

# Hand Gesture Recognition Algorithm for Smart Cities based on Wireless Sensor

<https://doi.org/10.3991/ijoe.v13i06.7022>

Thittaporn Ganokratanaa  
Chulalongkorn University, Pathum Wan, Bangkok  
acharisma\_sbunny@hotmail.com

Suree Pumrin\*  
Chulalongkorn University, Pathum Wan, Bangkok  
suree.p@chula.ac.th

**Abstract**—The relationship between humans and computers is called human computer interaction (HCI). HCI is a major research topic in the area of assistive technology. In the field of medical assistive technology specifically, a hand gesture is considered a suitable method to convey information. It can help elderly people, who are unable to walk or speak, communicate with caregivers whenever they need help. This paper proposes a system to recognize the hand gestures of elderly people using an inexpensive Raspberry Pi. A vision-based algorithm is developed to detect and classify dynamic hand gestures in real time on the Raspberry Pi embedded platform. There are three main procedures; contour detection, convex extraction, and rule-based classification. The system can detect six different gestures on both hands in various orientations. The experiment showed good results in detecting and classifying their meanings as lingual descriptions.

**Keywords**—Hand gesture recognition, assistive technology, Rule-based, Vision-based

## 1 Introduction

As the availability of low cost sensors and processors increase, computer vision systems are now being used as the primary interface in various assistive system applications. Human-computer interaction (HCI) techniques that utilize computer vision play an important role in helping medical systems analyze data collected from monitoring systems that track the patient's physical activities and emotions [1]. This type of interface is beneficial to the sick, physically handicapped, and elderly. However, even though HCI is used widely, it is not often calibrated for a specific type of patient, especially the elderly people. Hand gestures are especially useful for elderly people who stay alone and cannot be able to walk or express their feelings with words.

According to the United States census bureau [2], elderly people will reach the highest rate of population growth in 2043. One key concern is that many of the elder-

ly need to stay in a hospital or at home under full-time care and that there are many difficulties in taking care of them. For example, it is tough for the caregivers to keep careful watch in the hospital because they have many patients to take care of. As a result, some elderly patients are not able to contact anyone when they need help or when they want to eat, drink or use the restroom [3].

As mentioned above, a monitoring system is very useful for both the patient and the caregivers as it helps monitor the activities of patient and can generate alarms for the caregiver. One common design of these monitoring systems is to use an embedded platform that includes both a camera and a processor. A popular embedded system platform is the Raspberry Pi. The use of Raspberry Pi board and Raspberry Pi's camera are included for processing and monitoring real-time scenes as Raspberry Pi is a low cost device, easy to implement and has its computing processor. It can be integrated with vision-based techniques that provide effective results without interaction with additional devices. The key of a vision-based interface is to find the algorithm that suits for the hand gesture recognition in each environment and situation. The proposed vision-based system has many advantages such as silent, low cost, safety, and flexible use of algorithms.

For this application, the goal is to interact with the computer in a natural manner. Hand gesture is classified as an intuitive way to communicate with the computer because people can easily and instinctively control their hand movements for conveying information and conveying their thoughts [4]. There is a topic of research called hand gesture recognition that focuses on the development of algorithm-based machine learning applications. Its performances are able to detect, monitor, and classify or recognize the hand gestures. Also, it can interpret the meanings of the hand gestures as a lingual description.

For the hand gestures used in the system, simple six hand gestures are provided in order to help the elderly in remembering and demonstrating the correct gestures - complicated hand gestures with complex meanings are avoided. The main idea is that the meanings of the hand gestures should depend on the agreement between the elderly patient and the caregiver. Since the elderly patients might have different cultures and different types of knowledge, the patients can associate the hand gesture with the meaning by themselves. Thus, they will remember and understand each gesture correctly. Additionally, according to [5], some elderly patients might not be able to keep their hands steady because of their physical characteristics such as numbness or tingling and shaking hands. Thus, to reduce the risk of health hazard, the dynamic hand gesture is detected instead of the static hand gestures.

Given the importance of hand gesture recognition for elderly people, this paper proposes a system for dynamic hand gesture recognition that is integrated with a monitoring system using a Raspberry Pi that is suitable for elderly people. An algorithm for hand gesture recognition of elderly people is proposed in this paper. The main procedures are contour detection, convex extraction, and rule-based classification. The experimental results show good vision-based hand gesture recognition and good lingual descriptions for all six hand gestures experiments. This paper is provided into five sections; introduction, related works, proposed method, experimental results, and conclusion, respectively.

## **2 Related Works**

There are a variety of techniques for hand gesture recognition used by different researchers. Some may use additional devices considered as contact-based devices such as gloves, trackers for the hand, and orientation sensors for the wrist. Others are using the computer-vision-based interfaces that require the development of algorithms. In previous works, contact based devices are more likely to be used because they are easy to implement and provide accurate recognition results. However, they encumber the users and make an interaction more cumbersome. In the case of an elderly user, having to wear additional devices on their hand, wrist, arm, or body may cause health issues [6]. Hence, a vision-based algorithm is the focus of this paper and is considered to be more suitable for hand gesture recognition system than a contact based device. For a dynamic hand gesture, there are several conditions to consider such as the movement of the hand posture, location, and orientation at time. Some interesting results of dynamic hand gesture recognition are provided as below.

Shan, C. et al. [7] proposed the hand tracking system by the use of the mean shift embedded with the particle filter (MSEPF) implemented in real-time. They provided the advantages of two integrated approaches, the MS optimization and particle filter for improving the hand tracking for control interface for a robotic wheelchair in dynamic environments. MSEPF enhanced the sampling efficiency. The results of the hand tracking showed that the MSEPF was better than the PF and MS tracker. Five dynamic gestures were provided, including wave, move horizontally, move vertically, move clockwise, and move counterclockwise. The experiment implemented the sequences of the hand gestures more than 30 files. Each sequence had 400 frames at rate 12 fps. The MSEPF which required 50 particles gave better optimized tracking performance than the PF which required the MS with 200 particles. The percentage of the PF, MSEPF, and MS tracking were 100, 100, and 73 percent, respectively. The tracking performance was 85 percent.

Hsieh, C. C. et al. [8] proposed a robust system consisting of three modules: digital zoom, adaptive skin detection, and hand gesture recognition. The first module detected user's face and zoomed in the considering face region within the ROI included the hand gesture, thus the face and upper part were at the center part of the image. The second module analyzed the user's face color to define the other regions of skin color such as hands by using adaptive skin color. The last module was the most important part for processing both static and dynamic hand gesture recognition systems. Haar features with the use of SVM were used to enhance the accuracy results for recognizing the static hand gesture. In addition, for dynamic hand gestures, after digital zoom phase, it continued to recognition phase by using motion history image (MHI) combined with Haar features. The four dynamic hand gestures, including moving right, left, up, and down, and two static hand gestures including fist and waving hand gestures were applied in the experiment. The accuracy of the algorithm's results divided into two parts. For static hand gesture, after applying the face detection based skin color, its accuracy for recognition was 95.37 percent with the processing speed at 3.93 milliseconds per frame. For dynamic hand gesture, they used SVM to improve the

accuracy in each gesture. The overall accuracy of dynamic hand gesture was 95.66 percent.

Suk, H.I. et al. [9] proposed the dynamic Bayesian network DBN-based hand gesture model and the design of a gesture network model to enhance the media control and slide presentation. It focused on the extraction, modeling, motion tracking, and recognition. Both were believed to have a strong potential for developing applications related to sign language recognition. Many techniques were used such as discrete techniques for achieving the features, skin color model for detecting skin color pixels in YIQ color model, and blob for tracking the hand motion with optical flow for more accuracy. DBN were used with a cross-validation for the hand gesture recognition. However, it was quite complicated processes because it required the analysis of hand shapes. The accuracy of the recognition result by using DBN for the isolated hand gesture was 99.59 percent. While, the recall of continuous hand gesture was 84 percent and the accuracy was 80.77 percent.

According to the review, we found that most of vision based techniques in hand gesture recognition are suitable for either the hand is dynamic or statistic gestures and one or two-handed [10]. Some techniques are not flexible with the conditions. Also, it is still a challenge for real time dynamic two-handed gestures with various illuminations and for using with the embedded system in real world.

### **3 System Design**

To design the system, the limitation of the numbers of the gestures is needed to the users. This system is implemented with six hand gestures and designed by the developed algorithm with Raspberry Pi 2 Model B and its camera module. The hardware devices of the system are shown in Fig. 1.

#### **3.1. Camera Module**

Camera module is connected to Raspberry Pi 2 Model B board in order to monitor the hand gestures from the elder people in real-time and send the captured image to implement in the board. The output image size is equal to a camera resolution. This system sets the resolution at 800x600 pixels.

#### **3.2. Raspberry Pi 2 Model B**

This board receives the image from the camera module and then processes all image processing operations. Its power enables to charge with 5-volt power supply.

### **4 Proposed Method**

The proposed method of hand gesture recognition is divided into seven procedures as shown in Fig. 2.

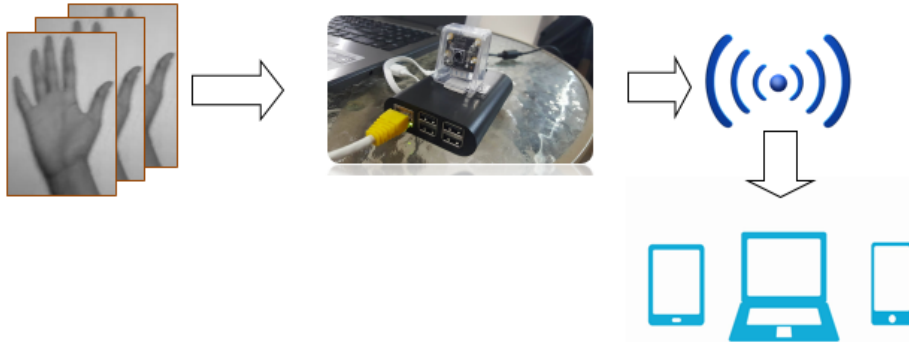


Fig. 1. Hardware devices of the system.

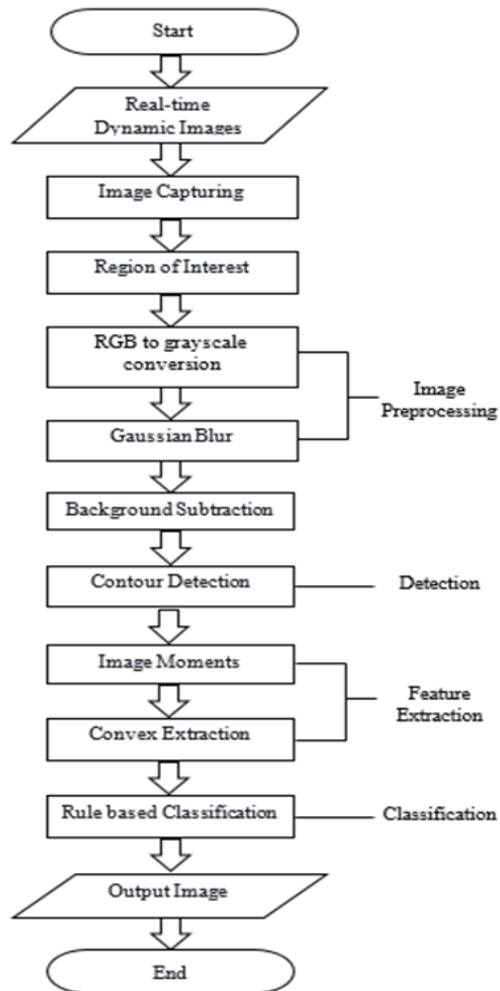


Fig. 2. Flowchart of the proposed method of the system.

#### 4.1. Image Requiring

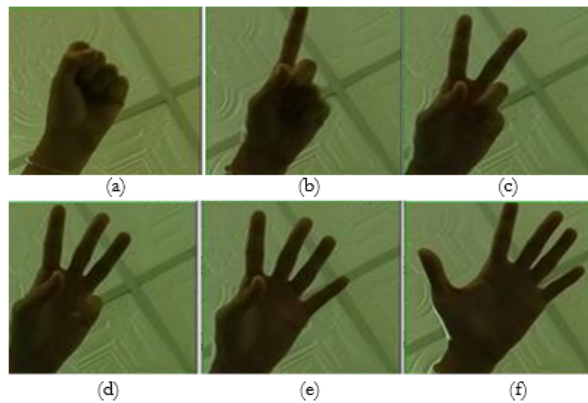
First procedure is to require the real-time image frames from a camera by splitting the image frames into single frame. The frame will be the same size as the camera resolution.

#### 4.2. Image Capturing

The system stores all sub-frames in the stacked array and divides to frames. After that, it collects all frames to display in the storage module in the frame buffer for analysis in the next process.

#### 4.3. Region of Interest

This procedure can be done by defining the rectangle bounding box to specify only the hand region as shown in Fig. 3.



**Fig. 3.** Results of region of interest; (a) the closed fist, (b) the pointing finger, (c) the two pointing fingers, (d) the three pointing fingers, (e) the four pointing fingers, and (f) the open palm.

#### 4.4. RGB to Grayscale Conversion

The RGB hand region image is converted into grayscale image as shown in Fig. 4. Grayscale image is easier to implement than RGB image. It only accesses in the gray levels that their channels are less than the RGB channels.



**Fig. 4.** Results of RGB to grayscale conversion.

#### 4.5. Background Subtraction

Background subtraction is applied with the static camera. It extracts the foreground object from the background and shows in binary image by calculating the difference between the current and reference frame and then removing those background pixels. Its method begins with the conversion of RGB to gray scale image. Gaussian filter is also used to remove noise signal. Then, threshold is defined to convert gray scale to binary image. For the color of pixels in binary image, white represents hand and black represents background as shown in Fig. 5. Its equation is

$$o_{ij} = c_{ij} - r_{ij} \quad 1)$$

where  $o_{ij}$  is the output result between the current and reference frames,  $c_{ij}$  is the current frame, and  $r_{ij}$  is the reference frame.

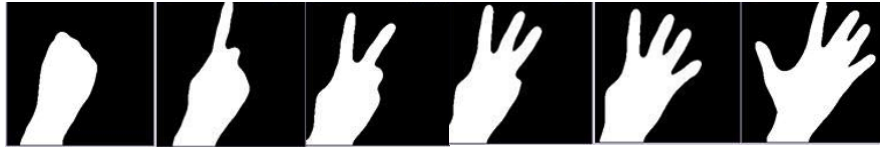


Fig. 5. Results of background subtraction.

#### 4.6. Image Moments

Moment features implemented without considering the image location and size of the hand [11]. Image moments are shown in Fig.6 and can be calculated by the following equations

$$M_i = \sum_{ij} x_i M(i, j) \quad (2)$$

$$M_j = \sum_{ij} y_j M(i, j) \quad (3)$$

The function  $M(i, j)$  defines the hand in the image and generates the moments  $M_{ij}$  to provide features of the hand. Image moments consist of center of mass, variance, and orientation as follow.

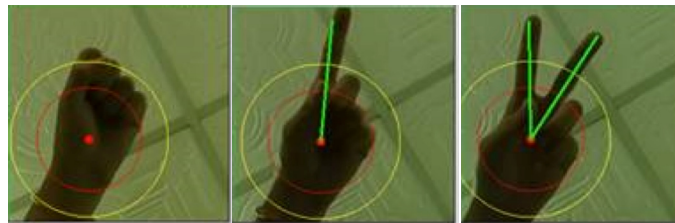


Fig. 6. Results of image moments.

**Center of Mass:** Center of mass is used to locate the center of palm called centroid and can be found by calculating the distance of the hand both of x- and y-axis. Then, using circle shape to determine the radius of the center of palm. This circle shape will be defined to a square frame that includes all pixels of the hand gestures.

$$M_{10} = \sum \sum x_i M(i, j) / M_{00} \quad (4)$$

$$M_{01} = \sum \sum y_j M(i, j) / M_{00} \quad (5)$$

For the calculation,  $M(i, j)$  is defined as the intensity at each point  $(i, j)$  of the given image ( $M$ ).  $M$  is represented the mass of  $(i, j)$ .

**Variance:** Variance ( $\sigma_2$ ) is obtained from the second moment of the centroid ( $M_{ij}^a$ ). It is calculated as

$$\sigma_x^2 = M_{20}^a = M_{20} - M_{10}^2 \quad (6)$$

$$\sigma_y^2 = M_{02}^a = M_{02} - M_{01}^2 \quad (7)$$

**Orientation:** Orientation represents the angle of axis of the least moment of the inertia. The calculation is shown in equation (8), (9), and (10)

$$\tan(2\theta) = \frac{2M_{11}}{M_{20} - M_{02}} \quad (8)$$

unless  $M_{11} = 0$  and  $M_{20} = M_{02}$ , consequently,

$$\sin(2\theta) = \frac{\pm 2M_{11}}{\sqrt{(2M_{11})^2 + (M_{20} - M_{02})^2}} \quad (9)$$

$$\cos(2\theta) = \frac{\pm (M_{20} - M_{02})}{\sqrt{(2M_{11})^2 + (M_{20} - M_{02})^2}} \quad (10)$$

#### 4.7. Contour Detection

After locating image moments, contour detection is the next procedure used to find the curve of the hand. The hand curve can be presented as many lines [12]. The green outline represents the largest contour of the hand which is shown in Fig. 7.

In modeling, the contour is defined as a parametric curve in the  $(x, y)$  plane of the image as shown in the equation (11)

$$\mathbf{v}(s) = (\mathbf{x}(s), \mathbf{y}(s)) \quad (11)$$

Contour detection provides energy ( $e_{snake}$ ) which is defined as the sum of the three energy terms including,  $e_{int}$ ,  $e_{ext}$ , and  $e_{cons}$ , representing internal, external and constraint, respectively. The energy is calculated as

$$e_{snake} = e_{int} + e_{ext} + e_{cons} \quad (12)$$

The internal energy ( $e_{int}$ ) is derived from the sum of the elastic energy ( $e_{elas}$ ) and bending energy ( $e_{bend}$ ) which are described below:



Elastic energy represents an elastic potential energy that decreases stretching and suits for shrinking the contour. The elastic energy is provided in the equation (13)

$$e_{elas} = \frac{1}{2} \int_S \alpha(s) |v_s|^2 ds \quad (13)$$

where weight is presented as  $\alpha(s)$  for controlling the elastic energy with different parts of contour.

Bending Energy ( $e_{bend}$ ) applied for a thin metal strip. It is calculated by the sum of squared curve of the contour. The equation of the bending energy is shown as

$$e_{bend} = \frac{1}{2} \int_S \beta(s) |v_{ss}|^2 ds \quad (14)$$

where  $b(s)$  is served as similar role to  $\alpha(s)$ . Thus, the internal energy can be calculated as

$$e_{int} = e_{elas} + e_{bend} = \int_S \frac{1}{2} (\alpha |v_s|^2 + \beta |v_{ss}|^2) ds \quad (15)$$

For finding an external energy, the external energy implements on small values including boundaries. It can be derived from the image defined by a function  $e_{img}(x,y)$ ,

$$e_{ext} = \int_S e_{img}(v(s)) ds \quad (16)$$

Thus, the energy ( $e_{snake}$ ) equation which finds a contour  $v(s)$  can be calculated to minimize the energy function of the hand by Lagrange equation as shown in following equation:

$$e_{snake} = \int_S \frac{1}{2} (\alpha(s) |v_s|^2 + \beta(s) |v_{ss}|^2) + e_{img}(v(s)) ds \quad (17)$$

$$\alpha v_{ss} - \beta v_{ssss} - \nabla e_{img} = 0 \quad (18)$$

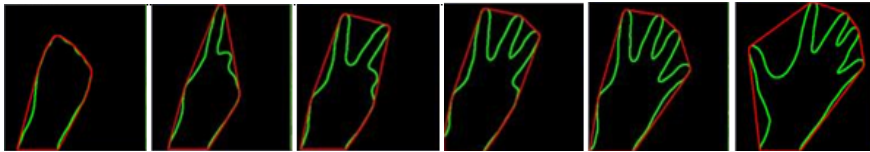


Fig. 7. Results of contour detection.

#### 4.8. Convex Extraction

Convex extraction is utilized as the feature extraction which includes convex hull and convexity defects. Fig. 8 shows the results after extracting the convex of the hand. The feature extraction is described as follow [13]:

**Convex hull:** It is the largest contour provided without the curve of the hand as shown in the red line in Fig.8. The convex hull can be divided into three parts, including lines, segments, and polygons as explained below:

$$ax + by + c = 0 \quad (19)$$

Lines can be represented as 1 performs a triple (a,b,c), in which a, b, and c are the coefficients of the linear equation

$$\frac{x-x_1}{x_2-x_1} = \frac{y-y_1}{y_2-y_1} \quad (20)$$

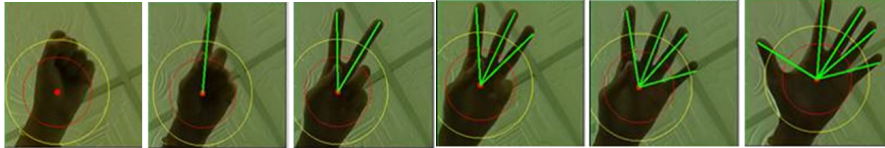
then we derive:

$$a = (y_2 - y_1) \quad (21)$$

$$b = -(x_2 - x_1) \quad (22)$$

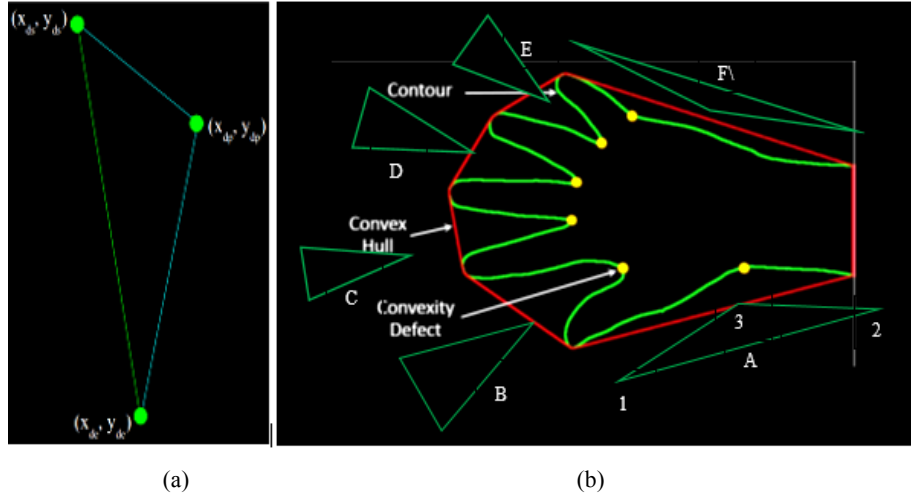
$$c = y_1(x_2 - x_1) - x_1(y_2 - y_1) \quad (23)$$

where segment s is represented by giving the pair (p,q) of points in the (x,y) plane that forms the endpoints of s and gives the line through the points with a range of xy coordinates that restricted to s. And polygon P is represented by performing a circular sequence of points called the vertices of P. The edges of P are the segments between P consecutive vertices. Polygon is considered as convex when it is simple shape, nonintersecting line, and all its angles are less than  $\pi$ .



**Fig. 8.** Convex extraction procedure

**Convexity Defect:** After drawing the convex hull around the contour line of the hand, the contour points are defined within the hull by using minimum points. The formation of defects is provided in the convex hull due to the contour of the hand. A defect is at the hand contour away from the convex hull. The set of values for every defect in the convexity is formed as vector. This vector contains the start, end, and defect points of the line in the convex hull. These points indicate the coordinate points of the contour [14]. Fig. 9 shows the convexity defects represented in yellow point.



**Fig. 9.** (a) The illustration of contour, convex hull, and convexity defect with the six defect triangles; A,B,C,D,E,F, and (b) the three coordinates; the defect start (xds, yds), defect end (xde, yde), and defect position points (xdp, ydp), labeled as 1,2, and 3, respectively.

In classification, the feature vectors are derived from the hull defects. The convexity defect in the hull is the distance between the contour line and actual hand. When the camera monitors the hand, the three coordinate points of defect, including the start (xds, yds), end (xde, yde), and position points (xdp, ydp), labeled as 1, 2, and 3, respectively, can be described by six defect triangles (A, B, C, D, E, F) in each frame as shown in Fig. 9. Thus, defect triangle A can be described as the following vector, called  $V_{td}^A$ , including  $[A(xds) A(yds) A(xdp) A(ydp) A(xde) A(yde)]$ .

From the convex defects, five largest defect areas of the hand is defined based on the defect position points (xdp, ydp) from the top six triangles: A, B, C, D, E, F. Accurate registration is important between the frames for further analysis.

#### 4.9. Euclidean Distance

The distance between two pixels can be found by Euclidean distance as shown in Fig.10. In x and y coordinates that restricted to  $s$ , we assume that the distance  $d$  can be calculated by defining the value of pixel one at the  $(i_1, j_1)$  and pixel two at the  $(i_2, j_2)$  as shown in equation (24)

$$d = \sqrt{(i_2 - i_1)^2 + (j_2 - j_1)^2} \quad (24)$$

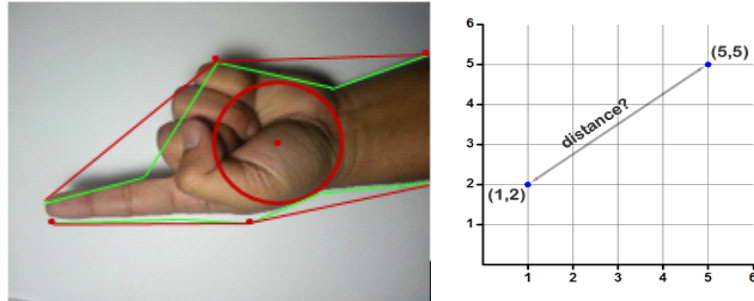


Fig. 10. The illustration of Euclidean distance.

#### 4.10. Rule-Based Classification

Rule based classification provides a set of encoded rules extracted from input gestures and compared with feature inputs. The output gesture is shown by matching the input gestures and the rules [15]. The rules consider the distance between the centroid of the hand and any point of fingertips. The threshold is set at 200 pixels. If the distance is less than 200 pixels, it will classify the gesture as stop. Conversely, if the distance is more than 200 pixels, it will classify the gesture according to the number of lines that occurred. The rules are divided into six procedures as shown in Fig.11.

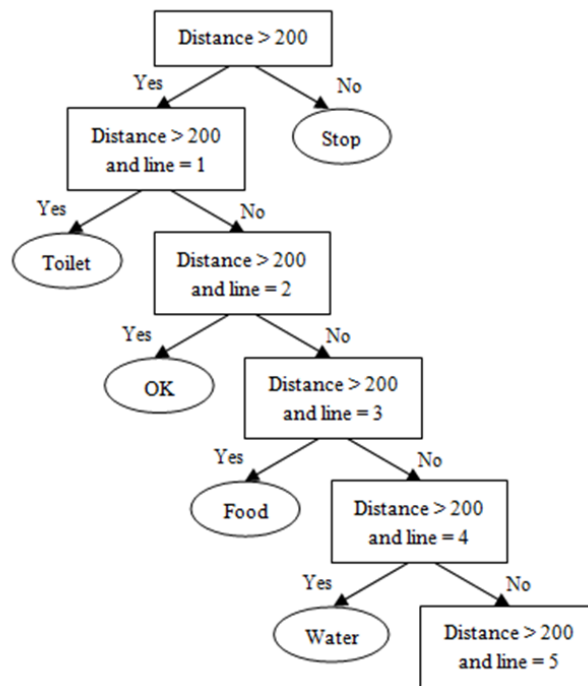


Fig. 11. The decision rules of rule based classification.

Rule based classification provides a set of encoded rules extracted from input gestures and compared with feature inputs. The output gesture is shown by matching the input gestures and the rules [14]. The rules consider the distance between the centroid of the hand and any point of fingertips. The threshold is set at 200 pixels. If the distance is less than 200 pixels, it will classify the gesture as stop. Conversely, if the distance is more than 200 pixels, it will classify the gesture according to the number of lines that occurred. The rules are divided into six procedures as shown in Fig.11 and can be described below:

Rule 1: The system interprets the lingual description of the close fist image that has no straight line longer than 200 pixels from the centroid of the hand drawn to any point of fingertips as “stop”.

Rule 2: The system interprets the lingual description of the pointing finger image that has a straight line longer than 200 pixels drawn from the centroid of the hand towards the pointing fingertip as the requirement of “toilet”.

Rule 3: The system interprets the lingual description of the image of pointing and middle fingers that has two straight lines longer than 200 pixels drawn from the centroid of the hand towards the pointing and middle fingertips as “ok”.

Rule 4: The system interprets the lingual description of the image of pointing, middle, and forth fingers that has three straight lines longer than 200 pixels drawn from the centroid of the hand towards the pointing, middle, and fourth fingertips as the requirement of “food”.

Rule 5: The system interprets the lingual description of the image of pointing, middle, forth, and little fingers that has four straight lines longer than 200 pixels drawn from the centroid of the hand towards the pointing, middle, fourth, and little fingertips as the requirement of “water”.

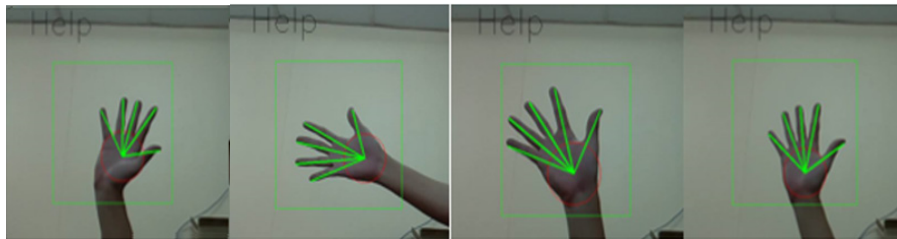
Rule 6: The system interprets the lingual description of the open palm image that has five straight lines longer than 200 pixels drawn from the centroid of the hand towards the five fingertips as the requirement of “help”.

## **5 Experimental Results**

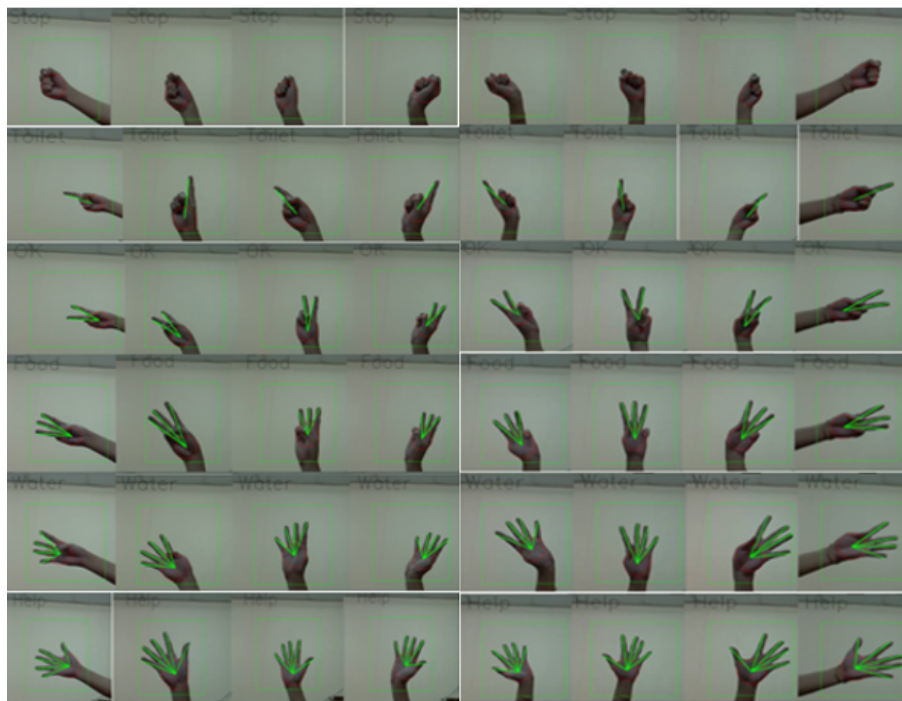
The hand gesture recognition system is implemented by analyzing the images of the six different static hand gestures under the closed environment. The camera is set from the top level above the hand about 30 centimeters. A static white paper placed at the reference level to be used as a scene in the background. The light source comes from fluorescent bulb that was about two meters above the reference level. The results show the good performance of classification as follows:

From the experimental results, the hand gesture recognition system provides good results in hand detection and classification for all the six hand gestures in various degrees as shown in Fig. 12 and 13. From Fig. 14, it can be seen that the centroid of palm, contour line of the hand, and the points of the fingertips are detected precisely. Moreover, the results of the hand gestures indicate the meaning very well, including the close fist represents “stop”, the pointing finger represents “toilet”, the pointing and middle fingers represent “ok”, the pointing, middle, and forth fingers represent

“food”, the pointing, middle, forth, and little fingers represent “water”, and the open palm represents “help”. These meanings have sent to the line application over wireless sensor to alert the caregiver as shown in Fig. 15.



**Fig. 12.**The results of the open right palm gesture of the real-time dynamic hand gesture recognition system for elderly people in various degrees; (a) 65 degrees (b) 45 degrees (c) 0 degrees and (d) -45 degrees.



**Fig. 13.**The examples of the results of the hand gesture recognition system from the right hand to the left hand in various degrees.

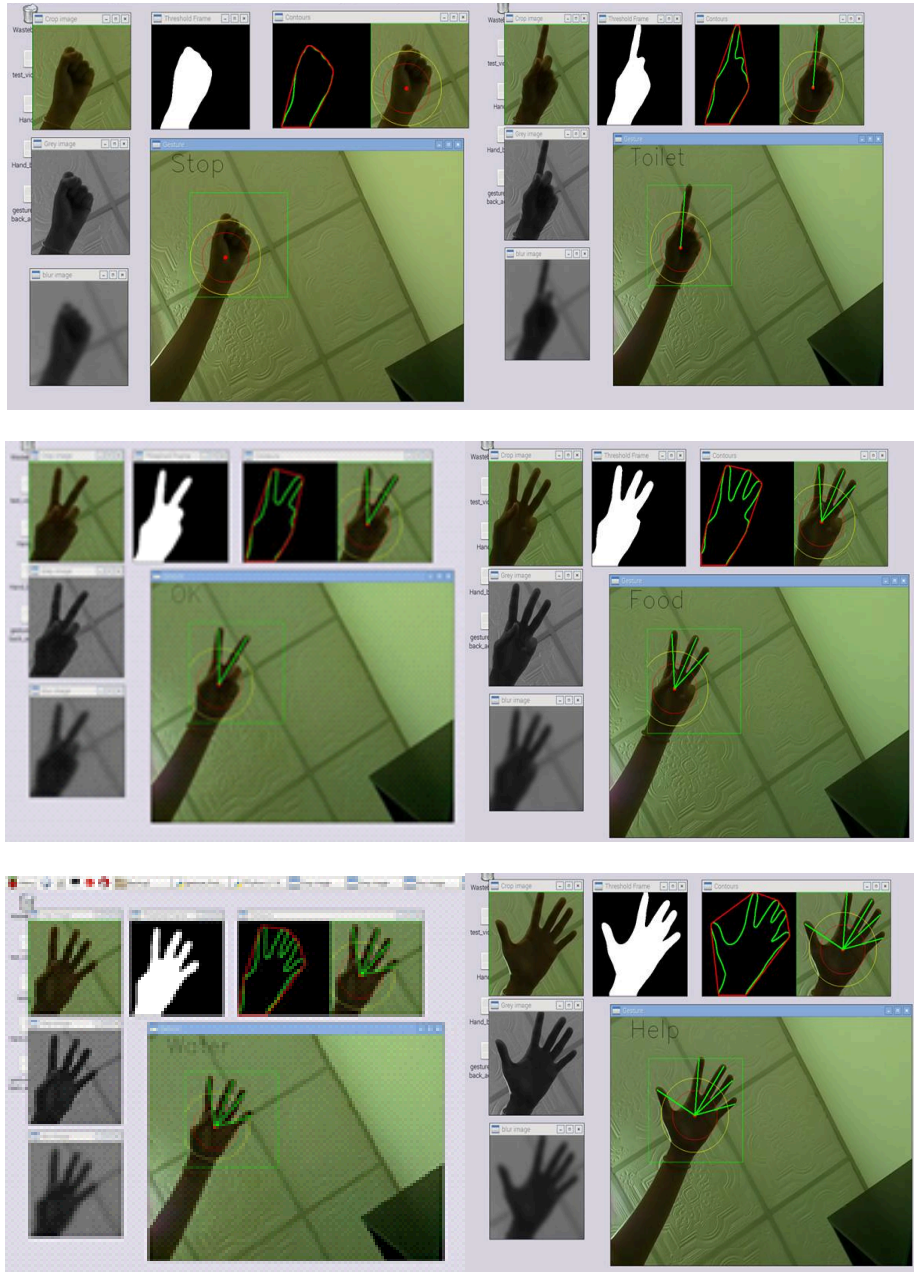


Fig. 14. The results of the hand gesture recognition system.

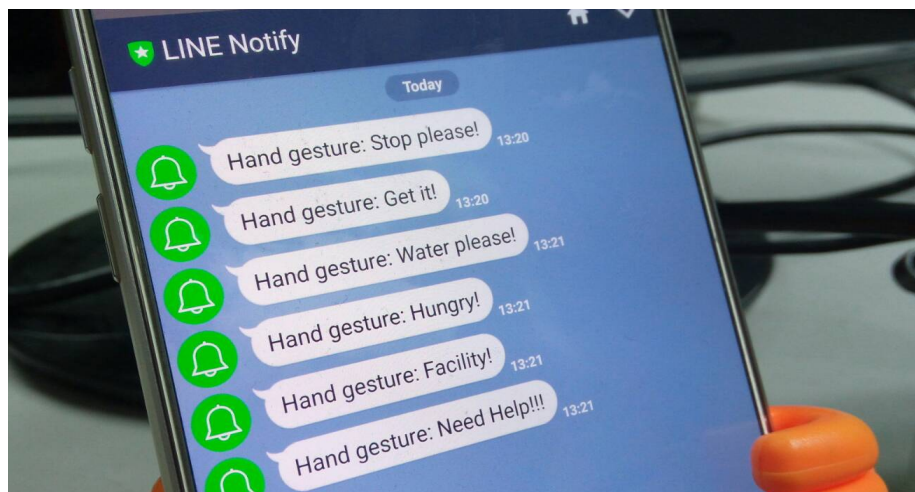


Fig. 15. The results of the hand gesture recognition system in the line application.

## 6 Conclusions

This paper proposed the hand gesture recognition techniques applied for the elderly. The process consists of three main parts, including detection, feature extraction, and classification using contour, convex hull, and rule-based algorithms, respectively. The system focuses on vision-based hand gesture recognition of the six static hand gestures. From the experimental results, the hand gesture recognition system provides good results in detecting and classifying all the six hand gestures as the lingual descriptor. The system can work in real time application for elderly. The results depend on some conditions such as the static background, the distance between the camera and hand, and the light conditions.

## 7 References

1. C. Acharya, H. Thimbleby, and P. Oladimeji, "Human computer interaction and medical devices," in *Proceedings of the 24th BCS Interaction Specialist Group Conference*, 2010, pp. 168-176.
2. J. M. Ortman, V. A. Velkoff, and H. Hogan, "An aging nation: the older population in the United States," *Washington, DC: US Census Bureau*, pp. 25-1140, 2014.
3. Ankit Chaudhary and Jagdish L. Raheja, "A Health Monitoring System for Elder and Sick Persons," presented at the International Journal of Computer Theory and Engineering, Vol. 5, No. 3, June 2013, 2013.
4. Siddharth S. Rautaray and Anupam Agrawal, "Vision based hand gesture recognition for human computer interaction: a survey," *Artificial Intelligence Review*, vol. 43, pp. 1-54, 2015. <https://doi.org/10.1007/s10462-012-9356-9>
5. S. Mitra and T. Acharya, "Gesture recognition: A survey," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 37, pp. 311-324, 2007. <https://doi.org/10.1109/TSMCC.2007.893280>



6. S. Athavale and M. Deshmukh, "Dynamic Hand Gesture Recognition for Human Computer interaction; A Comparative Study," *International Journal of Engineering Research and General Science*, vol. 2, pp. 38-55, 2014.
7. C. Shan, T. Tan, and Y. Wei, "Real-time hand tracking using a mean shift embedded particle filter," *Pattern recognition*, vol. 40, pp. 1958-1970, 2007. <https://doi.org/10.1016/j.patcog.2006.12.012>
8. C.-C. Hsieh and D.-H. Liou, "Novel Haar features for real-time hand gesture recognition using SVM," *Journal of Real-Time Image Processing*, vol. 10, pp. 357-370, 2015. <https://doi.org/10.1007/s11554-012-0295-0>
9. H.-I. Suk, B.-K. Sin, and S.-W. Lee, "Hand gesture recognition based on dynamic Bayesian network framework," *Pattern Recognition*, vol. 43, pp. 3059-3072, 2010. <https://doi.org/10.1016/j.patcog.2010.03.016>
10. G. Plouffe, A.-M. Cretu, and P. Payeur, "Natural human-computer interaction using static and dynamic hand gestures," in *Haptic, Audio and Visual Environments and Games (HAVE), 2015 IEEE International Symposium on*, 2015, pp. 1-6. <https://doi.org/10.1109/have.2015.7359473>
11. Chonbodeechalermroong, A., Chalidabhongse, T. H.: Dynamic contour matching for hand gesture recognition from monocular image. In: Computer Science and Software Engineering (JCSSE), 2015 12th International Joint Conference on, pp. 47-51. IEEE (2015) <https://doi.org/10.1109/jcsse.2015.7219768>
12. Peterfreund, N.: Robust tracking of position and velocity with Kalman snakes. In: Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 21, no. 6, pp. 564-569 (1999) <https://doi.org/10.1109/34.771328>
13. Goodrich, M. T., & Tamassia, R.: Algorithm design: foundation, analysis and internet examples. John Wiley & Sons (2006)
14. M. M. Youssef, "Hull convexity defect features for human action recognition," University of Dayton, 2011.
15. T. Hachaj and M. R. Ogiela, "Rule-based approach to recognizing human body poses and gestures in real time," *Multimedia Systems*, vol. 20, pp. 81-99, 2014. <https://doi.org/10.1007/s00530-013-0332-2>

## 8 Acknowledgement

This research has been supported by Embedded System and IC Design research laboratory, Department of Electrical Engineering and Chulalongkorn Academic Advancement into Its 2nd Century Project. The Scholarship from the Graduate School, Chulalongkorn University to commemorate the 72nd anniversary of his Majesty King Bhumibala Aduladeja and the 90<sup>th</sup> Anniversary of Chulalongkorn University Fund (Ratchadaphiseksomphot Endowment Fund) are also gratefully acknowledged.

## 9 Authors

**Thittaporn Ganokratanaa** received the B.Sc. with First-class Honors in Media Technology from King Mongkut's University of Technology Thonburi, Thailand, in 2014. She is studying in the M.S. in Electrical Engineering at Chulalongkorn University, Thailand, where she receives Chulalongkorn University Graduate Scholarship to Commemorate the 72<sup>nd</sup> Anniversary of His Majesty King Bhumibol Adulyadej and

The 90<sup>th</sup> Anniversary of Chulalongkorn University Fund (Ratchadaphiseksomphot Endowment Fund).

**Suree Pumrin** received the B.Eng. in Electronics Engineering, King Mongkut's Institute of Technology Ladkrabang, Thailand, M.S.E. in Electrical Engineering, Arizona State University, U.S.A and Ph.D. in Electrical Engineering, University of Washington, U.S.A. Dr. Suree Pumrin is an assistant professor in Department of Electrical Engineering, Chulalongkorn University, Thailand.

Article submitted 18 April 2017. Published as resubmitted by the authors 20 May 2017.