CHAPTER -6

FLOW CHART

Our proposed system solves the problem of hand recognition in the purpose of the interaction with mobile devices. The input of our system is a user's hand gestures taken from the front-facing in real time steaming. The recognition algorithm follows three steps: hand area detection, features extraction and gesture recognition. The flowchart is shown in Figure 6.1 below.

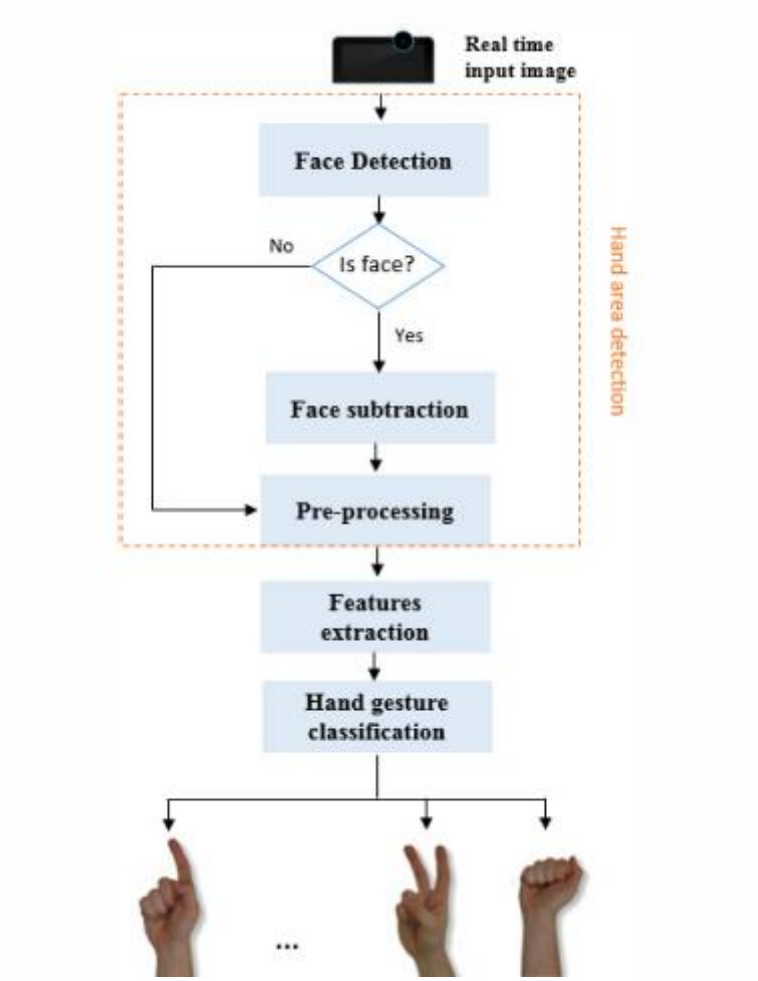

Figure 6.1: Proposed System Flowchart

6.1.1 Hand area extraction

This system extracts the hand region by applying skin color algorithm. As our algorithm is implemented on mobile devices, the computational complexity and the different uses cases and scenarios that a user can be situated. the proposed algorithm is based on applying a binarization based on HSV color space. This step is critical because it is important to localize the hand with the highest performance. Actually the human hand is

specified by it's specific color. The skin color detection is one of the most algorithm of segmentation used in literature mainly in the case of face and hand detection. The choice of the corresponding color space is very important in the performance of this method. The invariance with the illumination is a major factor of a successful detection. In this case, the use of RGB color is not a good choice because it is sensitive to llumination condition in contrast of another color space like HSV, YcbcCr, YUV and YIQ which separate between the illumination and chrominance components. The segmentation based on skin color is based generally on the use of threshold to detect skin areas.

6.1.2 Image preprocessing

Image processing is a method to convert an image into digital form and perform some operations on it, in order to get an enhanced image or to extract some useful information from it. Pre-processing is a common name for operations with images at the lowest level of abstraction — both input and output are intensity images. These iconic images are of the same kind as the original data captured by the sensor, with an intensity image usually represented by a matrix of image function values (brightness). The aim of pre-processing is an improvement of the image data that suppresses unwilling distortions or enhances some image features important for further processing, although geometric transformations of images (e.g. rotation, scaling, translation) are classified among pre-processing methods here since similar techniques are used.

6.1.3 Feature extraction

Once the face is detected, we aim to substrate it, from the image to avoid any conflict with the hand area localisation. We replace all the face pixel with black ones. This step eliminates the face in the binarization module and as a consequence provide the subtraction of the face. Feature extraction takes place through the different types of image detectors and sends this image to the detection module for processing in the form of frames. The commonly used methods of capturing input are data gloves, hand belts and cameras. In our system, we use the webcam inbuilt which is cost efficient to recognize both static and dynamic gestures.

6.1.4 Gesture recognition

The processed image is fed as input to the recognition algorithm to classify the gesture into respective classes. HAAR cascade classifier has been used as the classifying

algorithm. Haar-like features are digital image features used in object recognition. They owe their name to their intuitive similarity with Haar wavelets and were used in the first real-time face detector. A Haar-like feature considers adjacent rectangular regions at a specific location in a detection window, sums up the pixel intensities in each region and calculates the difference between these sums. This difference is then used to categorize subsections of an image. For example, let us say we have an image database with human faces. It is a common observation that among all faces the region of the eyes is darker than the region of the cheeks. Therefore a common Haar feature for face detection is a set of two adjacent rectangles that lie above the eye and the cheek region. The position of these rectangles is defined relative to a detection window that acts like a bounding box to the target object. In the detection phase, a window of the target size is moved over the input image, and for each subsection of the image the Haar-like feature is calculated. This difference is then compared to a learned threshold that separates non-objects from objects. The key advantage of a Haar-like feature over most other features is its calculation speed. Due to the use of integral images, a Haar-like feature of any size can be calculated in constant time. Use case diagram, activity diagram and interface design diagram are shown in Figure 6.2, Figure 6.3 and Figure 6.4 respectively.

6.2 USE CASE DIAGRAM

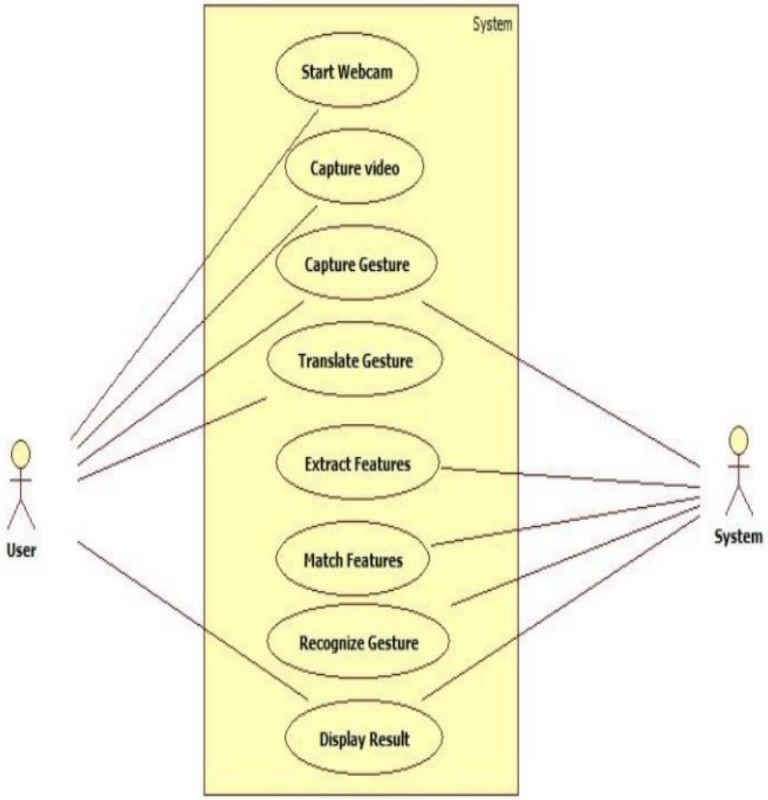

Figure 6.2: Use Case Diagram

6.3 Activity Diagram

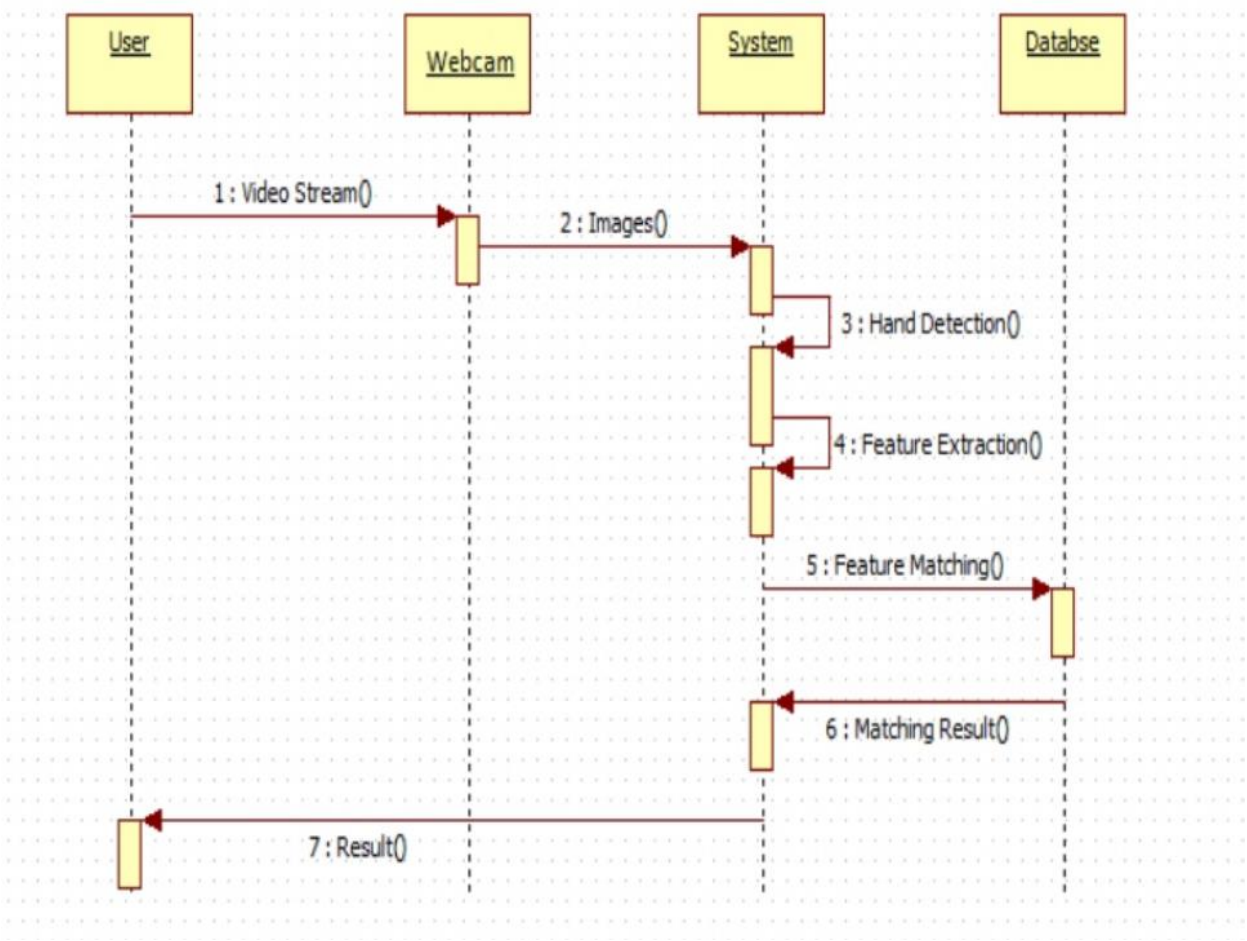

Figure 6.3: Activity Diagram

6.4 Interface Design

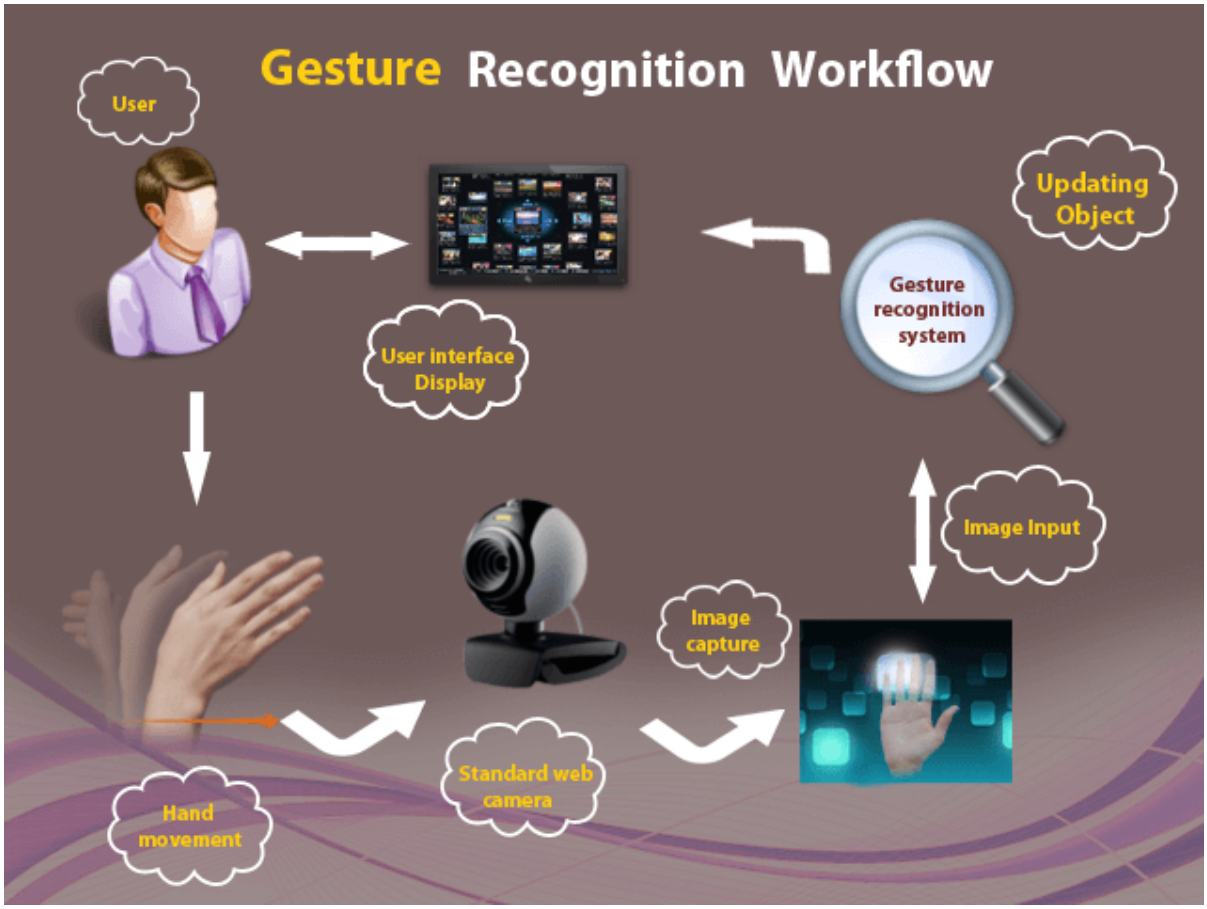

Figure 6.4: Interface Design

6.5 Modules

- MODULE 1 – Image acquisition: Image acquisition is the first step in any vision system, only after this process you can go forward with the image processing. In this application it is done by using IPWebCam android application. The application uses the camera present in the phone for continuous image capturing and a simultaneous display on the screen. The image captured by the application is streamed over its Wi-Fi connection (or WLAN without internet as used here) for remote viewing. The program access the image by logging to the devices IP, which is then showed in the GUI. Gestures of the hand are read by an input sensing device such as camera. It reads the movements of the human body and communicates with computer that uses these gestures as an input.

- MODULE 2 - Image processing: With the help of OpenCV library and feeding the image and hand features as inputs, image is transformed into simpler form and the useful part is extracted out of the image as shown in Figure 6.5.

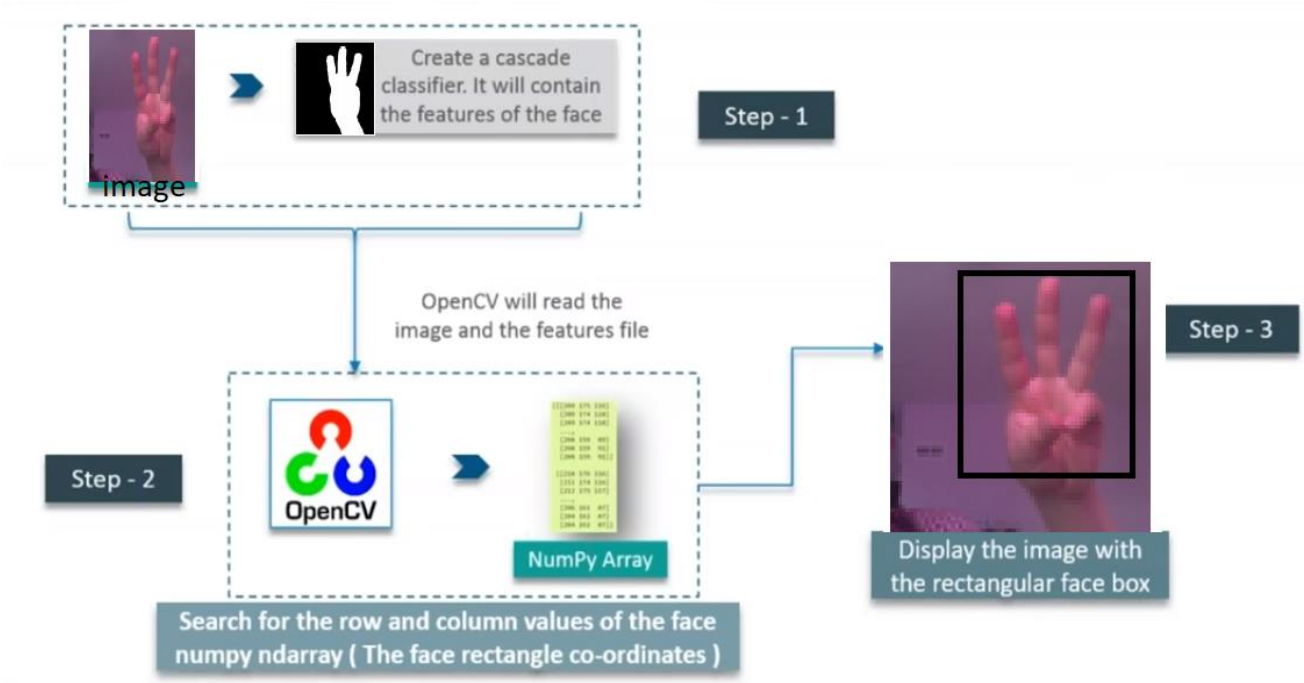

Figure 6.5 Functioning Of OpenCV

The process of generating the hand image, which is fed to the classifier, from the image captured by a camera. The process includes background subtraction, color filtering, Gaussian blurring, thresholding, morphological transformation, contour extraction, hand region extraction, image resizing.

- MODULE 3 - Gesture Recognition: The algorithm used for the classification purpose is HAAR cascade classifier. A Haar-like feature considers adjacent rectangular regions at a specific location in a detection window, sums up the pixel intensities in each region and calculates the difference between these sums. This difference is then used to categorize subsections of an image. For example, let us say we have an image database with human faces. It is a common observation that among all faces the region of the eyes is darker than the region of the cheeks. Therefore a common Haar feature for face detection is a set of two adjacent rectangles that lie above the eye and the cheek region. The position of these rectangles is defined relative to a detection window that acts like a bounding box to the target object. In the detection phase, a window of the target size is moved over the input image, and for each subsection of the image the Haar-like feature is calculated. This difference is then compared to a learned threshold that separates non-objects from objects.

- MODULE 4 – Action implementation: Once the gesture has been recognized, the corresponding action related to the gesture is ready to be performed. It is up to the user to decide the action to be performed corresponding to the gesture recognized. If the gesture recognized by the application has some corresponding action attached to it, the action is performed and the application has done its job perfectly otherwise an error is notified to the user.