An Efficient Approach for the Recognition of Hand Gestures from Very Low Resolution Images

D.K. Vishwakarma, Rockey Maheshwari, Rajiv Kapoor
Electronics and Communication Engineering
Delhi Technological University
Delhi, India
(dkvishwakarma@dce.ac.in, rockeymahesh93@gmail.com, rajivkapoor@dce.ac.in)

Abstract—**In this paper, a simple and effective approach for the recognition of hand gestures from very low resolution images is proposed. Enhancement of the low resolution images has always been the key focus in the processing of the digital images. Images with resolution as low as [50×50 pixels] are also considered for recognition. The gestures under consideration here are the number of fingers (one, two, three, four or five) raised by the person. The low resolution gesture image captured from web camera, mobile phone, or low cost cameras is processed systematically to output the number of fingers raised. Simple concepts of the geometry of the hand have been used for the recognition of hand gesture from the input low resolution images. The proposed method extracts the hand gesture directly from the low resolution image without the need of reconstruction to a high resolution image or use of any classifier. The proposed method is based on the generation of a mask for the image which is critical in the recognition of the hand gesture. This method is tested on publically available dataset of Marcel-Triesch. The high accuracy of the experimental results show the superior performance of the proposed method for the recognition of hand gesture from low resolution images.**

Keywords-geomectry; hand gesture; mask generation; recognition; low resolution

I. INTRODUCTION

The recognition of hand gestures has become an area of active research in the field of computer vision and machine learning. Hand gesture recognition plays a central role in the interaction between the digital and physical world. Over the years, the human machine interaction has evolved from the physical touch to virtual gesture based interactions. The hand gesture recognition system have many applications such as automation, smart systems, virtual gaming etc. Application of hand recognition system in smaller devices require smaller digital storage space requirement, simpler and very fast processing. Our main focus for this method has been to accommodate all these features so that the designed system can be used under all constraints of the system.

Many difficulties are faced during the recognition of hand gestures such as background complexity, blurring of image, orientation of the gesture. With the increasing digital information required to be stored in system, the focus of the developers is to reduce the space required for the functionalities of the system. This supports the concept of using low resolution images which occupies lesser space for the purpose of gesture recognition.

The main focus of the proposed method is the development of a simple, robust, and effective method for the recognition of hand gestures. Thus, the proposed method does not include the use of any classifier as classifiers needs training time and their execution time also increases. Information extracted is based on the geometrical structure of the hand and not on colour, size or brightness. This ensures the methods reliability and efficiency in the cases of low lighting conditions. The orientation of the hand gesture is also taken into consideration so that the hand gesture need not to be necessarily in upright position. This is an extra advantage of the proposed method.

The rest of the paper is as follows. Section 2 highlights the previous work carried out for the gesture recognition. Section 3 describes the technical details of the proposed method. The experimental results obtained by the application of the proposed method are discussed in Section 4. The work that can be done in this field to further improve robustness of the system are explained in Section 5. Lastly, the paper is concluded and the references for the paper are cited.

II. RELATED WORK

Many researchers in the past have proposed the use of different algorithms for the recognition of human hand gesture from images. Reference [1] brought the concept of using depth images for the recognition of the hand gesture. Reference [2] presented a gesture recognition system for dynamic user interface with the use of multiple algorithms such as Harr like features, Camshaft, Lucas kanade etc. Reference [3] used the extraction of the fingertip for the Human-Robot interaction. Reference [4] compared the efficiency of the various classifiers for the recognition of gestures on two different datasets. All these efforts show the importance of the effective hand gesture recognition in the present world. Many attempts have also been made in the enhancing of the low resolution images using various techniques. Reference [5] used the method of removing quantization error in frequency domain from the low resolution image to generate the high resolution images. Reference [6] introduced the combined concept of Particle Swarm Optimization and discrete wavelet transform for the enhancement of low resolution images which can be further used for various functionalities. Reference [7] recognized gestures from low resolution videos. In [8], for the

978-1-4799-1797-6/15 $31.00 © 2015 IEEE
DOI 10.1109/CSNT.2015.84

recognition of the hand gesture from the images, the shape parameters of the hand are extracted. This concept of shape parameters of the hand is developed further to decrease the recognition time of the gesture and improve the accuracy of the system. The above discussed methods have the limitation that they require either the reconstruction of the low resolution image to a higher resolution or the recognition of the hand gesture requires the use of classifiers which involves the training of the system and hence slows the performance.

III. PROPOSED METHOD

This section explains in detail the various steps involved in the recognition of the hand gesture from the low resolution images. The low resolution images that can be recognized can have a resolution of as low as $[50 \times 50 \ pixels]$.

Fig. 1 shows the step by step working of the proposed method. These steps are explained one by one further in this section.

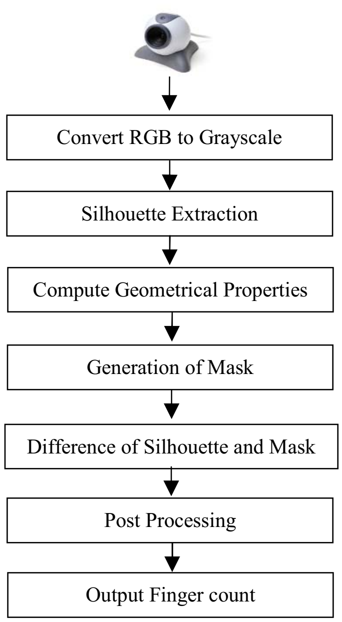


Figure 1: The work flow diagram for the proposed method for the recognition of the hand gesture from the low resolution images.

A. Pre-processing

The low resolution image containing the hand gesture needs to be pre-processed before the generation of the mask for the gesture. This involves the conversion of the RGB image to grayscale. The next important step for the recognition of hand gesture is the extraction of silhouette of hand.

- *Extraction of Binary Silhouette:* To extract the binary silhouette from the grayscale image, the concept of histogram is used. Image is a 2–dimensional function and thus functions representing an image will be a 2-D function. Let $R(x,y)$ represent the grayscale gesture image. The histogram of grayscale image, $R(x,y)$ determines the binary image based on a threshold value. If the value of $R(x,y)$ is greater than the threshold value then it is give a value equal to '0' in binary image else '1'. This is shown by (1).

  Let the threshold value for the grayscale image, $R(x,y)$ calculated dynamically be $'t'$. Now,

$$\begin{aligned} B(x,y) &= 1 \quad ,if \quad R(x,y) < t \\ B(x,y) &= 0 \ ,else \end{aligned} \tag{1}$$

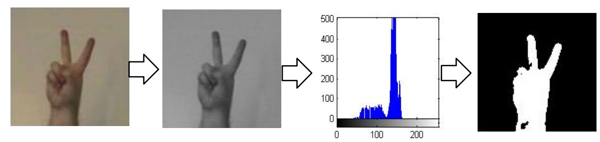


Figure 2: Steps of Pre-processing

After the formation of binary image, the next step is the application of morphological operations to improve the quality of the silhouette. The application of these morphological operations ensures that all the discontinuities from the binary image are removed. For these operations, the size of the structuring element taken is very small because of the low resolution of the image which makes the information content of the image in very small regions. Next, we generate of the mask for the binary image.

B. Mask Generation

Mask generation is the most critical step in the proposed method for the gesture recognition. The mask for the binary image, $B(x,y)$ is calculated based on the geometrical structure of the binary image.

The centroid of the silhouette represents the center of mass for the silhouette. It is calculated by the image moment as in [9],

$$G_{p,q} = \sum_i \sum_j i^p j^q V(i,j) \tag{2}$$

Where, $V(i,j)$ is the intensity of the coordinate (i,j) and the centroid (α,β) is given as:

$$\begin{aligned} \alpha &= \frac{G_{1,0}}{G_{0,0}} \\ \beta &= \frac{G_{0,1}}{G_{0,0}} \end{aligned} \tag{3}$$

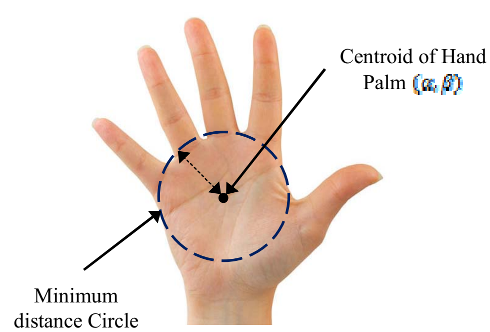


Figure 3: Hand Palm showing the minimum distance calculations

Once the centroid (α, β) of the binary image, $B(x, y)$ is obtained, the smallest distance of the boundary of the silhouette from this centroid point is calculated as per (4). This minimum distance is taken in all directions from the centroid and the area thus formed represents the mask for the binary silhouette. This area usually covers the palm region of the hand and thus forms the base of the mask. Further, morphological operations are performed on the mask thus obtained to incorporate and small section near the boundary of the silhouette.

$$\lambda = sqrt(|x_s - \alpha|^2 + |y_s - \beta|^2) \quad (4)$$

From (4), a set of distances is computed where (x_s, y_s) is the surface coordinate of palm and the minimum distance is used to mark a circle which helps in the formation of mask.

Let $M(x, y)$ represent the mask for the binary image, B(x, y). A sample mask obtained is shown in fig. 4.

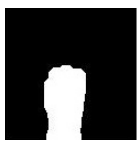


Figure 4: The mask generated for the gesture shown in fig. 2.

C. Mask Subtraction

Once the mask is obtained the next step is the extraction of the remaining information from the binary image. This is done by subtracting the mask obtained in previous step from the binary silhouette of the gesture.

$$S(x, y) = B(x, y) - M(x, y) \quad (5)$$

here, $S(x, y)$ is the subtracted image. Now, all the information about the hand gesture is contained in the binary image, $S(x, y)$.

Figure 5: Silhouette left after the removal of mask.

As it can be seen, the binary image, S(x, y) still requires some further processing before the count of fingers can be obtained from it.

D. Post-processing

Post-processing is required to remove all small unwanted regions present in the binary image. For this morphological operations are applied to the image to enhance it information content.

Figure 6: The final binary image after the processing for the gesture recognition.

Fig. 6 shows the binary image after the application of mask subtraction and morphological operations.

E. Gesture Recognition

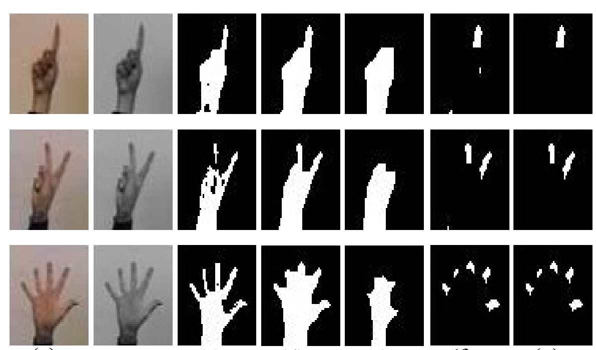
The final step is the recognition of the hand gesture i.e. the count of the number of fingers raised by the person in the image.

From the fig. 6, it can be observed that the number of fingers raised by the person is the number of white objects present in the binary image, $F(x, y)$. To count the number of white objects present in the binary image the region properties of the image are calculated. This returns the number of white objects present in the image which is equal to the number of raised fingers. Hence, the hand gestures are recognized from the low resolution images.

IV. Experimental Results

The proposed method is tested on the publically available dataset of Marcel-Triesch. The standard dataset of Marcel-Triesch has hand gestures with a resolution of $[50 \times 50]$ pixels. The dataset also consists of many gestures which have are based on the shape of hand and thus are not considered for the experimentation results. The gestures that are used for the recognition are one, two and all five fingers raised. A similar concept for the recognition can be applied on dataset with three or four fingers raised.

Fig. 7 shows three sample cases, one each for one, two and five fingers raised from the Marcel-Triesch dataset to explain the working of the proposed algorithm step by step. Fig. 7(a) shows the original low resolution image. Subsequent images in the sequence show the various stages of the processing algorithm as explained above.



(a) (b) (c) (d) (e) (f) (g)

Figure 7: Different stages of the proposed method are shown pictorially in the figure.

The proposed method is tested for accuracy on the marcel-Triesch dataset considering only the relevant gestures that are one finger raised, two fingers raised, and five fingers raised. The results obtained on the low resolution images of the dataset are shown in table1.

The average accuracy is calculated as ratio of Total number of successful cases to Total number of cases in the experimental calculations. This is obtained for the standard dataset of hand gestures and is found to be 97.47% which is considered as a high recognition rate. This recognition rate is comparable to the hand gesture recognition rate of other algorithms.

TABLE I. EXPERIMENTAL RESULTS ON STANDARD DATASET

Hand Gesture	**Gestures Under Consideration**		
	One Finger Raised	***Two Fingers Raised***	***Five Fingers Raised***
Total No. of Hand Gestures	65	57	76
Correct Detection	64	55	74
False Detection	1	2	2
Accuracy	98.46%	96.49%	97.37%
Average Accuracy : 97.47%			

A. Comparison with the other techniques

The accuracy obtained by the application of our method is compared with the accuracy of previously proposed methods for the recognition of hand gestures.

TABLE II. STATE-OF-ART COMPARISON OF ACCURACY OF DIFFERENT METHODS

Methods	**Accuracy**
S. Marcel [10]	93.70%
D.K. Vishwakarma [9]	93.50%
A. Just [11]	92.79%
Our Method	**97.47%**

Table 2 shows the comparison of various methods proposed previously for the hand gesture recognition. These methods are not necessarily applied on low resolution images and thus the accuracy of 97.47% of our system is a great advantage as this is modifiable for higher resolution images also.

V. FUTURE WORKS

The proposed method recognizes the hand gesture from very low resolution images. The method is made robust to both the lighting conditions as well as the orientation of the hand gesture by taking into consideration the geometrical structure of hand. Hand gestures recognized by this method are the number of fingers raised by the person i.e. one, two, three, four or five fingers. Further work can be done in this area to accommodate even a larger number of gestures which considers the shape of hand or the motion trajectory of hand.

This idea of the recognition of the hand gesture can be extended further to generate sets of code from multiple images to develop even a larger number of applications for the human machine interaction. A great example of this is presented in [9]. Thus, a system can be easily designed to understand a large number of sequences as required in the applications.

VI. CONCLUSION

With the rapid pace of development and digitalization of the systems, the gesture recognition has gained a huge popularity for the human machine interface. Hand gesture recognition from low resolution images not only saves a huge storage space but also reduces the processing time of the system. This method is particularly lightweight and suitable for scenarios with static background resources. The proposed system has an execution time of less than a sec (approx. 0.865 sec, running on i5 2.50 MHz processor with 4 GB RAM, MATLAB 2012a) which makes it feasible for real time implementation in different systems. The method is robust for the variable illumination conditions and orientation of the hand gesture. Furthermore, the average accuracy of the proposed system is 97.47% which shows a superior performance of our method as compared to the existing techniques used for gesture recognition. This type of systems can be efficiently used for controlling various application like home automation by capturing image from a very far distance such as a corner of the room.

REFERENCES

[1] Zhenyao Mo and Ulrich Neumann, "Real-time Hand pose Recognition Using Low-Resolution Depth Images", Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2006, pp.1499-1505, doi: 10.1109/CVPR.2006.237.

[2] Siddharth S. Rautaray and Anupam Agrawal, "Design of Gesture Recognition System for Dynamic User Interface", Proc. IEEE International Conference on Technology Enhanced Education, 2012, pp 1-6, doi: 10.1109/ICTEE.2012.6208628.

[3] Lae-Kyoung Lee, Su-Yong An and Se-Young Oh, "Robust Fingertip Extraction with Improved Skin Color Segmentation for Finger Gesture Recognition in Human-Robot Interaction", Proc. IEEE World Congress on Computational Intelligence, June 2012, pp.1-6, doi: 10.1109/CEC.2012.6256140.

[4] Paulo Trigueiros, Fernando Ribeiro and Luis Paulo Reis, "A comparison of machine learning algorithms applied to hand gesture recognition", Proc. 7th Iberian Conference on Information Systems and Technologies, pp 1-6, June 2012.

[5] Nobuhide Okabayashi, Hideya Takahashi, Tahito Aida and Hiromitsu Hama, "High-Resolution Image Restoration from Low-Resolution Images, Preprocessing for Pattern Recognition", Proc. Second International Conference on Innovative Computing, Information and Control, Sept. 2007,pp. 312, doi: 10.1109/ICICIC.2007.335.

[6] Md Iqbal Quraishi, Krishna Gopal Dhal, J Paul Choudhury, Kamal Pattanayak and Mallika De, "A Novel Hybrid Approach to Enhance Low Resolution Images Using Particle Swarm Optimization", Proc.

2nd IEEE International Conference on Parallel, Distributed and Grid Computing, pp. 888-893, 2012, doi: 10.1109/PDGC.2012.6449941.

[7] Myung-Cheol Roh and Seong-Whan Lee, “Gesture Detection in Low-Quality Video”, Proc. 18th International Conference on Pattern Recognition, 2006, pp.791-794, doi: 10.1109/ICPR.2006.607.

[8] Meenakshi Panwar, “Hand Gesture Recognition based on Shape Parameters”, Proc. International Conference on Computing, Communication and Applications, pp. 1-6, Feb. 2012, doi: 10.1109/ICCCA.2012.6179213.

[9] D.K. Vishwakarma and Rajiv Kapoor, “Simple and intelligent system to recognize the expression of speech-disabled person”, Proc. 4th International Conference on Intelligent Human Computer Interaction, Dec. 2012, pp. 1-6, doi: 10.1109/IHCI.2012.6481804.

[10] Sebastien Marcel, “Hand Posture Recognition in a Body-Face centered Space”, Proc. International Gesture Workshop, 2009.

[11] Agnes Just, Yann Rodrigues and Sebastien Marcel, “Hand Posture Classification and Recognition using the Modified Census Transform”, Proc. of 7th International Conference on Automatic Face and Gesture Recognition, April 2006, pp. 351-356, doi: 10.1109/FGR.2006.62.