Hand Gesture Segmentation and Recognition Based on Graph Cut

R. Keerthana, Dr. S. Mythili
1PG Student, 2Associate Professor
1,2Department of Electronics and Communication Engineering
1,2PSNA College of Engineering and Technology

Abstract

Hand gesture recognition system’s performance and accuracy are increased by one of the pre-processing techniques, i.e., gesture segmentation. In some real-time applications, due to low accuracy in gesture matching process used to lead unexpected response. Earlier days, contour model based hand gesture recognition, accelerometer based hand gesture recognition and scene classifications based on the contextual semantic information of an image are the recognition techniques used in hand gesture recognition system. Here, these techniques arise many problems. They are lagging in focus on detecting hands, some of them directly use marker based motion capture devices, it cannot provide a rotation and scale invariance, it does not provide the hand parts segmentation, and it capture the same gestures and manually labelled them for each subject. These drawbacks can be overcome by implementing the new image segmentation algorithm. The proposed algorithm uses binaryzation and K Nearest Neighbor (KNN) algorithms. The binaryzation used for background subtraction. And KNN classifier for classifying the hand features. Here, a sample image is processed with the graph cut algorithm. And then the parameter is observed. The parameter has been compared with parameter observed in CSS (Curvature Scale Space) algorithm. Finally, the result shows that graph cut algorithm gave the better accuracy than CSS algorithm. In real-time application, the obtained contour descriptor will be matched with contour descriptor in the data base. Once it matched, then it will trigger an event to drive an application.

Keyword- Gesture Segmentation, Binaryzation, K nearest Neighbour (KNN)

I. INTRODUCTION

Gesture recognition is the mathematical interpretation of a human motion by a computing device. Gesture can be recognised in many ways. They are facial recognition, voice recognition, eye tracking and lip movement recognition are the elements referred as a perceptual user interface (PUI). The aim of PUI is to improve the efficiency and make it more easy to use.

In many applications, gestures are used as input commands. Recognizing gestures make computers to be more accessible for the physically-impaired and makes interaction between two persons. It is used in a gaming or 3-D environment. Hand gesture recognition is gaining great importance in Human Computer Interaction (HCI) and Human Robot Interaction (HRI). Different approaches have appeared making use of different sensors and devices. Hand wearable devices such as sensor gloves have been used although they are usually expensive and user intrusive. Other less intrusive wireless devices like the sensing rings have appeared to overcome these drawbacks. Cameras and computer vision have proved to be useful tools for this task. In addition, other contact free sensors have emerged lately to detect hand motion and interact with different devices. [1]

In this paper, hand gesture segmentation experiment can be considered to be comprised of three layers: detection, tracking and recognition. The detection layer is responsible for dinning and extracting visual features that can be attributed to the presence of hands in the head of view of the camera. The tracking layer is responsible for performing temporal data association between successive image frames. Moreover, in model based methods, tracking also provides a way to maintain estimates of model parameters, variables and features that are not directly observable at a certain moment in time.

Fig. 1: Hand Segmentation and Feature Extraction
Last, the recognition layer is responsible for grouping the spatiotemporal data extracted in the previous layers and assigning the resulting groups with labels associated to particular classes of gestures. Exact shape of the hand obtained lead to good feature extraction and fast and powerful results from the graph cut algorithm and K Nearest Neighbour (KNN) algorithm. This paper is organized as follows:
1) In section II we describe the graph cut segmentation and we present the modules of pre-processing technique.
2) In section III we conclude the paper with results.

II. METHODOLOGY

The proposed system overcomes the drawback on the image segmentation. This condition is achieved by using new algorithm for image segmentation process, so that proposed algorithm may get better performance and accuracy during contour matching and classification process.

Image segmentation has come a long way. Using just a few simple grouping cues, one can now produce rather impressive segmentation on a large set of images. Behind this development a major converging point is the use of graph based technique. Graph cut provides a clean, flexible formulation for image segmentation. It provides a convenient language to encode simple local segmentation cues, and a set of powerful computational mechanisms to extract global segmentation from these simple local (pair wise) pixel similarity. Computationally graph cuts can be very efficient.

Graph based segmentation in four topics:
- General graph cut framework for image segmentation.
- Normalized Cuts, Typical Cuts, and Min Cuts; Image statistics.
- Grouping cues: intensity, texture.
- Data human image segmentation and segmentation benchmark

A. Graph cut
Let consider an undirected graph is represented as $G = < V \ E >$, here $V$ represented as vertices and $E$ represented as graph edge, which connects the every two neighbour vertices.

![Fig. 2: Illustration of s-t graph](image)

The vertex is comprised of two different type nodes (vertices). One type of vertices called neighbourhood nodes, which correspond to the pixels and other type is called terminal nodes. These terminal nodes have another two kinds of nodes named source (s) and sink (t). A Graph consists of s-t called s-t graph. Fig 2 represents the image pixels correspond to the neighbour nodes in the graph (except s and t nodes). The solid lines in the graph are n-links and the dotted lines are t-links.

B. Graphical Representation of an Image
In an image, set of points of the feature space represented as a weighted undirected graph $G = (V \ E)$, the points of the feature space are the nodes of the graph. Make edges between every pair of nodes and determine the weight of each edge. This weight of each edge is denoted as $w (i, j)$ and also it is a function of the similarity between the nodes $i$ and $j$. Then, partition the set of vertices into sets, where similarities within the sets is high and across the set is low. Therefore segmentation is equivalent to graph partition. Cut in a graph is a set of edges whose removal, disconnects the graph. A minimum cut is a cut with a minimum number of edges.

In the graph, $\text{Cut}(A, B) = \sum \ w (u, v) \ and \ u \in A, v \in B$ where &are two partitioned disjoint sets. A highly connected sub graph (HCS) is an induced sub graph $H$ in $G$ such that $H$ is highly connected.
By min Cut method, the graph is partitioned into clusters. Each cluster is considered as an image segment. Min Cut method uses the HCS (Highly Connected Sub graphs) Algorithm to find the clusters.

C. Modules
The matching process composed four modules followed as 1.pre-processing 2.Background subtraction 3.feature extraction 4.Classification.

1) Pre-Processing
Pre-processing is the process to prepare the image for further process. Then change the size of the image for our desire level, it helps to easily extract the features. And then remove the RGB (Red, Green, and Blue) in sign image it helps to easily identify the features in image. Use the Gaussian filter is used to remove noise and smooth the input images.

2) Background Subtraction
The background subtraction is removing the background in the image. It helps to extract the feature only the sign in image. The background subtraction is done by the binaryzation method. In the binaryzation the image is change logical if the foreground changes to 1 and background is changes to 0. Using this binary image can be extracted the background in gray scale image. In the background subtraction the sign is only get the values in image then the background get zeros.

3) Feature Extraction
Identify the unique of image. The Principle component of analysis is the using good algorithm for extract the features. It extracts the surface of the features in image, that surface features give the details of the sign shapes. It is helps to identify the unique of the different sign.

4) Classification
The classification process to analysis feature and give the better result for identify the word in sign. The K Nearest Neighbour is the best classifier for classify the features. The KNN (K Nearest Neighbour) is trained by train features and class, the train features is the group of sign for analysis the result. The class represents the features for specify sign. Then the test feature is classified by the train feature and give the specify class value it helps to identify the correct word from the sign.
III. RESULTS AND CONCLUSIONS

In this, a RGB image is given to the module section for further application. It has many stages. Initially, a RGB image is given as an input. And it sends to the pre-processing modules. There, using background subtraction technique skin colored region is segmented from the background.

<table>
<thead>
<tr>
<th>PARAMETER</th>
<th>VALUES</th>
</tr>
</thead>
<tbody>
<tr>
<td>CORRECT RATE</td>
<td>0.9785</td>
</tr>
<tr>
<td>ERROR RATE</td>
<td>0.0215</td>
</tr>
<tr>
<td>SENSITIVITY</td>
<td>0.09677</td>
</tr>
<tr>
<td>SPECIFICITY</td>
<td>1</td>
</tr>
<tr>
<td>POSITIVE LIKELIHOOD</td>
<td>0.9957</td>
</tr>
</tbody>
</table>

*Table 1: analysed parameter values using graph cut algorithm*

Then wrist region, which allows triggering an event, has been separated.

To develop the contour for corresponding image, the wrist region converted into binary image. For that binaryzation technique has been used. Using canny algorithm the edge has been detected. It creates contour and finally hand features extracted.

The extracted features called test features, matched with trained features. This project has a set of gestures contained 500 gestures. Thus the accuracy for matching with data set of 500 gestures is 97.8%.

REFERENCES