International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347 - 6435 (Online)) Volume 4, Issue 10, October 2015)

Vision Based Hand Gesture Recognition For Real Time Home Automation Application

Sagar Shimpi[1], Dashrath Mundkar[2], Rishikesh Shinde[3], Prajakta Nivangune[4], Priyanka Lanjewar[5], Asst. Prof. A. R. Sonawane[6]

Department of Computer Engineering, NBN Sinhgad School of Engineering, Pune
Savitribai Phule Pune University, Pune

***Abstract*— We are proposing a fast algorithm for automatically identifying a limited set of gestures from hand images for home automation purpose. Gesture recognition is a challenging problem in its normal form. We are going to consider a stable set of manual commands and a reasonably organized environment, and develop a basic, so far effective, method for hand gesture recognition. Our methodology contains steps for segmenting the hand region, locating the fingers, and finally classifying the gesture. The algorithm is constant to rotation, translation and scale of the hand gesture. Now We describe the efficiency of the technique on real imagery.**

***Keywords*— detection, hand gesture, home automation, recognition, robot control.**

I. Introduction

Vision-based automatic hand gesture recognition has been a very industrious research area in recent years with interesting applications such as human computer interaction (HCI), robot control, home automation, and sign language interpretation. The overall problem is bit difficult due a number of problems including the complex nature of static and dynamic hand gestures, complicated backgrounds, and obstructions. Attacking the problem in its broad view needs to elaborate algorithms requiring demanding computer resources. What encourages us for this work is a real time home automation problem, in which we are interested in operating a robot by hand pose signs given by a human. Due to real-time operational requirements, we are interested in a computationally proficient algorithm.

Early methodologies to the hand gesture recognition problem in a home automation context involved the use of markers on the finger tips [1]. An associated algorithm is used to determine the presence and color of the markers, through which one can identify which fingers are active in the gesture. The inconvenience of insertion markers on the user's hand makes this not a realistic approach in practice.

Latest methods use more innovative computer vision techniques and don't require markers. Hand gesture recognition is applied through allocating skin color tone range to the input assembler and Hand gesture acknowledgement is performed through a curve space technique [2], which implicates defining the boundary curves of the hand. This is a vigorous method that is scale, translation and rotation invariant on the hand pose, so far it is computationally challenging. In [3], a vision-based hand pose recognition method using minimum images is proposed, in which a multi-system camera is used to pick the midpoint of gravity of the hand and points with farthest distances from the center, as long as the locations of the finger tips, which are then cast-off to obtain a skeleton image, and lastly for gesture recognition. Computer vision tools used for 2D and 3D hand gesture recognition comprised of particular mappings architecture [4], principal component analysis [5], Fourier descriptors, neural networks, orientation histograms [6], and particle filters [7].

Main focus is on detecting stable set of manual guidelines by a robot, in a practically organized atmosphere in real time. Our approach involves fragmenting the hand based on skin color statistics, as well as scope constraints. We then find the center of gravity (COG) of the hand region as well the farthest point from the COG. Our algorithm is invariant to rotations, translations and scale of the hand. Furthermore, the procedure does not need the storing of a hand gesture record in the robot's memory. We demonstrate the effective and productive methodology on real images of hand gestures.

II. Block Diagram

An easy way to comply with the conference paper formatting requirements is to use this document as a template and simply type your text into it.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347 - 6435 (Online)) Volume 4, Issue 10, October 2015)

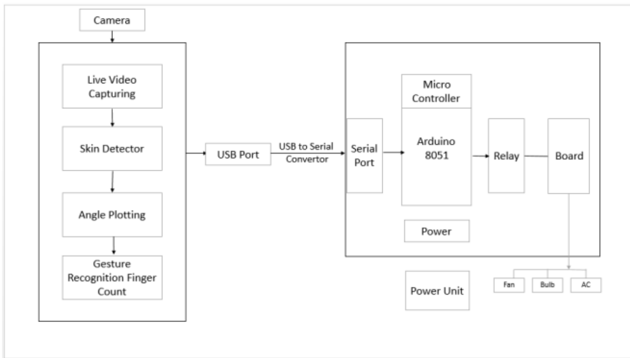

Figure I: Block Diagram

As shown above, a camera is connected externally to the computer system. Connection between computer system and controller board is successfully done via USB to Serial Convertor. Also, a power unit must be connected to controller board for power supply. Finally, home appliances will be connected to the micro-controller with the help of relay.

III. Hand Gesture Recognition

Consider a robot navigation problem, in which a robot responds to the hand pose signs given by a human, visually observed by the robot through a camera. This algorithm enable the robot to detect a hand pose sign in the image, as one of five possible command. The recognized command will then be used as a regulator input for the robot to accomplish a certain set of action or execute certain tasks. The hand poses could be linked with several significances reliant on the function of the robot. For instance, a "one" count could mean "move reverse", a "five" count could mean "fast". Furthermore, "two", "three", and "four" counts could be interpreted as "forward", "turn right," and "turn left."

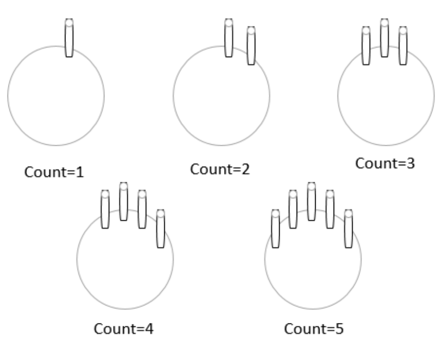

Figure II: Set of Hand Getures

Our projected technique of hand gesture recognition consists of the next steps:

- Localizing hand-like regions based on skin color [8], forming a Black and White image output.
- Performing region-based fragmentation of the hand, excluding small regions that were declared as "hand-like," based on their color statistics.
- Calculating the center of gravity (COG) of the hand region as well as the farthest distance in the hand region from the COG.
- Producing a circle centered at the COG that intersects all the fingers that are active in the count.
- Withdrawing a 1D binary signal by following the circle, and classifying the hand gesture based on the number of vigorous regions (fingers) in the 1D signal.

In the following subclasses we define each of the steps stated above:

A. Localizing Hand-like Regions by Skin Detection:

We shoulder that the part of the scene about the hand has previously been extract. Then our first chore is to divide the hand in image from the background. We secure that goal in two phases. Firstly, we find the pixels in the image that are probable to have its place to the hand region, which we mention in this section. Then we simplify that result, as we describe in the next section.

It has been detected that the Red/Green (R/G) ratio is differential characteristic feature for human skin color [8]. Our statistical analysis also uphold this claim. From this we spot that the R/G ratio lies within a thin band of values for skin pixels, however it is much more variable for non-skin pixels. Hence, we could use this ratio to determine whether a pixel belongs to the hand region or not. In precise, we pragmatically notice that the following two thresholds fruitfully capture hand-like intensities:

$$\boxed{\mathbf{1.05 < R / G < 4.00}}$$

We set all the pixels with color intensities within the thresholds to one, and all the rest to zero, with the help of this thresholding scheme; resulting in a black and white image output. Of course, this simple scheme could harvest many erroneous decisions, for example many contextual pixels having skin-like colors could be classified as "hand-like." We refine this output in the next section.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347 - 6435 (Online)) Volume 4, Issue 10, October 2015)

B. Fragmentation and False-Region removal:

The design described in the previous section could produce many disconnected areas in the image classified as hand-like. We use concepts from region-based fragmentation to reduce this problem. Our presumption is that the largest connected white region belongs to the hand. So we use a comparative region size threshold to remove the undesired areas. Specifically, remove the regions that contain less number of pixels than a threshold value. The threshold value is chosen as 25% of total number of pixels in the white regions. Note that this is an image-size unchanging scheme. The ideal outcome is the fragmented hand region.

C. Discovering the Centroid and Farthest Distance:

Given the fragmented hand region, we calculate its centroid, or center of gravity (COG), (x, y), as follows:

$$\overline{x} = \frac{\sum_{i=0}^{k} x_i}{k} \quad and \quad \overline{y} = \frac{\sum_{i=0}^{k} y_i}{k}$$

Where xi and yi are x and y coordinates of the ith pixel in the hand region, and k indicates the number of pixels in the region. After COG is achieved, we compute the distance from the farthest point in the hand to the center; generally this extreme distance is the distance from the centroid to tip of the lengthiest active finger in the particular gesture (see Figure 3).

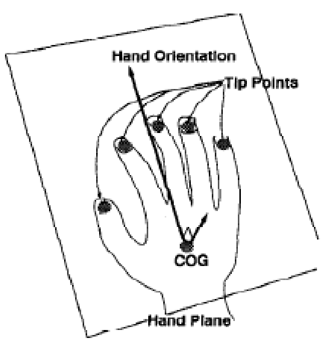

Figure III: Centre of Gravity and farthest points of the hand.

D. Generating a Circle:

We draw a circle whose diameter is 1.4 of the farthest distance from the COG. Such a circle is likely to coincide all the fingers active in a specific gesture or "count." Just to provide a visual flavor.

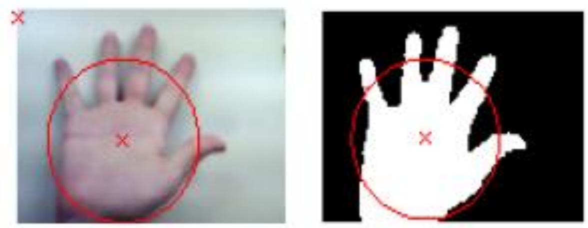

Figure IV: Illustration of the execution of the processing steps described in sample image

E. Withdrawing a 1D Signal and Classification:

We now extract a 1D binary signal by tracking the circle constructed in the previous step. Ideally the uninterrupted "white" portions of this signal belongs to the fingers or the wrist. By counting the number of zero-to-one (black-to-white) transitions in this 1D signal, and subtracting one (for the wrist) leads to the estimated number of fingersre active in the gesture. Determining the number of fingers leads to the detection of the gesture.

Note that our algorithm just counts the number of active fingers without regard to which fingers are acting. For example, Figure 5 shows three dissimilar ways in which our algorithm would discover three counts; rotation, orientation, or any other combination of three fingers would also give the same result. So user does not have to recall which three fingers he/she needs to use to articulate the "three count." While this feature may be desirable in some chores, in other tasks one might be interested in associating different meanings to different finger combinations. We could revise and adapt our algorithm to such a setting by different modifications.

F. Scale, Rotation, and Translation Invariance:

Our offered algorithm is not reliant on on scale. It means that actual size of the hand and its distance from the camera does not affect interpretation. It is rotation invariant, since the orientation of the hand does not hamper the algorithm from detecting the gesture. In addition, the hand position is also not a problem.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347 - 6435 (Online)) Volume 4, Issue 10, October 2015)

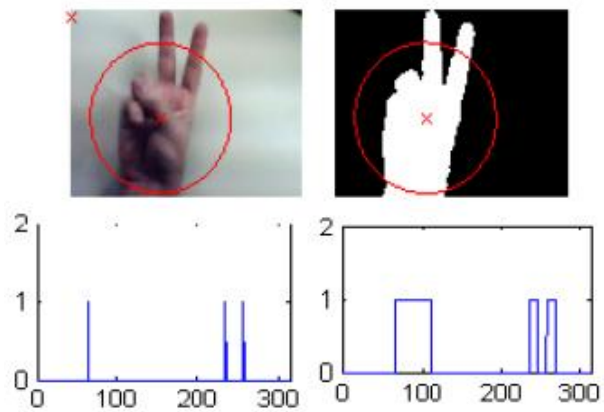

Figure V: Sample result for a "two" count. Top left: Original image and the circle constructed. Top right: Segmented hand and the circle constructed. Bottom left: Zero-to-one transitions in the 1D signal extracted. Bottom right: The 1D signal itself.

IV. Experimental Results

We have conducted experiments based on images, acquired using 5 Mega-Pixel simple webcam. We have collected these data on uniform as well as disordered backgrounds.

The number of these transitions minus one (for the wrist) produces the estimated count. Also, the computation time needed to obtain these results is very small, since the algorithm is quite simple. We have noted that images taken under inappropriate light (especially using the webcam) have led to the false results. In these cases the failure mainly stems from the flawed fragmentation of some background areas as the hand region. Our algorithm appears to perform well with somewhat complex backgrounds, as long as there are not too many pixels in the background with skin-like colors. Overall, we find the performance of this simple algorithm quite satisfactory in the context of our interesting and home automation applications.

V. Conlcusion

We proposed efficient algorithm for a hand gesture recognition problem of given detected images of the hand, the hand regions are sectioned, and then inference is made on the movement of the fingers involved in the gesture.

We have to show the efficiency of this well-organized algorithm on real images we have developed.

Based on our inspiring home automation applications, we have only considered a limited set of gestures. Our algorithm can be stretched in a number of techniques to classify a wider set of gestures. The dissect portion of our algorithm is too simple, and would need to be enhanced if this technique is to be used in challenging operating conditions. However we should note that the separation problem in a general setting is an open research problem itself.

Acknowledgement

We would like to thank all the researchers working on this field who in one way or another guided us on achieving our goals. We would also like to thank all the professors at NBN Sinhgad School of Engineering (Savitribai Phule Pune University, Pune) who were kind enough to share their views and offered some suggestions in making this project success.

Refernces

[1] J. Davis and M. Shah "Visual Gesture Recognition", IEE Proc.-Vis. Image Signal Process., Vol. 141, No.2, April 1994.

[2] C.-C. Chang, I.-Y Chen, and Y.-S. Huang, "Hand Pose Recognition Using Curvature Scale Space", IEEE International Conference on Pattern Recognition, 2002.

[3] A. Utsumi, T. Miyasato and F. Kishino, "Multi-Camera Hand Pose Recognition System Using Skeleton Image", IEEE International Workshop on Robot and Human Communication, pp. 219-224, 1995.

[4] R. Rosales, V. Athitsos, L. Sigal, and S. Sclaroff, "3D Hand Pose Reconstruction Using Specialized Mappings", IEEE International Con. on Computer Vision, pp. 378- 385, 2001.

[5] C. Tomasi, S. Petrov, and A. Sastry, "3D = Classification + Interpolation", IEEE International Conf. on Computer Vision, 2003.

[6] W. T. Freeman and M. Roth, "Orientation Histograms for Hand Gesture Recognition", IEEE International Conf. on Automatic Face and Gesture Recognition, 1995.

[7] L. Bretzner, I. Laptev, and T. Lindberg, "Hand Gesture Recognition using Multi-Scale Color Features, Hierarchical Models and Particle Filtering", IEEE International Conf. on Automatic Face and Gesture Recognition, 2002.

[8] J. Brand and J. Mason, "A Comparative Assessment of Three Approaches to Pixel-level Human Skin Detection", IEEE International Conference on Pattern Recognition, 2000.