I J C T A, 9(37) 2016, pp. 23-28

An Efficient System for Assisting the Elderly People

Saira Chaithra Jacob* and L.N.B Srinivas**

Abstract: Gestures are visible bodily actions that convey a particular message which is in a non-verbal form of communication instead of speech. Human–computer interaction (HCI) focuses on the interfaces between users and computers. HCI research designs and use of computer technology by observing various ways of interaction between human and computers and design technologies that allow humans to interact effectively with the computers. This project investigates in creating an assistance system that assists the elderly for indicating their basic necessities such as the need to use the toilet, requesting for food or water, or need emergency assistance of a doctor. The system is developed in MATLAB platform with a windows application. Image frames are extracted at a certain time delay from the video. The entire video is converted into a number of image frames and processed further. Then the image frame is converted into color space model and the threshold value for that image is selected. The hand region is segmented by using superpixel segmentation. The effortlessness of our approach makes it exceptionally simple to use. The Connected Component Analysis Technique is used to identify the gesture indicated by the user.

Keywords: Gesture recognition; Superpixel segmentation; Human Computer Interaction; Connected Component Analysis Technique.

1. INTRODUCTION

Gesture-based communication is a very common and widely used form of interaction in human societies. Gestures support us to interact with other people and objects as well as a substitute every other form of communication like communicating with the deaf people. On the other hand, computers have become an inseparable part of our society, influencing many aspects of our daily lives for communication and interaction. There have been many psycholinguistic studies, trying to describe and analyze human hand gestures. Every gesture is the physical expression of the mental concept. Webster dictionary definition for gestures as: "… the use of motion of the limbs or body as a means of expression; a movement usually of the body or limbs that express or emphasizes an idea, sentiment, or attitude". In general, gestures can be conceived as a non-verbal form of communication and expressions of emotions and information.

Hand gestures have a wide variety, depending on the context. Thus, several categorizations can arise such as conversational, controlling, manipulating and communicative gestures. Sign languages for deaf or a navigation gesture "go there" are examples of communicative and controlling gestures, respectively. Communicative and controlling gestures are the classes of gestures that research is mainly turned to, as vision-based recognition systems can efficiently help. A general problem faced by using computer vision is to build a robust system that endures imprecise and vague results of lower level modules. Existing recognition methods require either the position of the gesturing hand or the start and end frames of each gesture as an input. These requirements make it difficult to deploy gesture recognition systems in many real-world scenarios as the availability of the requirements is unrealistic. The presence of skin-colored objects and background motion increases the problem of localizing the hand. Moreover, detecting the start and end of a gesture automatically from a continuous sequence of images (video), automatically is a challenging problem.

* Student, M.Tech Database Systems, Department of Information Technology, SRM University, Kattankulathur-603203 Chennai, India. *Email: sairachaithrajacob@gmail.com*

** Assistant Professor (S.G), Department of Information Technology, SRM University, Kattankulathur-603203, Chennai, India. *Email: srilnb@gmail.com*

2. RELATED WORK

Gesture analysis has become a very widely used concept in today's world. This system focuses on hand gesture analysis which uses two main modules ie image segmentation and gesture detection. Hand gesture has become a significant part of the interaction of communication. Gesture interaction in earlier systems used additional pointing devices such as data gloves and markers. These sensor-based devices were used in order to perceive movement and to provide accurate measurements of hand pose but they require extensive standardization, are costly and restrict ordinary hand movement. Locating and segmenting the hand from the background is usually a very difficult task. Factors that affect this include fast motion, lighting variances, or the presence of similar objects. [1] provides algorithms to detect hand on a plain, uniform and complex background. But it requires a constant background which is static. The recent systems do not require any physical attachments. Gestures are performed by hand freely in 3D space captured by various cameras which are analyzed and recognized with video-based solutions. Our system uses a webcam to get real time videos from which frames are generated for further analysis. Various algorithms based on vision-based hand gesture recognition has been proposed in the previous years which provides robust and reliable systems, these algorithms are reviewed in [2]. Superpixels are widely used for image segmentation. There are various projects based on superpixels. [3] uses Kinect depth camera that provides depth and skeleton information and superpixel earth mover's distance that measures the dissimilarity between hand gestures. [4] Introduces an algorithm that clusters pixels and generates uniform superpixels. Our system uses this algorithm which is fast, easy to use and produces high-quality segmentation. Due to over segmentation done using super pixel we can lose meaningful image edges, hence, our system includes morphology in order to enhance the image [5]. Once the segmentation is done the next step is to analyze the gesture. [6] discusses various techniques and algorithms related to the gesture recognition. [7] provides connected component algorithms used to extract regions which are not separated by a boundary. Our system uses this algorithm to find the number of fingers shown by the user.

3. PROPOSED SYSTEM

Human–computer interaction (HCI) focuses on the interfaces between users and computers. One of the most crucial steps in any hand gesture recognition system is the hand segmentation. All the steps that follow the hand segmentation depend on the quality and accuracy of segmentation. Improper segmentation can lead to losing of valuable data which cannot be recovered later.

A. Image Acquisition

In the proposed system a continuous input of the live video is streamed. Image frames are extracted at a certain time delay from the video. The entire video is converted into a number of image frames and processed further.

B. Pre-processing

The image frames acquired is then preprocessed.

Imaging applications commonly require conversion of images between classes of data and image type. Image Processing Toolbox provides functions in order to convert between data classes, including signed or unsigned 8-, 16-, and 32-bit integers and single- and double-precision floating point. Along with these functions it also provides algorithms in order to convert between the types of images including grayscale, binary, true color, and indexed color. It also supports various color spaces (such as YCrCb, YIQ and HSV) as well as high dynamic range images and Bayer pattern encoded.

In this step, we convert the image to unsigned 8-bit integers. And reduce the noise present in the image to get an enhanced clear image of our gesture.

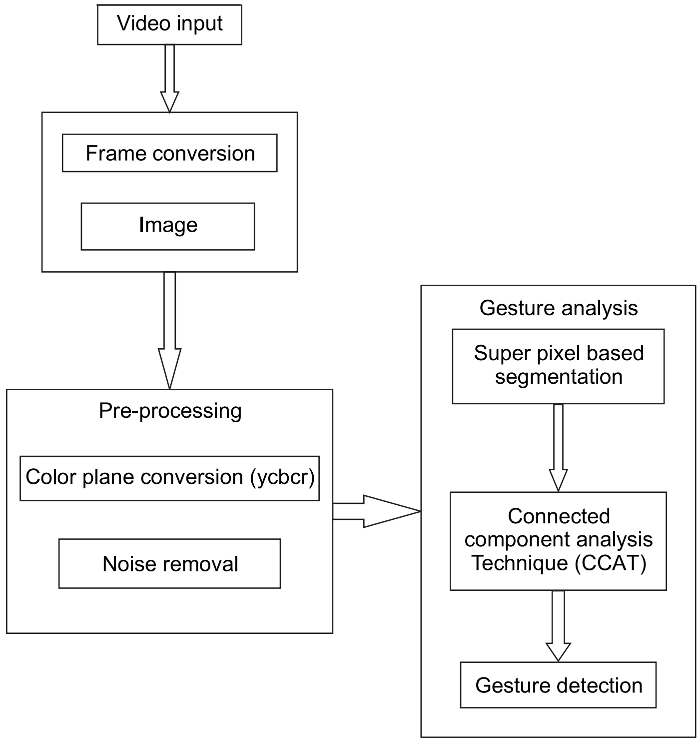

Figure 1: Architecture diagram

C. Gesture Analysis

There are different ways to interpret a gesture depending on the type of the input data. The 3D coordinate system represents various key pointers. Depending on the relative motion of these key pointers, the quality of the input and the algorithm's approach, the gesture can be detected with high accuracy.

1. *Superpixel based segmentation:* The smallest element of a picture displayed on the screen is called a pixel. Most image processing tasks represent the image pixels as their basic unit. However, they are an outcome of the dissimilar representation of images and not natural entities. Perceptual grouping of pixels i.e. the over segmentation of image result in superpixels. Superpixels contain more information than pixels and align better with image edges than rectangular image patches. Superpixel can increase the performance of the following processes since it replaces hundreds of thousands of pixels into a group of 25 to 2500 super pixels. Superpixels are becoming increasingly popular in many computer vision applications.

   There are various benefits of this segmentation. It is computationally efficient; the image complexity is reduced to a few hundred super pixels from hundreds of thousands of pixels. It is also representational efficient, pairwise constraints between units, while only for adjacent pixels on the pixel-grid, can now model many longer-range interactions between superpixels. The superpixels are perceptually meaningful each superpixel is a perceptually consistent unit, i.e. all pixels in a super pixel are most likely uniform in, say, color and texture. It is near-complete: because super pixels are results of an over-segmentation, most structures in the image are conserved. The loss of data in moving from the pixel-grid to the superpixel map is very little.

   The computational effort for the computation of super pixels and the risk of losing meaningful image edges by placing them inside a super pixel are some of the disadvantages of using superpixel segmentation. If we compare the human-marked segmentation (the ground truth) to the one reconstructed from the superpixel map, we may find that some contour details are lost in the process of over-segmentation.

2. *Morphology:* Morphology contains large number of image processing operations. In morphology the images are processed based on shapes. A same sized output image is formed by applying a structuring element to the input image. A morphological operation compares the pixels in the input image with its neighbors and sets the value of the corresponding pixel in the output. A morphological operation is constructed by choosing the size and shape of the neighborhood, which is sensitive to specific shapes in the input image. In this paper we use two of the common morphological processes dilation and erosion. Pixels are added on to the objects boundaries of the image in dilation whereas pixels are removed from the object boundaries in erosion. Depending on the size and shape of the element used to process the image, the number of pixels is added or removed from the objects in an image. Noise present in segmented output is removed based on morphological operations to obtain a more clear & distinct output.

3. *Connected Component Analysis Technique (CCAT):* Connected component analysis is an algorithmic application of graph theory, where based on a given heuristic, subsets of connected components are uniquely labeled. Even though color images and high dimensional data can be processed, connected-component labeling is widely used to detect connected regions in binary digital images. Blob extraction is generally performed on the resulting binary image from a threshold step.

   CCAT is used to extract regions which are connected after the boundary has been detected. Any set of pixels which is not separated by a boundary is called connected. Each maximal region of connected pixels is called a connected component.

   Connected components labeling initially scans an image and based on the connectivity of the pixel, groups the pixels into components *i.e.* all pixels in a connected component are connected with each other since they share similar pixel intensity values. After determining the groups, each pixel is given a gray level or color label according to the component it was assigned to.

   Many automated image analysis applications depend on extracting and labeling of several disjoint and connected components in an image. Once the connected components are identified & region properties are determined for all objects. The centroid is calculated using region props technique & bounding box are determined for all connected components for continuous tracking of finger movement from live video.

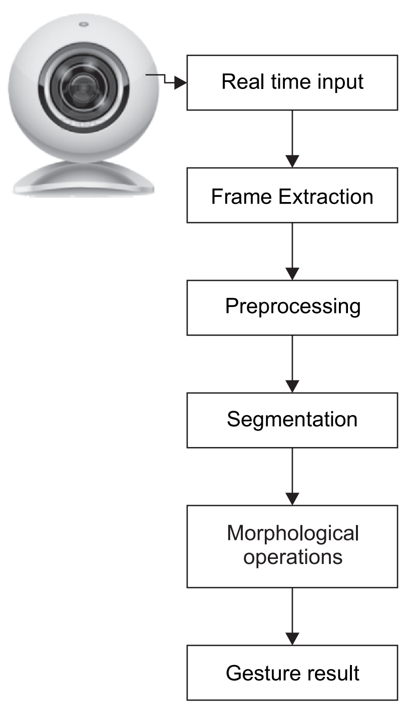

Figure 2: Data-Flow diagram

4. IMPLEMENTATION AND EXPECTED RESULTS

A continuous input of the live video is streamed. Image frames are extracted at a certain time delay from the video. The entire video is converted into a number of image frames and processed further. In the preprocessing step, we perform noise removal and smoothening of the image to get an enhanced clear image of our gesture. Once the image is preprocessed the segmentation of the users hand from its background is done. For this purpose, we include an image segmentation algorithm known as super pixel-based segmentation. The segmented image is then enhanced by performing dilation and erosion processes of morphology. Once the image is enhanced using morphology we perform connected component labeling in order to find the gesture type.

The following explains the connected component labeling algorithm

```
linked = []
labels = structure with dimensions of data, initialized with the value of Background
First pass
  for row in data:
    for column in row:
      if data[row][column] is not Background
        neighbors = connected elements with the current element's value
        if neighbors is empty
          linked[NextLabel] = set containing NextLabel
          labels[row][column] = NextLabel
          NextLabel += 1
        else
          Find the smallest label
          L = neighbors labels
          labels[row][column] = min(L)
          for label in L
            linked[label] = union(linked[label], L)
Second pass
  for row in data
    for column in row
      if data[row][column] is not Background
        labels[row][column] = find(labels[row][column])
  return labels
```

We use two pass algorithm for finding the connected components. In the first pass we give a value to the background and iterate through each element. If an element whose value is not the background is encountered, we find its neighbors. If there is no neighbor present then we label the element. If any neighbor is present then we have to find the smallest neighbor and assign its label to the current element. The neighboring labels store the equivalence. In the second pass we again iterate through each of the element. And we re-label the elements in accordance to the lowest equivalent label. The detailed explanation of this algorithm is given in [7].

Once the gesture is recognized the text is displayed on the screen for the required assistance. We can further connect the MATLAB application to the assistant's android mobile to notify the assistant about the gesture.

5. FUTURE WORK

This technology can not only be used by the elderly but can also be implemented in the hospitals as well as for disabled people. By using a 3-D camera, we can increase the efficiency of the system. We can implement a database to the system to analyze the previous operations of the user which will be useful for medical purposes. It can be used for automated house systems by integrating embedded systems to the gesture results.

6. CONCLUSION

In the methodology "An Efficient System for Assisting the Elderly People" provides an HCI to analyze the gesture of the user and to display the text onto the screen for assistance. Here a real time video is taken and image frames are continuously captured. The resulting image is then preprocessed enhanced segmented and using CCAT we analyze the gesture and find the corresponding assistance required.

References

1. A. A. Randive, H. B. Mali, S. D. Lokhande, "*Hand Gesture Segmentation*", *International Journal of Computer Technology and Electronics Engineering*, Vol. 2, Issue 3, June 2012.
2. G. Simion, V. Gui, and M. Otesteanu, "*Vision-Based Hand Gesture Recognition: A Review*," *International Journal of Circuits Systems and Signal Processing*, Vol. 6, Issue 4, pp. 275-282, 2012.
3. Chong Wang, Zhong Liu, and Shing-Chow Chan, "*Superpixel-Based Hand Gesture Recognition with Kinect Depth Camera*", *IEEE Transactions On multimedia*, Vol. 17, No. 1, pp. 29-39, January 2015.
4. Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Süsstrunk, "*SLIC Superpixels Compared to State-of-the-art Superpixel Methods*", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 34, num. 11, p. 2274 - 2282, May 2012.
5. Megha Goyal, "Morphological Image Processing", IJCST Vol. 2, Issue 4, Oct. - Dec. 2011.
6. R.Pradipa, Ms. S.Kavitha, "*Hand Gesture Recognition – Analysis of Various Techniques, Methods, and Their Algorithms*", *International Journal of Innovative Research in Science*, Engineering and Technology Volume 3, Special Issue 3, March 2014.
7. Tetsuo Asano, Hiroshi Tanaka, "*In-place Algorithm for Connected Components Labeling*", *Journal of Pattern Recognition Research*, 1, pp. 10-22, July 2010.