CS512 F17 Project Proposal

Ronit Rudra
A20379221
Prerna Kapoor
A20371648

October 26, 2017

Abstract

We propose the implementation of Hand Gesture Recognition using only a webcam. The detected gestures would then be used to perform trivial tasks. The challenge would be to accurately detect hand and fingertips. To that effect, background subtraction, skin color detection and convex hull fitting would be used. A set of training instances would be created using the detected features and a classifier would generate a model for the different poses selected. The features would need to be scale and translation invariant.

1 Problem Statement

Gesture recognition systems are an indispensable asset in the field of engineering and entertainment especially in robotics and video games. These systems have become ubiquitous over the years and are slowly replacing tactile hardware. Now, you can wave your hand to move a robot instead of operating a joystick.

The gesture recognition problem is divided into two types: **Hardware Based** and **Vision Based**. We shall be tackling the Vision Based approach as the only input is an image frame. This approach can be implemented in various ways using stereo-vision cameras, cameras with depth sensors (like the *XBOX Kinect*) or with just a generic mono-vision camera.

As mentioned previously, we have chosen to create a gesture recognition framework using only a mono-vision webcam.

The next section describes the process flow of a typical gesture recognition task and elucidates on the method that we shall use.

2 Proposed Solution

Gesture Recognition can be broken down into multiple steps as shown in the flowchart below:

Figure 1: Process Flow
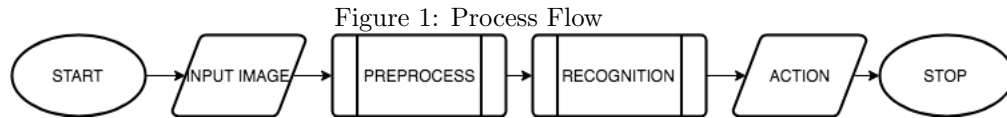
The **Preprocessing** and **Recognition** phases determine how well the system performs. Preprocessing refers to the transformation of the input image to a viable form on which *feature extraction* can be performed. Only when relevant features are available, the recognition system can produce meaningful outputs.

As mentioned in the paper chosen [1], the system is split into the following:

Preprocessing Perform *Additive Background Subtraction* [4] and *Skin Color Detection* [3]

Detection Convert processed image to set of feature vectors using *Haar-like Features* [2] and *Convex Hull Fitting*

Training Use Supervised AdaBoost classification to generate a model for classifying test images containing gestures.

The training data would be collected using the webcam and would be manually labeled from a set of pre-defined labels. A simple script can be written to automate this process.

3 Responsibilities

Ronit :

- Skin Color Detection: Detect only skin color by converting RGB to YCbCr, thresholding and finding ROI by erosion/dilation.
- Convex Hull Features: Extract features by fitting a convex hull/contour fitting and detecting fingers using deformities in the hull.

Prerna :

- Background Subtraction: Write an algorithm to construct a static background image model and extract ROI of a moving hand as foreground. frame.

- Haar Features: Use the Haar Wavelet Transform to generate Haar-features for static and dynamic gesture detection. Use motion history image (MHI) for each dynamic directional hand gesture and design four groups of Haar-like patterns.

Common :

- Data Mining: Collect and label hand gestures using webcam in different settings (occlusion, noise etc.)
- Classification: Design a classifier and improve model generalization using AdaBoost
- Testing: Ensure crash-free program through error/exception handling and thorough unit testing.

References

[1] Ruchi Manish Gurav and Premanand K. Kadbe, *Real time Finger Tracking and Contour Detection for Gesture Recognition using OpenCV*, 2015 International Conference on Industrial Instrumentation and Control (ICIC).

[2] Qing Chen Nicolas, D. Georganas, and Emil M. Petriu, *Hand Gesture Recognition Using Haar-Like Features And A Stochastic Context-Free Grammar*, IEEE ,Vol. 57, No. 8, August 2008.

[3] Amit Kumar and Shivani Malhotra , *Real-time human skin color detection algorithm using skin color map*, 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom), 2015, p2002 - 2006

[4] Lianqiang Niu and Nan Jiang, *A Moving Objects Detection Algorithm Based on Improved Background Subtraction*, 2008 Eighth International Conference on Intelligent Systems Design and Applications, 2008, Vol. 3, p604 - 607