance from the center or origin.
The two main advantages of using Radial basis function classifier are,

- This classifier can be implemented using simple mathematics, i.e. linear Algebra.
- The computations are relatively cheap i.e. it can do complex mathematical computation in a relatively short time.
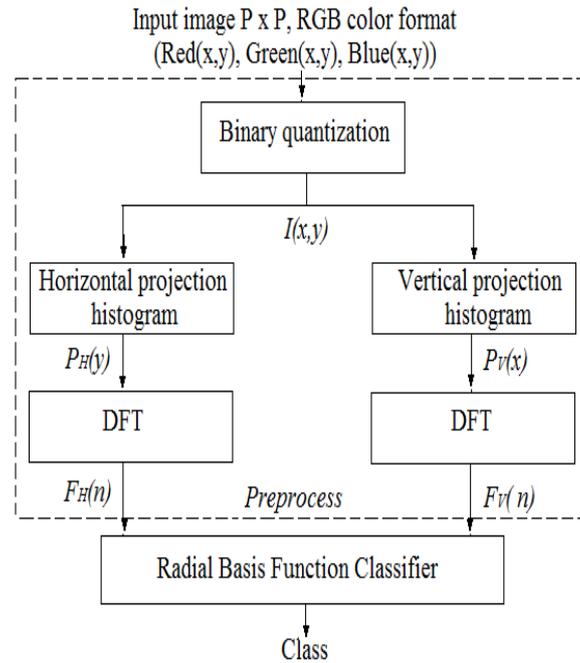
Figure 1. Flowchart of Hand sign recognition system

V.   METHODOLOGY

DISCRETE FOURIER TRANSFORM (DFT):

A DFT transforms discrete samples of a wave or signal into its frequency spectrum. It is a statistical method that degrades a signal into sinusoidal components of various frequencies ranging from 0 to maximum. This is a very important transform, and hence many applications are widely using this transform to perform Fourier analysis. It basically deals with limited amount of data, and so computers can be implemented by various techniques such as numerical algorithms or sometimes as specially designed hardware. DFT's are of great importance as they show periodicity of input response as well as strengths of periodic components.

The formula for DFT is:

$$X[k] = \sum_{n=0}^{N-1} x[n]e^{-j2\Pi kn/N}$$

It is extensively used in signal processing applications and related fields:
- To analyze signal and determine its frequency components, data compression of the information signal, Spread spectrum.
- To solve multiplications of large integers, partial differential equations etc.
- To perform various analytical operations such as convolutions, cross-correlation, power constraints, match filtering, analysis of LTI systems.

RADIAL BASIS FUNCTION:

The RBF is a real-valued function of which value is determined based on the distance from center or origin. It satisfies,

$$\phi(x) = \phi(\|x\|)$$

Distance from center point 'c' is given as,

$$\phi(x,c) = \phi(\|x - c\|)$$

Therefore, the function ϕ which satisfies the above property, is called as radial function. Basically, Euclidean distance is used as distance function.

The RBF is a type of neural network that can be used as nonlinear classifier. Usually, neural networks are referred to as Multilayer Preceptor (MLP) i.e. "Artificial Neural Networks". The weighted sum of the input values is taken by each neuron in MLP. The multiplication of input value and co-efficient is performed, and the results are added

together. One MLP neuron form a simple linear classifier, and combining these neurons in a network makes a complex non-linear classifier.

The RBFN approach is superior than the MLP. The RBF network compares the input with the example from the training set and classifies the neurons accordingly. The "prototype" is stored in a network as an example from the training set. While classifying new input, a Euclidean distance between input and prototype is computed by each neuron.

Radial basis function network architecture:

The architecture of Radial basis function works on concept of rough estimation i.e. approximation strategy. These strategies approximates the unknown function along with linear combinations of function which is non-linear called as **Basis functions.** This basis functions are called radial functions as they have radial symmetry with respect to the center.

The operation of RBF network or classifier is explained as. One of the vector from the training set vectors is stored in each RBF neuron as a *"Prototype"*. Comparison of input vector with the prototype vector which is usually in the center is done using RBF neuron which gives output in the form of values which is either '0' or '1'. As the distance between prototype and an input vector increase the response of the system goes towards 0.

The RBF network can be employed in any model, i.e. linear or nonlinear and also as any network i.e. single layer or multi layer. However, Broom head and Lowe's RBF networks were traditionally associated with a single layer network of radial functions as shown in figure.
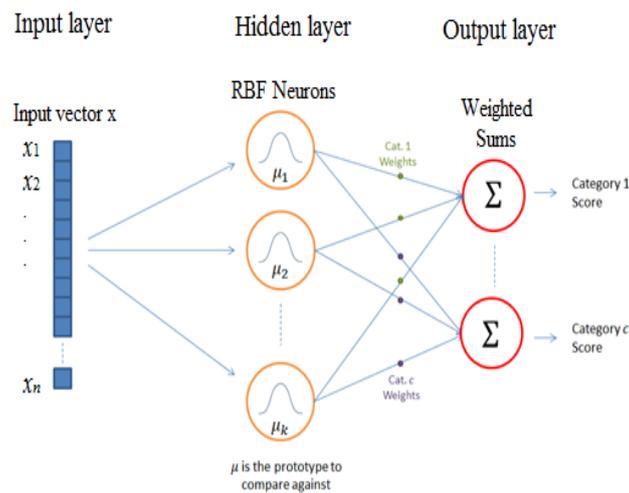
Figure 2. RBF network architecture

The above figure shows the RBF network architecture. The architecture composed of three layers those are input layer, hidden layer, i.e. RBF neuron layer, and output layer with class of data.

A. The Input Layer:

The sequence of values, i.e. x_1, x_2, ... a_n with n components forming n-dimensional vector that needs to be classified acts as input vector of the system. The complete input vector space is applied to respective RBF neurons of the hidden layer.

B. The Hidden Layer:

One of the vectors from the training set acts a "prototype" vector that is stored in single RBF neuron. Further, the comparison of RBF and the input vector to its prototype is done that gives output value 0 or 1. If both the input and the prototype are found equal, then the RBF neuron will be 1 at its output. If the distance between input and its prototype increases, it leads to falling in response exponentially to zero. The response of RBF neuron is a shape of the bell curve that is shown in above architecture diagram. The response value of neuron is known as "activation" value. The prototype vector that has a value at the center of the bell curve is also called as neuron's "center".

C. *The Output Layer:*

The set of nodes, one per category which is to be classified forms the output layer of the network. A score for the respective category is computed with each output node. Typically, a classification decision is made by assigning the input to the category with the highest score. The score is computed from every RBF neuron from which activation values are taken by calculating a weighted. The weighted sum means that a weight value is associated with respective RBF neuron with an output node and then this weight is multiplied with neuron's activation before adding it to the total response. As the score for a different category is computed by each output weights. At the output layer the RBF neurons that belong to its category is assigned a positive weight whereas, the others are assigned a negative weight.
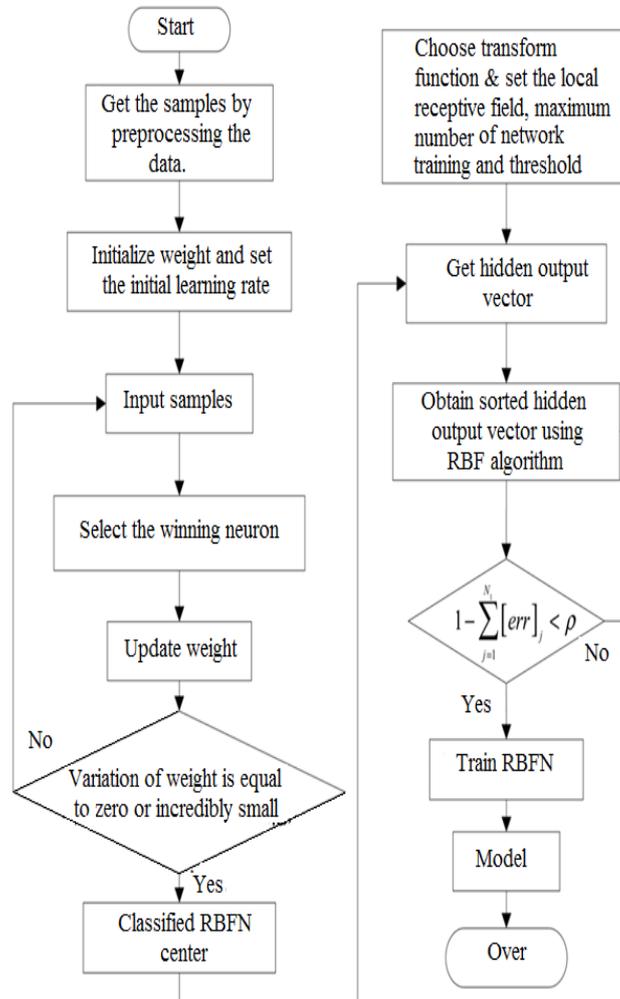
Flowchart of RBF Classifier:

Figure 3. Flowchart of RBF Classifier

VI.   CONCLUSION

The study of literature as discussed above shows all the work implemented using different approaches and techniques for recognizing various hand gestures or postures the conclusion can be drawn as hardware as well software platform is required to get accurate and faster results and to maintain the lowest possible recognition rate. The drawback of the system is explained in the summary of literature survey. To deal with all the problems of existing systems a new system with new algorithm is proposed. The captured image of a hand gesture is quantized into its horizontal and vertical components and then the extracted values are converted in spectrum domain representation using discrete Fourier transform. The processed data are then classified using radial basis function classifier to get the recognized image by hand. The main focus of many researchers is on redesigning feature vectors to make the system rotational and invariant of individual function i.e. scaling invariant. But this approach leads to increase in hardware cost. In contrast, proposed work focuses on classifier tuning by the new training method. A new technique is proposed to increase the accuracy, speed, and adaptability of a gesture recognition system reducing the power consumption. This system can surely prove its usefulness by helping millions of deaf people to communicate with normal people.

REFERENCES

[1] S. Oniga, A. Tisan, D. Mic, A. Buchman, and A. Vida-Ratiu, "Hand postures recognition system using artificial neural networks implemented in FPGA," in Proceedings. 30th Int. Spring Seminar Electron. Technology., 2007, pp. 507–512.

[2] Y. Taki, H. Hikawa, S. Miyoshi, and Y. Maeda, "Hand sign recognition system based on hybrid network classifier," in Proc. IEEE Int. Joint Conf. Neural Netw. (IJCNN), Jun. 2009, pp. 3074–3081.

[3] G. Tofighi, S. Amirhassan. Monadjemi, and Nasser. Ghasem-Aghaee, "Rapid hand posture recognition using adaptive histogram template of skin and hand edge contour," in Proc. 6th Iranian Mach. Vis. Image Process. (MVIP), Oct. 2010, pp. 15.

[4] Annamária R. Varkonyi-Koczy and Balazs. Tusor, "Human–computer interaction for smart environment applications using fuzzy hand posture and gesture models," IEEE Trans. Instrum. Meas., vol. 60, no. 5, pp. 1505–1514, May 2011.

[5] N. Dardas and N. D. Georganas, "Real-time hand gesture detection and recognition using bag-of-features and support vector machine tech-inquest," IEEE Trans. Instrum. Meas., vol. 60, no. 11, pp. 3592–3607, Nov. 2011.

[6] Q. Gu, T. Takaki, and I. Ishii, "Fast FPGA-based multiobject feature extraction," IEEE Trans. Circuits Syst. Video Technology., vol. 23, no. 1, pp. 30–45, Jan. 2013.

[7] J. Triesch and C. Vonder Malsburg, "A system for person-independent hand posture recognition against complex backgrounds," IEEE Trans. Pattern Anal. Mach. Intel., vol. 23, no. 12, pp. 1449–1453, Dec. 2001.