Recognition of Hand Gestures Using Image Processing

Maria Anu V[1]., L.Mary Gladence[2], G.Nagarajan[3], J.Refonaa[4], Velgonda Laasya[5], Vinta Jahnavi[6]

[1,2,3,4] Associate Professor, Department of Computer Science and Engineering,
Sathyabama Institute of Science and Technology, Chennai, India
[5,6] UG Student, Department of Computer Science and Engineering,
Sathyabama Institute of Science and Technology, Chennai, India
maria.it@sathyabama.ac.in,
lvelgonda@gmail.com,jahnavireddy44@gmail.com,

Abstract. This paper presents a far better approach in recognizing the hand gestures supporting curvature of perimeter. Hand gestures are an integral part of communication, especially, if we are communicating any foreign language. Recognition of mitt is a characteristic and natural approach to connect with the PC, where collaborations with the computer can be expanded through various dimensional utilization of hand gestures as contrasted with elective info techniques. The aim of this project is to investigate various techniques for HGR (hand gesture recognition) using detection of the finger tips.This system algorithms makes use of image processing toolboxes of Mat lab, computer vision to access the webcam. The difficult and clumsy interface devices for human-computer interaction (HCI) can be solved by using webcam hand gestures.

Keywords: Gesture recognition, curvature of perimeter, human-computer interaction (HCI)

1   Introduction

With increasing technology, it is becoming easier to build applications for sign languages, one such application is being engineered using image processing. We will be able to help deliver a wide range of information across to the other person via hand gestures[1].
Based on Kendon and Quek et al. researches, the hand gestures can be categorized as the following

- Deictic – These are a type of hand gestures which involves a pointing activity to establish
the identity or spatial location of an object within the context of an application domain;
- Manipulative – These are usually performed by freehand movements to mimic
manipulations of physical objects, such as in virtual or augmented reality interfaces;
- Semaphoric – These are specific hand gestures that define a set of commands and/or
symbols to interact with machines. They are often used alternatively to the speech modality, when the latter is unusable or ineffective;
- Gesticulation – It is one of the most natural forms of gesturing and is commonly used in
combination with conversational speech interfaces. These hand gestures are often unpredictable and difficult to analyze;
- Language -It is the hand gestures used for sign language. They are performed by combining
a set of gestures to form grammatical structures for conversational style interfaces. In case of finger spelling, these gestures are often thought of like semaphoric ones.

Hand motion acknowledgement paves us an approach to comprehend the information communicated by the previously mentioned classifications, which are basically used to connect with applications that are innovative like intuitive games, gesture based communication recognizers, identifying emotional expression, robotics remote controllers, high level computer interfaces, and others[2].

By and large, the methodologies utilized close by hand gesture are regularly separated into two principle classes: Appearance based and Three-dimensional model-based[3,4]. The essential procedure utilizes important components of the body parts to secure important three-dimensional information, whereas the subsequent method utilizes pictures or video arrangements to obtain the data[5,6]. Back previously, numerous colored picture cameras were important to get a three-dimensional model of the body parts, just as hands[9]. Later experiments, upheld by cutting edge gadgets, e.g., LMC or Microsoft Kinect, in like manner as novel displaying calculations bolstered by profundity map idea, have empowered the use of 3D models among regular application areas[7,8].

Through this paper, we further justify the survey created to build the project in section II, existing system and proposed system in section III and IV respectively, followed by, block diagram in section V, inferred conclusion from the project in section VI.

2   Literature Survey

2.1 Intelligent Approaches to interact with Machines using Hand Gesture Recognition in Natural way: A Survey

This paper we will be dealing with the work done in the area of recognition of hand gestures. Here the main focus is on the soft computing primary based strategies like artificial neural network, fuzzy logics, genetic algorithms and intelligent approaches[16]. The strategies within the hand image construction and preprocessing for segmentation are also considered in the study. Many researchers used finger tips for hand gesture detection in appearance which are primarily based on modeling. Finally the various comparisons of results given by completely different researchers are additionally given[10].

The disadvantages with this paper are that in the paper we will work in the area of individual finger position bending detection and movements, as work done in this area are very few[11].

An Analysis of Features for Hand-Gesture Classification

The human-computer interaction, also called as HCI, depends totally on physical devices. The goal of this work is the evaluation and analysis of methods that permits the user to interact to machines employing a hand gesture based on natural language[14]. Here we present some approaches which are used in HCI systems employed on hand gesture and a replacement proposal that uses geometric shape descriptors for hand gesture classification[17]. The results analysis shows that this new proposal beats some limitations of different known HCI methods.

The disadvantages of this paper are that, in this paper the user has to wear special gloves that measure the hand pose and the joint angles[15]. The problem with this type of technique is that after the user wears a glove, the system becomes invasive, besides the very fact of special gloves being high cost.

A Vision-based Remote Control

This paper presents vision-based system touch free interaction with a display at a distance. A camera will be fastened on top of the systems screen and it will be pointing in the direction of the

user. A mechanism which alerts the user to start, permits the user to begin the interaction[18]. The user can manage the screen pointer by moving their mitt in a fist pose which is directed towards the camera[12].

The disadvantages of this page are in this method the initialization to search for the hand automatically is needed at the start of the interaction and after tracking frequent loss in tracking the first happens.
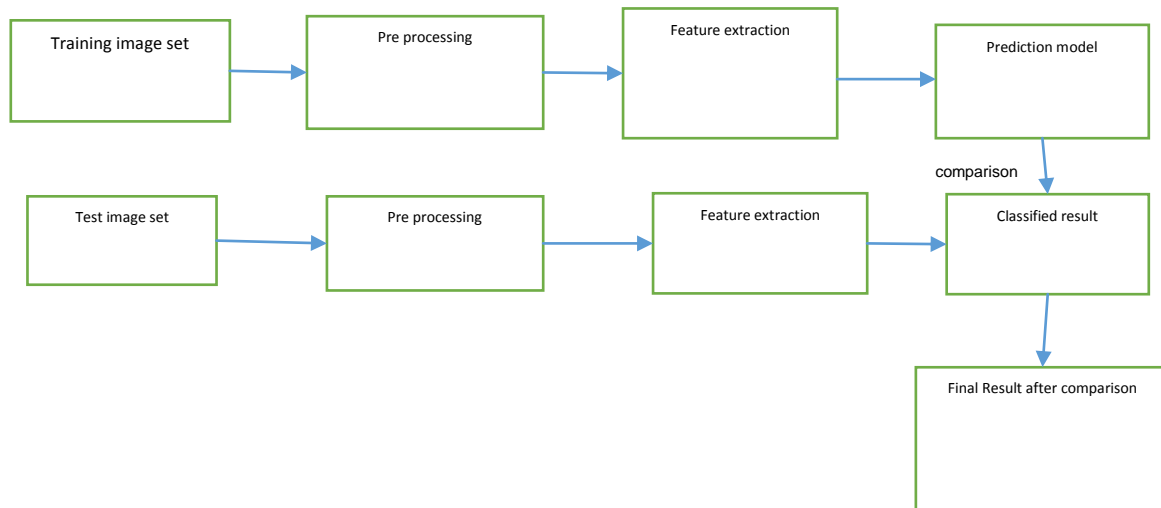
3   Existing System

Fingertip detection is used by many researches in recognition of hand gestures. During this existing system convex hull algorithms are used for recognition of these gestures[13,19]. The convex hull technique is quick however it's not that strong while considering a fist the knuckles are recognized as finger tips since they appear like the points in convex hull. Another thing is when moving the finger inwards, it losses the fingertips.

4   Proposed System

Web camera captures the hand actions without the need of gloves and saves the images in MATLAB. It transfers the commands via wireless and control the applications using wireless technology. During this project two algorithms are used as SVM and Curvature of Perimeter. For dynamic gesture recognition, using curvature of perimeter can prove to be robust. Curvature of Perimeter with its application can be bestowed as a virtual mouse.

5   Block Diagram

```
[Training image set] → [Pre processing] → [Feature extraction] → [Prediction model]
                                                                          |
                                                                   comparison
                                                                          ↓
[Test image set] → [Pre processing] → [Feature extraction] → [Classified result]
                                                                          |
                                                                          ↓
                                                         [Final Result after comparison]
```

**Fig. 1.** Basic understanding of how the image is interpreted and output is given

6   Strategy

Consider, each motion of the hand received from a participant, is represented by a $X = fx0; x1; ::::; xT1g$ of feature vectors, where T contains the maximum number of time instants, inside a time interval _, in which the features are separated by a LMC.

See that, a LMC is picked as reference gadget for the acquisitions since it is optimized for the hands and the acquired skeleton model gives exceptionally precise unique information about the bones of the finger. A Delayed long short-term memory technique is utilized and used to display these kind of groupings of information, where a period arrangement of one vector for each time instant is changed into a pattern of output probability vectors Y = fy0; y1; ::::; yT1g. Each yt 2 Y gives the probability of class of the motion completed at time t, with 0 _ t _ T 1. At last, the order of the motions is finished by a soft max layer [46] utilizing K = jCj classes, where C is the arrangement of the considered signal classes.

6.1 Feature extractions

Motion is a composition of different kinds of postures, where each posture is shown in a specific angle. Similar kind of concepts have been already used in a few works, utilizing the edges shaped by the body joints to perceive human activities. Along these lines, each element vector xt 2 X, with 0 _ t _ T 1

1. Inward edges ω1, ω2, ω3, and ω4 of the joints between moderate and distal phalanges. The interior edge ω0, considered for the thumb, is figured between distal and the proximal phalanx;

2. Inward edges β1, β2, β3, and β4 of the finger joints between moderate and proximal phalanges. The inward edge β0, taken for the thumb, is registered between metacarpal and proximal phalanx.
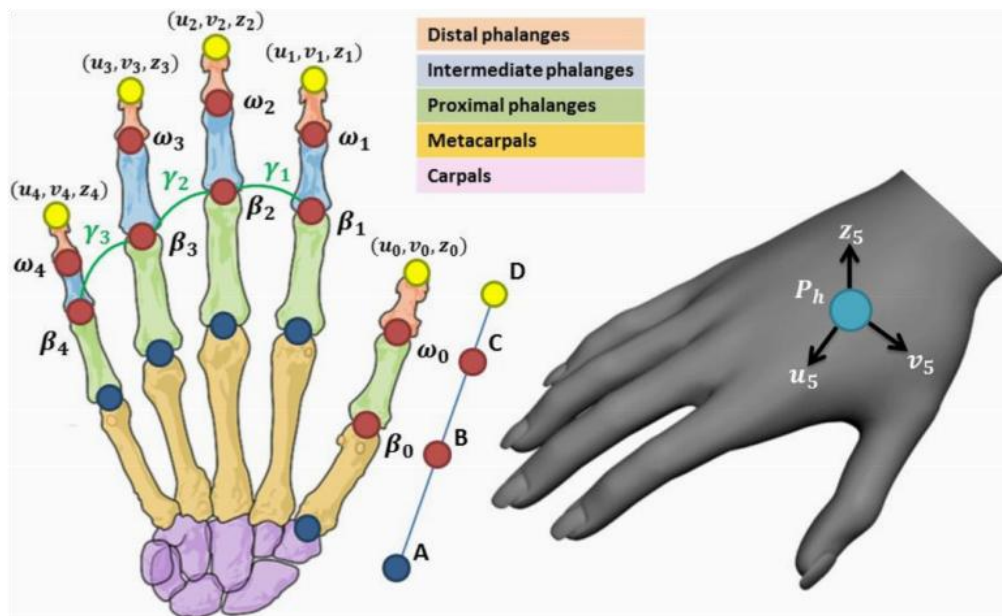
Fig. 2. Depiction of the features

Fig. 6.1. A depiction of the features extricated: joint points and position of finger tips. the edge yellow focuses show that the position of finger tips on which the three-dimensional removal is processed. the area with red focuses show us that the joints on which the edges are processed.

6.2 Sampling process

As every individual can play out a similar signal at various rates, and since the proposed technique

necessitates all the recordings that must analyzed are comprised of the similar number T of tests, an examining procedure known as Savitzky-Golay Filter.

6.3 DLSTM network

An important aspect in our planned system is that the network utilized in the hand motions arrangement, is employed on different long short-term memories, which differ from different kinds of Neural Networks (NNs). They can productively analyze time arrangements of information. Various components, similar to the mistake exploding issue and the disappearing slope, don't permit the utilization of basic initiation capacities (e.g., sigmoid) to appropriately prepare a system made by numerous recursive neural networks. This issue can be understood with the long short-term memory units.

The long short-term memory can be defined as a progression of memory obstructs which consists of three multiplicative units and at least one self-associated memory cells which are the input and the output, and forget gates. These entryways gives us consistent analogs of compose, read, and reset activities for the cells. Although a long short-term memory permits to handle the issue of the diminishing gradient, the information time arrangement regularly have a fleeting chain of command, with data that is spread out over numerous time scales that can't be appropriately perceived by basic repetitive systems, for example, LSTMs. Along these lines, profound LSTMs have been presented. Truth be told, by developing repeating systems framed over various layers, a higher deliberation on the info information is come to. Expanded information reflection doesn't continually bring benefits, on the grounds that the adequacy of these systems rely upon both assignment and dissected information.

In different works, for example, it was seen that profound long short-term memories work superior to the ones which are shallow on discourse acknowledgment. Broke down for example in discourse to-content assignment, can be expounded on more deliberations extending from the whole articulated expression to the syllables of each word. In addition, every deliberation can be caught in various time scales inside the allocated period. On account of sound successions broke down in the discourse acknowledgment issue, hand signals can be seen over numerous time scales. Each signal can be considered as created by numerous little developments and sub-motions of the hand and, as watched, this sort of information preparing is especially appropriate for this sort of system. In light of these contemplations, the LSTM stack-based arrangement was tested and afterward contrasted with the presentation of a solitary level system. The initial step was the meaning of the actuation elements of memory cell of the LSTM0 (the principal layer of the proposed neural system), just as the calculation of the info, yield, and overlook entryways controlled by assessing iteratively the accompanying conditions (from t = 0 to T 1):

$$i0;t = \_(Wxixt + Whih0;t1 + Wcic0;t1 + bi) \quad (4)$$
$$f0;t = \_(Wxfxt + Whfh0;t1 + Wcf\ c0;t1 + bf\ ) \quad (5)$$
$$c0;t = ft \oslash ct1 + i0;t \oslash tanh(Wxcxt + Whch0;t1 + bc) \quad (6)$$
$$o0;t = \_(Wxoxt + Whoh0;t1 + Wcoc0;t1 + bo) \quad (7)$$
$$h0;t = o0;t \oslash tanh(c0;t)$$

where, i- is the input gate, f- is the forget gate and o-is the output gate and c- is the cell activation vectors, respectively.

The concealed vector h and these vectors have a similar length. Rather, Wxi, Wxf ,Wxo, and Wxc are the loads of the information door, overlook entryway, yield door and cell to the info. What's more, Wic, Wfc, and Woc are the askew loads for peep-opening associations. At long last, the terms bi, bf ,bc, and bo demonstrate the info, overlook, cell and yield

inclination vectors, separately. Here _ denotes strategic sigmoid capacity and ⊙ denotes component astute result of the vectors. When the enactment capacities for the main level are characterized, then the the upper level initiation capacities are to be characterized.

7   Experimental Analysis

This segment depicts the exploratory tests directed to assess the exhibition of the proposed strategy. The DLSTM arrange and the BPTT calculation, were utilized to process the minimization utilized on the stochastic angle drop, were actualized by utilizing the Keras1 structure. The principle target of the trial meeting where the approval of the proposed technique, includes the joint points appraisal as notable highlights for the acknowledgment of hand motion, and providing a solution which proves to be best than the works of the present best in class. The principal objective had been acquired by making a difficult dataset dependent on the communication via gestures on which the ideal number of stacked Long short-term memory's and viability of the chose joints highlights were investigated. What's more, on the equivalent dataset, a lot of notable measurements was figured to assess the general execution of the methodology. Rather, the subsequent objective was gotten by contrasting the proposed technique and other extensive chips away at the premise of the SHREC dataset. Below are the screenshots of the following experimental analysis:
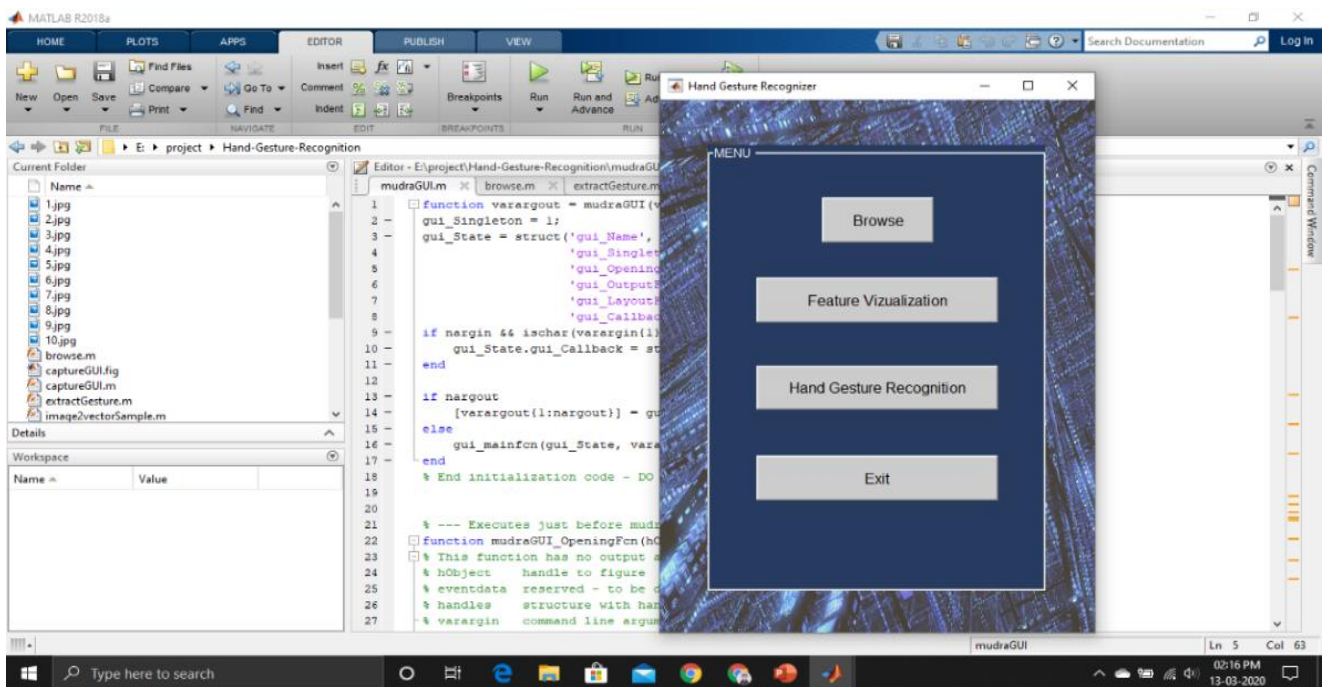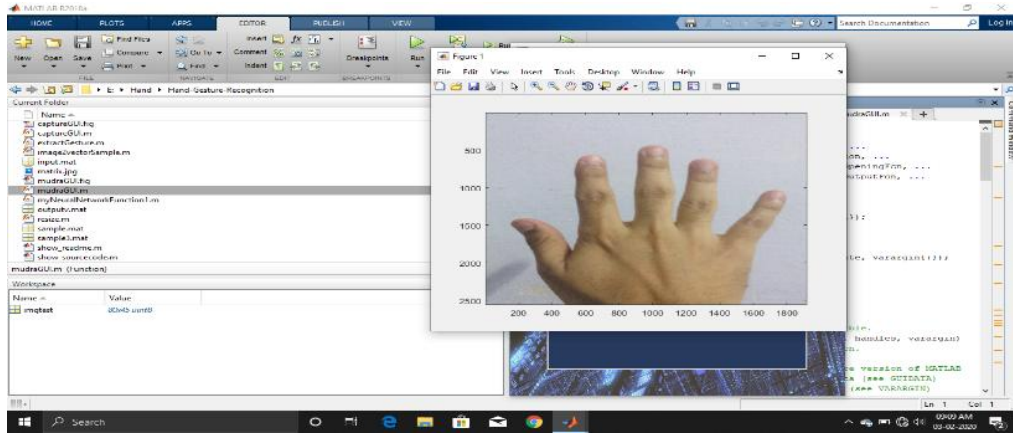

Fig. 3. Browsing Dataset
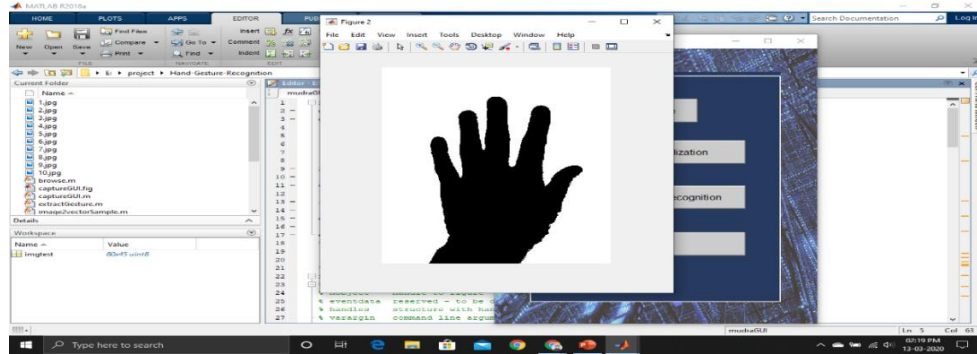
Fig. 4. Selected image is shown

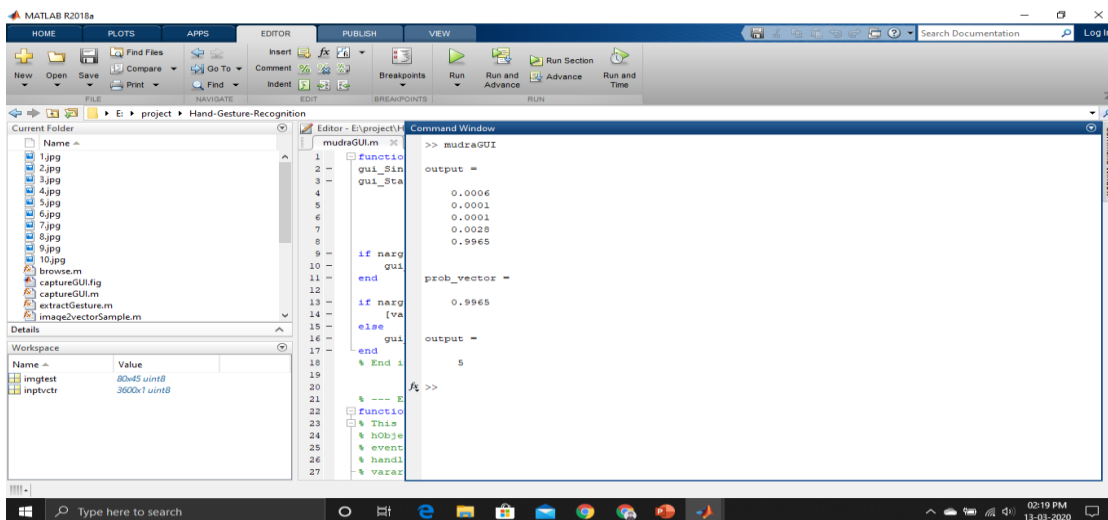**Fig. 5.** Feature Visualization

Fig. 6. Hand gesture recognition

8   Conclusion

Right now, unique hand signal acknowledgment strategy dependent on dLSTM is exhibited. Specifically, a full of feeling set of distinguishing highlights dependent on both   fingertip positioning and joint edges is utilized in mix with a Long short-term memory Recurrent neural networks to acquire high exactness results. The technique we proposed beats contending takes a shot at the SHREC dataset. This dataset has been additionally used to break down the power of the separated highlights and the conduct of the system when the quantity of Long short-term memory's change.

At a subsequent stage, we plan on making another open dataset, constantly depending on the American Sign Language, where many more hand signals are embedded at regular intervals. RGB outlines will be included in this new dataset, profundity maps, and the entire hand skeleton model. This dataset ought to have the option to help various ambiguities study cases (e.g., the acknowledgment of hand motions). As of now, as future improvement of the proposed strategy, we are attempting to blend highlights removed from the hand.

References
1. Ankit Chaudhary, J. L. Raheja , Karen Das , Sonia Raheja,"Intelligent Approaches to interact with Machinesusing Hand Gesture Recognition in Natural way". International Journal of Computer Science & Engineering Survey (IJCSES) Vol.2, No.1, Feb 2011.
2. K. G. Derpanis, "A Review of Vision-Based Hand Gesture,"Department of Computer Science, York University, February, 2004.
3. Thiago R. T rigo, Sergio Roberto M. Pellegrino, "An Analysis of Features for Hand-Gesture". IWSSIP 2010 17th International Conference on Systems, Signals and Image Processing.
4. Jong-Min Kim; Woong-Ki Lee, "Hand Shape Recognition Using Fingertips" Fuzzy Systems and Knowledge Discovery, 2008. Fifth International Conference on, vol.4, no., pp.4448, 18-20 Oct. 2008.
5. Oka, K.; Sato, Y.; Koike, H., "Real-time fingertip tracking and gesture recognition," Computer Graphics and Applications, IEEE, vol.22, no.6, pp. 64-71, Nov/Dec 2002.
6. Nolker C., Ritter H., Visual Recognition of Continuous Hand Postures, IEEE Transactions on neural
   networks Vol 13, No.4, July 2002, pp. 983-994.
7. Nguyen D.D., Pham T.C., Jeon J.W., Fingertip Detection with Morphology and Geometric Calculation,
   IEEE/RSJ International Conference on Intelligent Robots and Systems, St. Louis ,USA, Oct 11-15, 2009,
   pp. 1460-1465.
8. Lee D. and Park Y., Vision-Based Remote Control System by Motion Detection and Open Finger Counting, IEEE Transactions on Consumer Electronics, Vol. 55, issue 4, Nov 2009.