REAL-TIME NUMERICAL 0-5 COUNTING BASED ON HAND-FINGER GESTURES RECOGNITION

[1]ABD ALBARY SULYMAN, [2]ZEYAD T. SHAREF, [3]KAMARAN HAMA ALI FARAJ3, [4]ZAID AHMED ALJAWARYY, AND [3]FAHAD LAYTH MALALLAH

[1]Computer and Information Engineering, Electronic Engineering, Ninevah University, Iraq.
[2] College of Engineering, Ahlia University, Bahrain.
[3]Computer Science / Cihan University / Sulaimaniya, Kurdistan Region, Iraq.
[4]Faculty of Science and Technology University of Human Development / Sulaimaniya, Kurdistan Region, Iraq.
[4]fahad.layth.86@gmail.com ,

ABSTRACT

A well Pointing out by hand for originating some gestures is highly useful in terms of human computer interactions especially when mute people desire to speak something, here a difficulty is raised by delivering their message to the outside world. Therefore, these people can do easily some tasks by drawing a gesture in air using their hands in front of a computer camera which translates these gestures to a speech or text to be understood by other people. Part of hand gesture recognition is counting by hand. This paper proposes a new technique describing hand gesture numerals which are from 0 to 5 that are pointed out by people to be understood by a computer. The technique is implemented by reading a frame as an image then extracting only hand by using YCbCr colour space filter. Then, it is converting to black and white image. After that, number is assigned to a gesture by counting number of flip as white to black from left to right on an intelligently selected path to be scanned on. The experiment was conducted using 180 random hand gesture frames taken from random people, the result of this recognition rate is recorded as 98%.

Keywords: *Hand Gesture Recognition, Feature Extraction, Image Processing, Computer Vision, Data Science.*

1. INTRODUCTION

Recently there is a great emphasis in the human-computer interface (HCI) research to gain the most efficient method of interfaces between manipulation skills of humans and computer hardware. The interaction with the computing devices has progressed to as it has become necessity, as well as, human being cannot live without it. Now, the technology became embedded into human's daily lives such as for work, shopping, communicating and entertainment etc. [1]. The human-computer interface can be explained as the point of communication between the human user and the computer system, in which the flow of information between the human and computer is considered as the loop of interaction. One of this interaction can be done through human body language that is one of the most well-known ways through which humans are able to communicate nonverbally. Body language constitutes many forms such as face movement and pose, facial expressions and eye gaze [2], arm and hand gestures (our proposed idea about special part of hand gestures), and torse movement and posturing [3]. Gesture recognition refers to the process of understanding and classifying meaningful movements of a human's fingers, hands, arms or head [4]. Hand gesture as a natural, intuitive, and convenient way of human-computer interaction will greatly ease the interaction process [5] as well as hand gesture is deemed as biometric modality [6]. Based on the methods used to capture the gestures, the literature on hand gestures recognition techniques can be classified into two categories: Vision-based Gestures Recognition (VGR) [7] and Sensor-based Gestures Recognition (SGR) techniques [8].
In this research, a new technique is proposed and implemented to translate hand gesture into number

recognized by computer, which is able to work with low resolution cameras of systems. The proposed technique can detect motion of hand and count the number of fingers in sufficient expected environment, as well as, is able to work accurately for the hand placed at position of about half meter away from webcam. This type of recognition has many functions and services such as helping mute people of communicating with others, contactless control of different life sectors especially to avoid dangerous touched system and etc.

This paper is organized as follows; Section 2 is on literature review of counting by hand gesture recognition, Section 3 explains the methodology of the proposed technique with testing and analysis, Section 4 is dedicated with the experiments description, Section 5 comprises results and discussions. Finally, Section 6 concludes this research with a possible future work.

2.   LITERATURE REVIEW

Hand Gesture Recognition as it is available in the literature can be largely divided into two kinds, Touch-Based and Contactless-Based hand gesture. The first kind can be described as doing multi-hand gesture by touching the human hand to scanning devices [9], while the second one is characterized as hand gesture signals are transmitted to the computer system for recognition operation remotely without any touching [10-12]. Moreover, contactless-Based has also two types. Firstly, handheld Device-Based hand gesture which uses sensors attached to a human hand in order to transduce hand waving signals to the computer systems for determining the hand posture or trajectories. With this technique user carries set of cables connected to the computer, for example, glove based hand gesture as in [13], another example for camera handheld using for tracking trajectories information for the hand gesture [14], and also example of using an intuitive six degrees-of-freedom wireless inertial motion sensor that is used as a gesture and motion control input device with a sensor fusion algorithm [15], also in this work [16] an inertial sensor held with an human arm is used to input gesture to a kinect receiver, another work using Kinect sensor to control Robot with hand gesture is in [17]. Secondly, Vision-Based hand gesture, which is described as computer vision to the hand gesture by using camera [3], for example, air-writing characters can be recognized as the same as to motion gestures in free space by hand [18]. Furthermore, the Vision-Based recognition of the hand gesture falls also into three kinds as seen in the literature: statics hand gesture,

which is based on the shape, appearance and geometry of hand per image such as recognizing "stop" or "ok" sign [19]. For instance, static hand gesture was recognized by using 52 hand shape features comprises of bones length and width, palm characteristics and relative distance relationships among fingers, palm center and wrist with using ANN and SVM [20], dynamic hand gesture is a sequence of hand shapes as a sequenced features collected from a spatial transformation (hand movement) such as rotation, translation, scaling/depth variations etc, it is also can be defined as a spatial-temporal pattern [21]. The motion of the hand can be described as a temporal sequence of points with respect to any point of the hand [22], for example dynamic hand gestures have been recognized by using Leap Motion Controller (LMC) to extract feature vector to be fed to the Hidden Conditional Neural Field(HCNF) as a classifier [23]. Finally, the third type named hybrid gesture recognition, in this approach a combination of both static and dynamic gestures is executed in a real-time processing, for example in [24] a real-time solution for the recognition tracks and recognizes in real time hand gestures based on depth data collected by a Kinect sensor is presented, another hybrid gesture recognition is in [25], this work targets real-time recognition of both static hand-poses and dynamic hand-gestures in a unified open-source framework. In this paper an efficient algorithm for the hybrid gesture recognition type is proposed for recognizing finger-hand based count number 0 till 5 for the hand as it is noticed that there is no a sufficient algorithm for a computer to do this task.

3.   METHODOLOGY

Hand gesture data is input to a computer system by reading RGB image frame that contains the hand gesture as shown in Figure 1(1), the RGB image is converted into YCbCr color space by using these filtering: Cb_min=77, Cb_max=120 , Cr_min=137 , and Cr_max=163 as shown in figure. 1(2). YCbCr is beneficial for eliminating the illumination which is sensitive against light changes. After that the frame is undergone some image processing operations such as converting to black and white image with threshold 0.5 between "0" and "1". Then, median filter is used for noise removing. Then, removing any object which has 300 pixels or less than 300 as pixel area, afterward some morphological operations like image eroding in order to sharpen the object. Also, filling any hall which is surrounded by white pixel to have the only one completed object is applied. Then, searching

for the largest object which is certainly hand object and remove all the small objects which are smaller than hand object, from this it is ensured that the remained object is only hand object from the cluttered background as shown in Figure 1(3). Finally a border is drawn around the hand for the original image frame to set the target region of interest (ROI) by extracting the four border points, then plotted to the original image as shown in Figure1(4).

3.1. Recognition Algorithm (Proposed Technique)

Once the ROI is accessed, the proposed algorithm is kicked off to finally output 6 numbers starting from 0 until 5 based on the human hand gesture. Steps of the algorithm are as follows:

5- Combine left & right slops to be a scanned path.
6- Calculate number of flips based only on the scanned path.

Flip number is defined as the counter of alternating between "0" and "1" pixel of the image. Flip number is deemed significantly useful to describe what is inside black and white image. Below is a pseudo-code for extracting the number of flips of a black and white image:

```
first_value=(1,mid_y);
for x=min: x_max
    if ( first_value ~= Object_array(x,mid_y))
        first_value= Object_array(x,mid_y);
        flip_num=flip_num+1;
    end
end
```
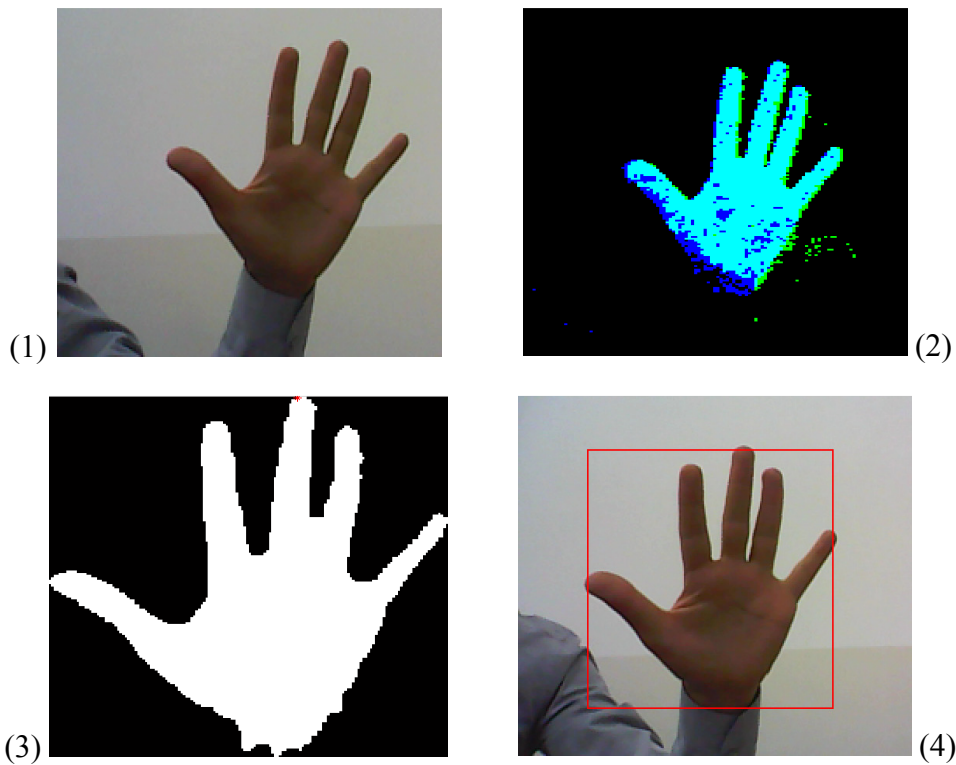

Figure 1: Hand Gesture Stages for Detecting and Foreground Segmentation

1- Get Top Middle Finger (TMF).
2- Move Down of the middle finger quarterly to be reference point (Ref_Pnt).
3- Draw right line slop from (Ref_Pnt).
4- Draw left line slop from (Ref_Pnt).

To elaborate the proposed technique, first step is searching for the top middle finger as shown below in Figure 2(1), the star (*) symbol is affixed on the top middle finger, also labeled with TMF, in which this operation is implemented by scanning from left to right and top to bottom of the black and white image, the first "1" pixel is considered as the top of

the middle finger, then store the address of this pixel. Second step, it is going down with a quarter distance of the middle finger of the selected point in step 1. Here, to guarantee drawing slops without missing any other finger of the hand. Step 2 is labeled as *Ref_Pnt* in Figure 2(1). Third step is to draw a right slop starting from *Ref_pnt* going down to the right end of the image as shown in Figure 2(2), the idea of extracting the right slop is by incrementing one to both rows and columns to get new extracted scanned right curve, the algorithm programming as pseudo-code is shown in the following:

```
    y_new=y_ref+x;
  ry=ry+1;
  right_slop_y(ry)= y_new;
end
```

Similarly, the forth step of the proposed technique is to draw a left slop starting from Ref_pnt going down to the left end of the image as shown in Figure 2(3), the idea of extracting the left slop is by decrementing one to image rows and incrementing one to the columns to get new extracted scanned left curve, the algorithm programming is shown below:
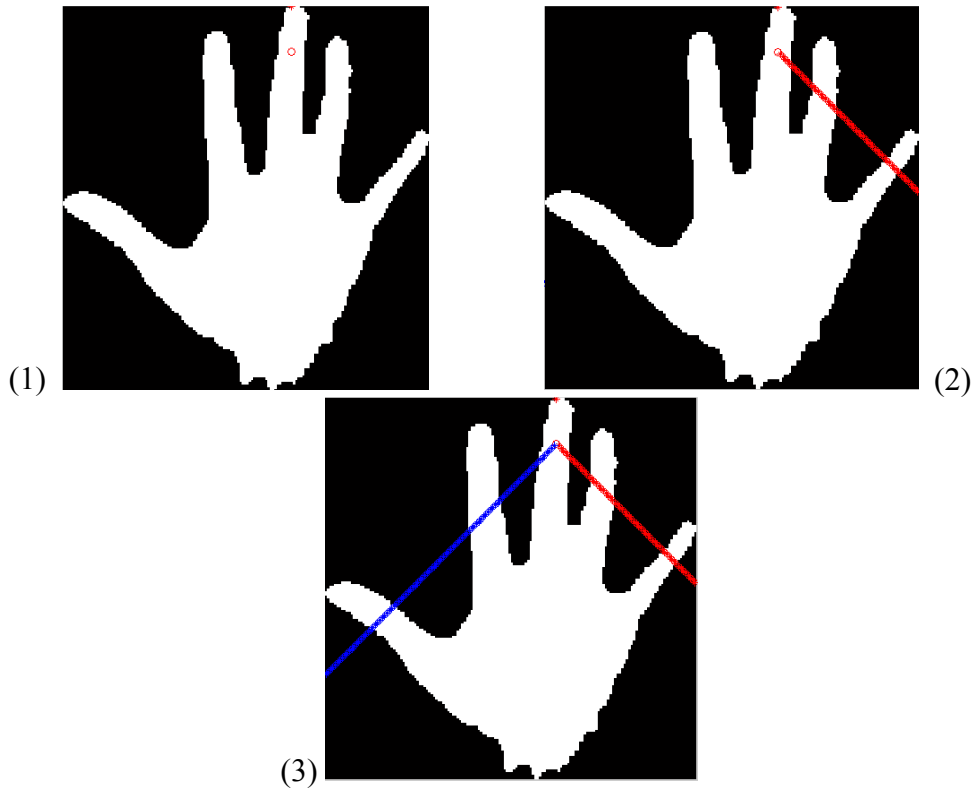

Figure 2: Extracting a Scanned Path of the Proposed Technique.

```
right_slop_x=0;rx=0;
  right_slop_y=0;ry=0;
  size_slop=1;
  for x=1:size_slop:(x_max-x_ref)-1
    x_new=x_ref+x;
  rx=rx+1;
  right_slop_x(rx)=x_new;
```

```
 size_slop=1;
 left_slop_x=0; lx=0;
 left_slop_y=0;ly=0;
 left_size=x_max-length(right_slop_y);
 while (left_size~=0)
   left_size=left_size-1;
   lx=lx+1;
```

x_new=x_ref-lx;
left_slop_x(lx)= x_new;
    ly=ly+1;
y_new=y_ref+ly;
left_slop_y(ly)= y_new;
  end

3.2. Hand Gesture Recognition

To explain all the states of the numerical hand gesture starting from number zero recognition. The recognition is by using expert system based on extracted features based on number of flips that have been extracted followed the scanned path (right and left flips). With number zero only the recognition is based on the solidity, which is defined as the ratio (S) of the foreground area to the background area as shown in equation (1), Here, the threshold is set to 0.8 after extensive experiments iterations. If solidity ratio (S) more than 0.8, then it means the hand gesture is zero,

otherwise will be other numbers ranging from 1 to 5 according to the flips numbers. Figure 3 depicts the tracked zero hand gesture at left, while at right zero hand gesture with applied proposed algorithm to it.

$$S = \frac{Area\_S}{Convex\_Hull\_Area} \qquad (1)$$

In order to predict number one to the hand gesture sign, number of flips should be double of the predicted assigned number. In other words, number of flips for one hand gesture is 2 flips. Figure 4 illustrates the tracked hand gesture of number one at the left and the right of the figure is the hand gesture together with the applied proposed algorithm. It is clear that the flip number is 2 following the scanned path (left & right slop drawn from the reference point.

To assign number two to the hand gesture, number of flips should be double of the predicted assigned number. That means, number of flips for one hand gesture is 4 flips. Figure 5 illustrates the tracked hand gesture of number two at the left and the right of the figure is the hand gesture together with the
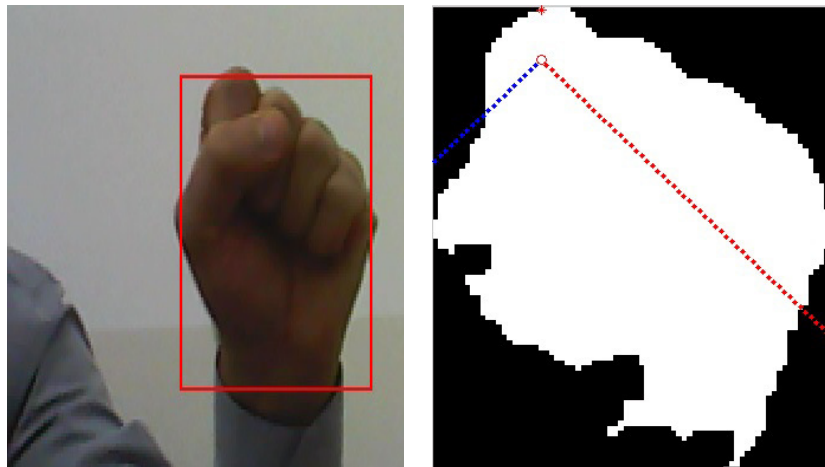

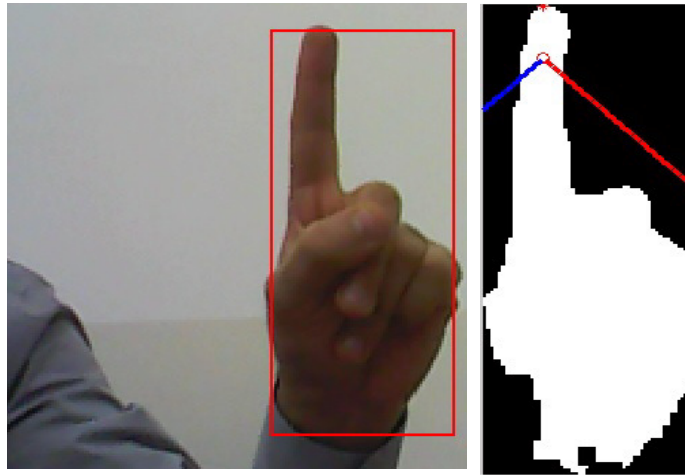*Figure 3: Zero Tracked Hand Gesture at Left and its Proposed Algorithm with Zero Case at Right.*

Figure 4: Number One Tracked Hand Gesture at Left and its Proposed Algorithm at Right.
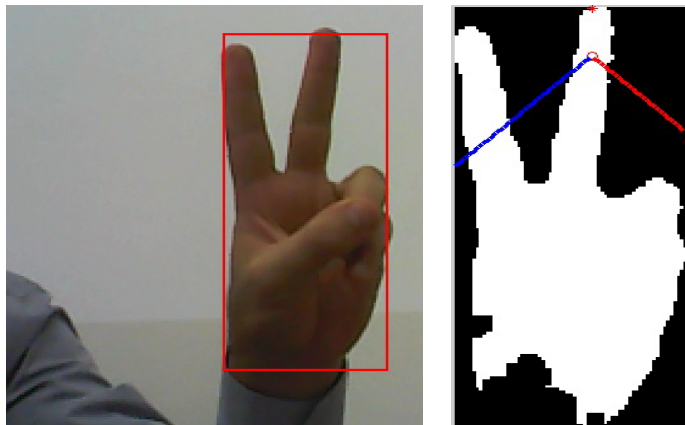

Figure 5: Tracked Hand Gesture Number Two at Left and its Proposed Algorithm at the Right.

About figure 6. it has also two illustrated images, the one lies at the left is related to the tracked hand gesture for number three while the image at the right is number three together with the preprocessing and after the proposed algorithm, which contains both left and right slops that are drawn from the *Ref_pnt* to be a scanned path. Here, the prediction is number three, if only the number of flips is 6. Next, if the number of flips is eight, which means the predicted number of the hand gesture is four. Figure 7 shows the hand gesture number four for both the tracked and after the proposed algorithm by drawing the two slops

starting from the reference point (*Ref_Pnt*). It is clear if the number of flips which is counted from left to right based on the slop is 8, and then the number will be recognized as four.

Finally, in case flips number is computed by the computer as 10, then, the predicted number is five of the hand gesture. Figure 8 illustrates both the tracked and proposed algorithm on the hand gesture showing the left and right slops. Briefly the prediction can be modeled by a mathematical equation as in (2), which describes the predicted number (P_N) based on flips number (*Flp*):
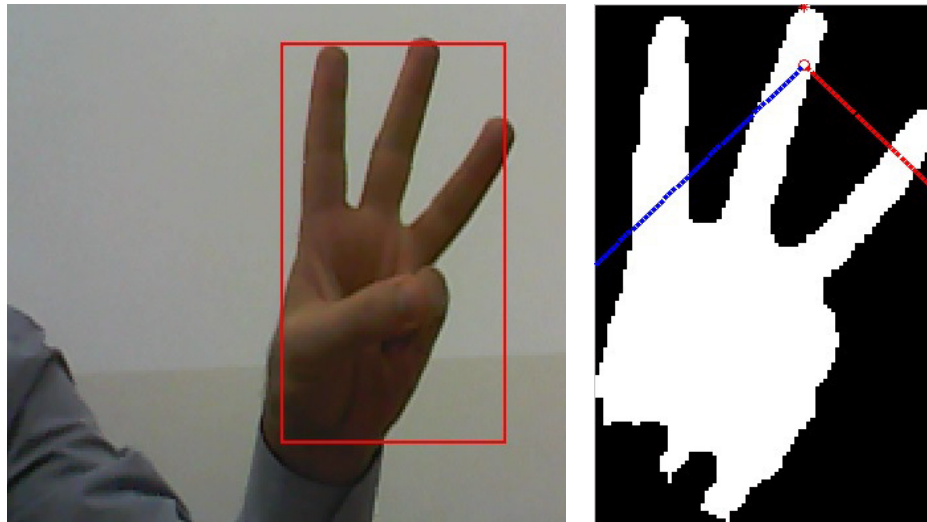
$$P_N = \frac{FLP}{2} \qquad (2)$$

Figure 6: Tracked Hand Gesture Number Three at Left and its Proposed Algorithm at Right.
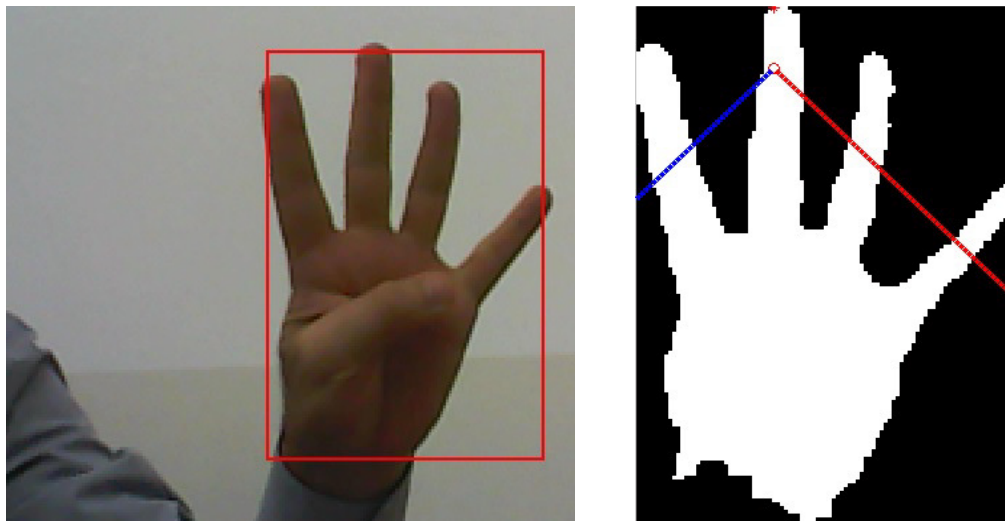

Figure 7: Tracked Hand Gesture Number Four at Left and its Proposed Algorithm at Right.
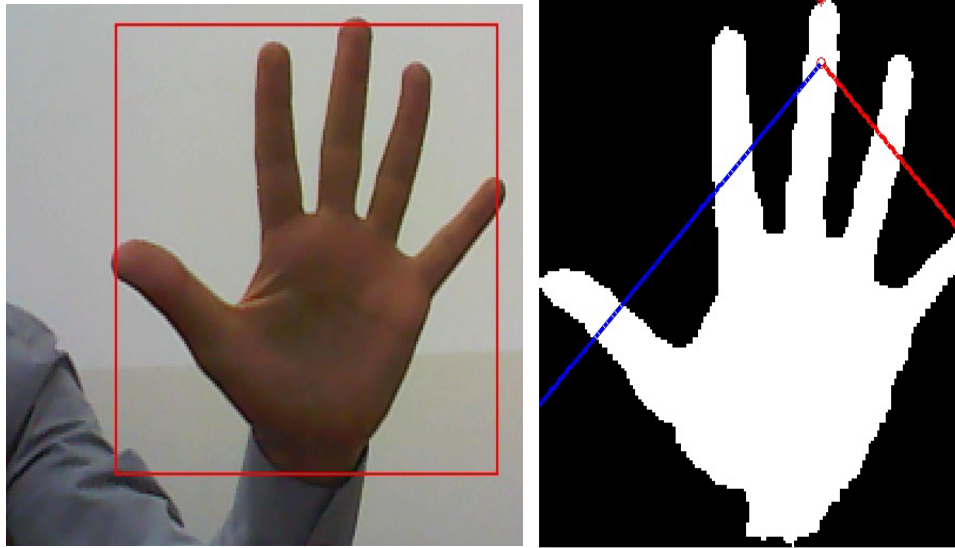
Figure. 8: Tracked Hand Gesture Number Five at Left and its Proposed Algorithm at Right.

4. EXPERIMENT

To evaluate the proposed method, 180 random frames as images have been captured containing hand gesture with different gesture ranging from 0 to 5. Hand gesture has been taken from random individuals asked to originate random hand numbers ranging from 0 to 5. Method of acquisition is by using Camera specified 320x250 pixels and processing by using Matlab 2013 installed into a computer has the following characteristics core2due, 4 G-RAM. Normally, in the verification or identification comparison, there are two possible errors to be measured: False Accept Rate (FAR), which results from the forged template that accepted by the computer system falsely during testing and the second error is False Rejection Rate (FRR), which results from the genuine template that the system recognizes as the query template wrongly [26, 27]. Then, the total accuracy of the system is calculated by subtracting the average error rate from 100% as in (3):

$$Accuracy\ \% = 100\ \% - \frac{FAR + FRR}{2} \quad (3)$$

In this type of research, FAR error does not exist, since there are no forge templates in this experiment. Therefore, FAR is considered to be zero. However, FRR is used for the testing measure to assess the recognition rate, because these hand gesture numbers are considered as genuine templates. In case they are wrongly recognized by computer system, then the FRR increases.

The equations that are used to measure the accuracy of this research are in (4) and (5):

$$Accuracy\ \% = 100\ \% - FRR\ \% \quad (4)$$

$$FRR\% = \frac{Total\_False\_\mathrm{Re}ject}{Total\_True\_Attempt} \times 100\% \quad (5)$$

5. RESULT AND DISCUSSION

The captured samples have been tested by using the proposed algorithm, the result reported in this research as the table 1, which contains hand gesture posture types (0,1,2,3,4,5), total number of one type (captured randomly) and the result field in the table

to output either Match or Mis-match result. For example in Table 1 number of gesture "0" has been repeated 36 times, and all of them have been recognized correctly so that in the match field is written as 36 times. Also, in case hand gesture sign "1", number of random gesture that has been repeated is 30 times, only two of them have been wrongly recognized so that it is written 2 in the field of Mis-Match. About hand gesture "3", only one sample has been wrongly recognized out of 12 samples. In case hand gesture number "2","4", and "5" as total samples captured as 22, 33,47 respectively, all of these samples have been correctly predicted by computer. As over all, only 3 samples are wrongly mismatched out of 180 samples in this research experiment.

Table 1: Overall Results As Match Or Mis-Match Of Hand Gesture Numbers.

Hand Gesture Type	Total Hand Gesture	Total Hand Match	Total Mis-Match	Matching Rate
0	36	36	0	100%
1	30	28	2	93.3%
2	22	22	0	100%
3	12	11	1	91.6%
4	33	33	0	100%
5	47	47	0	100%
Total Gesture Samples	180	177	3	

The results of the experiment showed as a successful accuracy is 98.3 % as False Reject Rate (FRR) is 1.6%. As it is clear from Table 1, four hand gestures (0, 2, 4 and 5) have scored 100% as successful recognition while hand gesture once scored 93.3% and hand gesture 3 scored 91.6 % recognition rate. Part of the faced obstacles during running this experiment was the noise taken with the captured image of the hand gesture. For example, below image in figure 9 has been captured with a noise. This noise affects on the recognition ability of the proposed algorithm. In this example, in spite of the noise taken, the predicted number is two, which is false, instead of one (true). It is worth to mention that the proposed algorithm has been

compared with a method as described here [7] which is considered as a benchmark. The comparison showed that the proposed method has a better accuracy for the 6 hand gestures as 0, 1, 2, 3, 4 and 5 (the scope of this paper). The compared paper has used a scale invariance feature transform (SIFT) algorithm to extract feature vectors. then, vocabulary tree along with gesture along with K-means clustering method that are used to partition the hand postures to ten simple sets as: "one", "two", "three", "four", "five", "six", "seven", "eight", "nine" and "ten" numbers based on the number of extended fingers. Furthermore, the benchmark paper used Kinect camera which has better characteristics for the recognition rate as it has depth factor that is able to consolidate the recognition.

It is clear from Table 2 adapted from [7], number 2 has recognition rate 86.28 %. As compared with the proposed algorithm, number "2" gesture has 100% recognition rate as shown in Table1. Also, it is noticed that number "5" has gotten 83.21% while the proposed method has 100% recognition rate.

Table2: Adapted From [7] Shows The Accuracy Of The Compared Paper .

Gesture	Recognition	(millisecond)
"two"	86.28%	86
"five"	83.21%	95
"seven"	81.33%	76
"eight"	75.56%	88
"ten"	76.14%	89

Accordingly, 6 hand gestures of hand have been recognized well and better than the state-of-the-art. The possible limitation is in lack recognition when the camera is affected by a heavy light during capturing that will be negatively affected on the
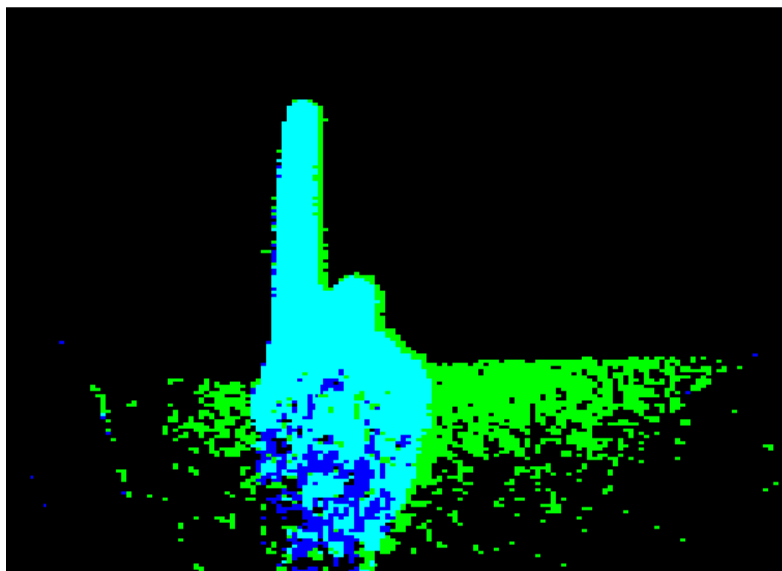
accuracy.

Figure 9: Illustrating Noise Effect on the Predicted Number of the Taken Hand Gesture Image.

6. CONCLUSION

Counting number is important in majority sectors of the life style. This research proved that counting by finger-hand gesture can be originated from mute people by using computer system, through just gesturing by their hands doing normal counting based on their fingers. Technically, an algorithm has been proposed so as to predict number based on the hand gesture. The operation is kicked off by converting the hand object into black and white then searching on the top of the middle finger then moving down quarter of the middle finger then drawing right and left slops starting from the reference point to both left and right sides. This will create a scanned path for calculating the number of flips which will be depended in the number prediction. In a brief, predicting number is recognized based on the flips count divided by two as explained above. The proposed algorithm has been tested and the performance of it, is 98% resulted from the experiment that has been conducted on 180 hand gesture images taken from random people and random numbers ranging from 0-5. This promises with an improvement in the computer vision techniques as Human computer Interaction.

For the future work, the proposed algorithm might be developed combining both hands at the same time to count from 0 to 9. Also it might be improved to have several reading print of the gesture to express on infinity numbers. This research also might be a starting point to improve autistic people communication.

REFRENCES:

[1] S. S. Rautaray and A. Agrawal, "Vision based hand gesture recognition for human computer interaction: a survey," *Artificial Intelligence Review,* vol. 43, pp. 1-54, 2015.

[2] L. Farmohammadi and M. B. Menhaj, "Facial Expression Recognition Based on Facial Motion Patterns," *Indonesian Journal of Electrical Engineering and Informatics (IJEEI),* vol. 3, pp. 177-184, 2015.

[3] S. Berman and H. Stern, "Sensors for gesture recognition systems," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews),* vol. 42, pp. 277-290, 2012.

[4] A. Akl*, et al.*, "A novel accelerometer-based gesture recognition system," *IEEE transactions on Signal Processing,* vol. 59, pp. 6197-6205, 2011.

[5] T. Lu, "A motion control method of intelligent wheelchair based on hand gesture recognition," in *Industrial Electronics and Applications (ICIEA), 2013 8th IEEE Conference on*, 2013, pp. 957-962.

[6] F. L. Malallah*, et al.*, "Irreversible Biometric Template Protection by Trigonometric Function," *International Review on Computers and Software (IRECOS),* vol. 11, pp. 1138-1146, 2016.

[7] M. Hamissi and K. Faez, "Real-time hand gesture recognition based on the depth map for human robot interaction," *International*

Journal of Electrical and Computer Engineering, vol. 3, p. 770, 2013.

[8] H. P. Gupta*, et al.*, "A continuous hand gestures recognition technique for human-machine interaction using accelerometer and gyroscope sensors," *IEEE Sensors Journal,* vol. 16, pp. 6425-6432, 2016.

[9] N. Sae-Bae*, et al.*, "Biometric-rich gestures: a novel approach to authentication on multi-touch devices," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2012, pp. 977-986.

[10] M. S. M. Asaari and S. A. Suandi, "Hand gesture tracking system using Adaptive Kalman Filter," in *2010 10th International Conference on Intelligent Systems Design and Applications*, 2010, pp. 166-171.

[11] J. Doublet*, et al.*, "Contactless hand recognition based on distribution estimation," in *Biometrics Symposium, 2007*, 2007, pp. 1-6.

[12] A. Morales*, et al.*, "Comparing infrared and visible illumination for contactless hand based biometric scheme," in *2008 42nd Annual IEEE International Carnahan Conference on Security Technology*, 2008, pp. 191-197.

[13] G. Murthy and R. Jadon, "A review of vision based hand gestures recognition," *International Journal of Information Technology and Knowledge Management,* vol. 2, pp. 405-410, 2009.

[14] H. Sajid and S. C. Sen-ching, "VSig: Hand-gestured signature recognition and authentication with wearable camera," in *Information Forensics and Security (WIFS), 2015 IEEE International Workshop on*, 2015, pp. 1-6.

[15] S. Patil*, et al.*, "Inertial sensor-based touch and shake metaphor for expressive control of 3D virtual avatars," *Sensors,* vol. 15, pp. 14435-14457, 2015.

[16] K. Liu*, et al.*, "Fusion of inertial and depth sensor data for robust hand gesture recognition," *IEEE Sensors Journal,* vol. 14, pp. 1898-1903, 2014.

[17] M. Biao*, et al.*, "A robot control system based on gesture recognition using Kinect," *Indonesian Journal of Electrical Engineering and Computer Science,* vol. 11, pp. 2605-2611, 2013.

[18] M. Chen*, et al.*, "Air-Writing Recognition—Part I: Modeling and Recognition of Characters, Words, and Connecting Motions," *IEEE Transactions on Human-Machine Systems,* vol. 46, pp. 403-413, 2016.

[19] Z. Ren*, et al.*, "Robust part-based hand gesture recognition using kinect sensor," *IEEE transactions on multimedia,* vol. 15, pp. 1110-1120, 2013.

[20] A. M. Bernardos*, et al.*, "A contactless identification system based on hand shape features," *Procedia Computer Science,* vol. 52, pp. 161-168, 2015.

[21] M.-C. Su, "A fuzzy rule-based approach to spatio-temporal hand gesture recognition," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews),* vol. 30, pp. 276-281, 2000.

[22] P. Vashist and K. Hema, "Dynamic Hand Gesture Recognition Using Neural Networks," *International Journal of Science and Technology,* vol. 2, 2013.

[23] W. Lu*, et al.*, "Dynamic Hand Gesture Recognition With Leap Motion Controller," *IEEE Signal Processing Letters,* vol. 23, pp. 1188-1192, 2016.

[24] G. Plouffe and A.-M. Cretu, "Static and dynamic hand gesture recognition in depth data using dynamic time warping," *IEEE Transactions on Instrumentation and Measurement,* vol. 65, pp. 305-316, 2016.

[25] F. Pedersoli*, et al.*, "XKin: an open source framework for hand pose and gesture recognition using kinect," *The Visual Computer,* vol. 30, pp. 1107-1122, 2014.

[26] F. L. Malallah*, et al.*, "Online handwritten signature recognition by length normalization using up-sampling and down-sampling," *International Journal of Cyber-Security and Digital Forensics (IJCSDF),* vol. 4, pp. 302-313, 2015.

[27] F. L. Malallah*, et al.*, "Non-Invertible Online Signature Biometric Template Protection via Shuffling and Trigonometry Transformation," *International Journal of Computer Applications,* vol. 98, 2014.