

Original Research Paper

Static Hand Gestures: Fingertips Detection Based on Segmented Images

¹Ahmad Yahya Dawod, ²Md Jan Nordin and ³Junaidi Abdullah

¹Pattern Recognition Research Group, Centre for Artificial Intelligence Technology (CAIT), Malaysia

²Pattern Recognition Research Group, Centre for Artificial Intelligence Technology (CAIT),

Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia (UKM), 43600 Bangi, Malaysia

³Assistive Technology SIG, Faculty of Computing and Informatics, Multimedia University (MMU), 63100, Cyberjaya, Selangor, Malaysia

Article history

Received: 17-11-2015

Revised: 16-01-2016

Accepted: 02-02-2016

Corresponding Author:

Ahmad Yahya Dawod

Pattern Recognition Research

Group, Centre for Artificial

Intelligence Technology

(CAIT), Malaysia

Email: ahmadyahyadawod@gmail.com

Abstract: Fingertips detection is important for recognition of static gesture in sign language. This paper presents a new method that is based on YCbCr colour space and skeletonization for fingertips detection towards gesture segmentation and recognition. The method begins with the conversion of an image in RGB colour space into YCbCr colour space. Then, the chrominance Cb and Cr are extracted from the YCbCr colour space. A thresholding technique, which is based on a pre-defined range for Cb and Cr components representing the skin colour value is used to extract the hand from the background and to achieve the binary image. A morphological processing is performed on the obtained binary image to remove noise and unwanted image pixels. The candidate fingertips position is then calculated based on skeletonization algorithm and tracing process. The centroid is subsequently found, after which the Euclidean distances between the pixels' coordinates that belong to the candidate fingertips and the centroid are calculated towards validating their representation of the actual fingertips. Based on the proposed method, the fingertips of twenty-six American Sign Language (ASL) alphabets hand sign samples were detected successfully and the gesture was correctly recognized with an accuracy of 96.3% during the conducted experiments.

Keywords: Static Gesture, YCbCr Colour Space, Morphological Processing, Hand Centroid, Skin Colour Segmentation

Introduction

For years, sign language has been among several communication options utilized by hard-of-hearing or deaf people. It is considered as a language employing signs created through the movement of the hands in combination with body postures and facial expressions (Nguyen *et al.*, 2013). Gesture language identification, with its high applicability, likewise belongs to the areas that are explored to assist deaf people's integration into the community. Researchers utilize specialized equipment (i.e., gloves or recognition techniques that are based on image processing with the use of computers and cameras) (Rautaray and Agrawal, 2015) towards detecting gestures.

Due to its potential use for human-computer interaction, hand gesture recognition has become a very active topic for research. A range of non-verbal communication is considered part of human hand gestures; for instance, simple ones, such as pointing,

to more complex gestures, such as sign language. Because hand appearance is variable, accurately detecting hands in video or still images continues to be a challenging problem. In terms of detecting fingertips towards recognizing gestures, a majority of the methods of detection are based on two steps. Hand segmentation is always involved in the first step, which is followed by fingertips detection (Bhuyan *et al.*, 2012). Through hand segmentation, the areas of interest, which is intended towards finding the fingertips, becomes smaller; thus, the process of detection becomes faster. Most methods of hand segmentation, however, cannot get good results because of certain conditions, such as cluttered background, fast hand motion and poor lighting condition. The results of fingertip detection are usually invalidated by poor hand segmentation method performance (Raheja *et al.*, 2012). Two trends emerge from the study of the problem of hand gesture

recognition. These trends correspond to two types of data: Static and dynamic gestures.

The hand shape and pose information are commonly utilized towards describing static gestures while hand motion is used towards representing dynamic gesture types, Hasan and Abdul-Kareem (2014). However, a close relationship has been observed between these trends; this is because many dynamic gesture identification approaches are developed according to solutions of static gesture recognition. Thus, there is consideration of static gesture identification as the research basis when it comes to general gesture recognition (Shen *et al.*, 2012). Most hand segmentation research is based on the detection of the skin colour. In this approach, the main idea rests on the assumption of skin colour's difference from other objects' colours and that its distribution leads to the formation of a cluster in a few specific colour spaces.

Research activities pertaining to skin segmentation have mainly progressed in these two directions: Statistical colour models and generic colour models (Bhuyan *et al.*, 2015; Cooper *et al.*, 2012). A fixed colour range is defined by the so-called generic colour model towards separating skin from non-skin pixels. In general, the fixed colour range is accomplished empirically utilizing a few collected training instances. Using generic colour models by itself cannot handle human skin variations and illumination. (Wenkai and Lee, 2012; Khan and Ravi, 2013) used a classic YCbCr colour space for hand segmentation. Chrominance Cb and Cr were extracted from YCbCr colour space; a binary hand segmentation result was obtained through taking the chrominance CbCr values as the elliptic equation model's parameters.

This paper is organized as follows. In the section, 2 states of the art of fingertips detection are presented. In section 3 Proposed Colour spaces used for skin detection is introduced. In section, 4 The Proposed Methodology is described. The experimental results are introduced in section 5. The paper is concluded and future work is presented in section 6.

State of the Art

These days, the hand gesture is counted among the non-verbal methods of communication which can be employed in Human-Computer Interaction (HCI). Moreover, this method has been extended for use in Human-Robot Interaction (HRI). The hand images can be easily obtained through the use of a camera attached to the computer or the robot. A few processing steps, however, should be done in order to recognize the meaning of the hand gesture. Before the hand gesture recognition process, some hand features, such as its centre, the finger directions and the fingertip positions, should be extracted. Ibraheem *et al.* (2013) put forward a proposal for a method towards hand segmentation based on skin colour with the use of the YCbCr colour space. In

order to detect the skin, YCbCr was utilized for the extraction of the luminance component of the colour space. However, the method is unreliable because of the difficulty to have a model for skin colour that can represent people of different ethnicity and can be applicable under various lighting conditions.

Vieriu *et al.* (2013) proposed a simple system that is developed by utilizing artificial neural network and skin colour segmentation. However, the extracted features resulting from these methods do not have sufficient hand shape information, which is a key attribute that has a bearing on the recognition accuracy. Thus, the level of generalization for these methods is not really high.

Dung and Mizukawa (2009) suggested finding the position of the fingertips according to a set of new feature pixels. This method emphasizes on finding the position of the fingertips in an image frame which includes both face and hands. The extraction of these feature pixels, which are called distance-based feature pixels, is made through a simple technique that is based on the connected component labelling image's distance transformation in order for the binary skin colour detection image to be found quickly and easily. Thus, the process of hand gesture recognition and hand tracking will be performed with more ease and accuracy. Furthermore, no complex calculations are required for all these above-mentioned processing steps; thus, hand gesture recognition or hand tracking system is allowed to achieve real-time performance.

Joshi and Vig (2015) gave a presentation about a computer vision system for the recognition of six American Sign Language (ASL) hand gestures through the use of a combined hand geometry and colour information. The overall accuracy rate that was given for this system is 89%.

Colour Spaces used for Skin Detection

Skin detection is defined as the process of looking for skin-coloured pixels and regions in a video or an image. Typically, this process is utilized as a pre-processing step towards finding regions in an image which potentially have human hands and faces. The pixels are usually transformed into appropriate colour space to detect skin in an image. Then, a skin classifier is used to label whether the pixel can be classified as skin or non-skin (Shakti, 2013).

RGB Colour Space

RGB colour space is a type of mixed colour space. It describes colour space through primary colours namely red, green and blue. In RGB, the chrominance and the luminance components are mixed, which is the reason for RGB's non-selection as a preferred colour space used in analysis and segmentation algorithm (Al Tairi and Sariipan, 2014).

YCbCr Colour Space

The YCbCr colour space is typically utilized for skin colour segmentation. Luminance is reflected by the component Y. The Cb component is reflective of the difference between a reference value and the blue component (Qiu-yu *et al.*, 2015). The Cr component, on the other hand, reflects the difference between a reference value and the red component. In the conventional algorithm that detects skin, no consideration is given to the luminance component Y since the chrominance components Cb and Cr is nearly independent of that colour space's luminance component (Ayala-Ramirez *et al.*, 2011). In a few works, therefore, only the Cb and Cr components were used towards classifying each pixel as skin or non-skin. In this research study, the YCbCr colour space is chosen instead of RGB because of the following reasons:

- The YCbCr possesses the same structure theory with a person's vision (Kaur and Kranthi, 2012)
- The transformation between YCbCr and RGB colour space is linear. The computation process is, therefore, simple. It is characteristically better-clustered than the other colour mode
- The colour model is extensively utilized in television as well as in other vision devices
- Under various conditions of illumination, YCbCr has a tiny overlap between the skin and non-skin colour
- RGB colour space has a very high sensitivity towards intensity differences

The formulas for the conversion from RGB to YCbCr and vice versa are illustrated in "Equation 1 to 3". Figure 1 shows YCbCr colour space is represented:

$$Y = 0.299R + 0.587G + 0.114B \quad (1)$$

$$Cb = (B - Y) * 0.564 + 128 \quad (2)$$

$$Cr = (R - Y) * 0.713 + 128 \quad (3)$$

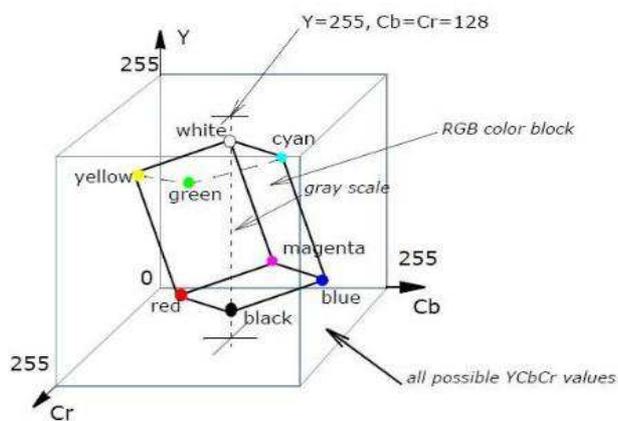


Fig. 1. YCbCr colour space

Proposed Methodology

The four key stages of this research study's proposed methodology are presented below. These consist of Image Acquisition, Segmentation, Morphological Operation and Fingertips Detection. Figure 2 illustrates the flowchart of our approach.

Image Acquisition

The first step involved the acquisition of the hand gestures of the American Sign Language alphabets. These images, which may include varying illumination and cluttered background conditions, were then captured utilizing a camera. Normal static hand gesture images were also photographed, aside from the static ASL. A Logitech USB 2.0 webcam was used with the resolution of 320x240 pixels. In general, the subsequent stages will be less complicated if, one, the image is shot using a simple background and, two, the contrast between the background and the hand is set high. The ASL images were also gathered from a few available online data sets 'http://www.lifeprint.com/asl101/pages-layout/handshapes.htm'. Only some images of symbols, i.e., B, C, D, F, I, K, R, V, W and Claw in the American Sign Language (ASL), are utilized, as illustrated in Fig. 3.

Segmentation

In this proposed method, the images underwent conversion from RGB to YCbCr colour space. When the background image is not complex, simple thresholding was used for hand segmentation. But, when the background is complex and people from various ethnicities were involved in the hand gestures, then the free-form skin colour model was used for hand segmentation, as described in (Dawod *et al.*, 2010a). The thresholding process' condition is also shown in "Equation 4", which produces a binary image where the white pixels (value of 1) represent the hand and the black pixels (value of 0) represent the background:

$$SkinColor(x,y) = \begin{cases} 1 & \text{if } (77 \leq cb \leq 127) \cap (133 \leq cr \leq 180) \\ 0 & \text{Otherwise} \end{cases} \quad (4)$$

The algorithm for the thresholding is stated below:

```

for i = 1:m
    for j = 1:n
        if(Cb(i,j) > 77 && Cr(i,j) < 127 && Cr(i,j) >
            133 && Cr(i,j) < 180)
            ImageWithHand (i, j) = 1;
        else
            ImageWithHand (i, j) = 0;
        end
    end
end
    
```

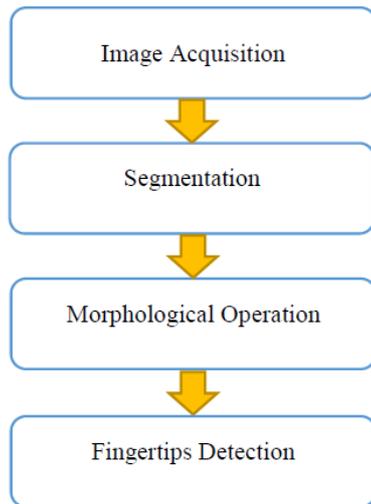


Fig. 2. Schematic view of the proposed approach

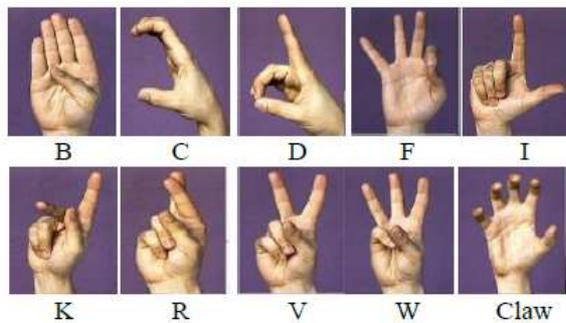


Fig. 3. Some images from the dataset

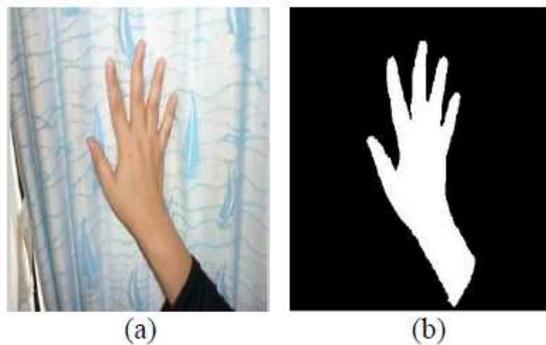


Fig. 4. (a) Before segmentation (b) After segmentation

In the algorithm stated above, C_b and C_r are the matrices that represent the C_b and C_r pixel values in accordance with the YCbCr colour space. ImageWithHand stands for the matrix representing the consequent binary image. Based on the good results that have been provided experimentally, the values of the threshold have been determined in this matrix. Figure 4 illustrates the result of the segmented hand.

Morphological Operation

Image pre-processing is the term employed for image operations at the lowest level of abstraction that intends to enhance the image data. This is done through the removal or minimization of undesired noises or through the enhancement of some image features, which are important for additional processing and analysis. Image pre-processing does not increase the image information's content.

The hand segmented binary image undergoes morphological operations, namely 'Opening' followed by 'Dilation'. In the Opening operation, the small connected components are eliminated; the narrow isthmuses are broken and the contours are smoothed. The Dilation operation employs a disk-shaped structuring element towards shaping the contour and increasing the skin region's size, which will be helpful in the performance of the region labelling operation. Dilation and Erosion are applied with 3×3 structuring elements in the case of this research study as shown in "Equation 5 and 6" This is towards the removal of small background objects and the separation of the hand from the background, as illustrated in Fig. 5.

Let $N \in x^2$, which is the 2D space of the (a, b) and also let $M \in x^2$ be the SE which controls the structure of the morphological operations, then for any binary image I where $N \subseteq I$, dilation and erosion can be defined as in (5) and (6) respectively (Hasan and Mishra, 2012).

Dilation Operator Given N and M sets in x^2 , the dilation of N by M , is defined by:

$$N \oplus M = \left\{ X \mid \left(\overset{\cdot}{M} \right)_x \cap N \neq \emptyset \right\} \quad (5)$$

Erosion Operator of N by M is defined as.

Given N and M sets in x^2 , the erosion of N by structuring element M , is defined by:

$$N \ominus M = \left\{ X \mid (M)_x \subseteq N \right\} \quad (6)$$

The erosion of N by structuring element M is the set of all points x , such that M , translated by x , is contained in N .

Even after the performance of morphological operations, a few background objects which share a similar colour with the skin still remains. A blob detection is done to solve this problem. A unique label is assigned to each blob towards separating it from other blobs. The same labels are also assigned to all the pixels that exist within a blob of 1s that is spatially connected. After selecting the blob that has the highest number of pixels as the blob to represent the hand, all other blobs are removed.

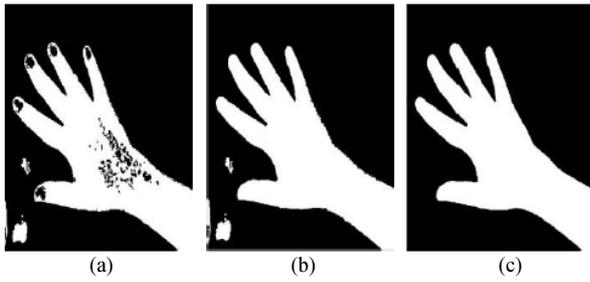


Fig. 5. (a) Before morphological operation (b) After dilation (c) After erosion

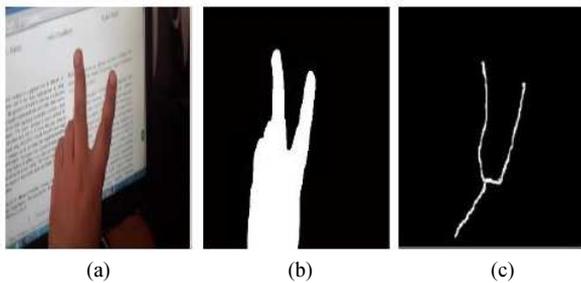


Fig. 6. (a) Source image (b) Hand segmentation image (c) Fingertips skeleton image

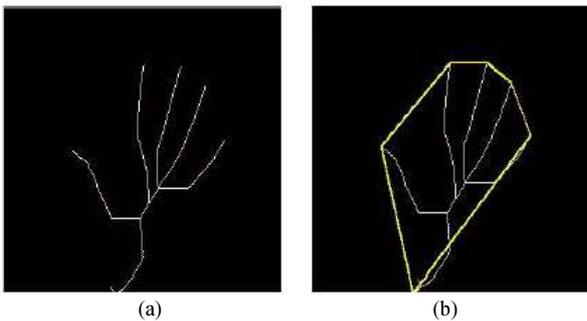


Fig. 7. (a) Skeleton image (b) convex hull

Fingertips Detection

In the skeletonization process, which is used to identify fingers on the hand image, the hand blob was ‘thinned’ until it produced a size of the one-pixel skeleton. Additional details on the algorithm can be seen in (Dawod *et al.*, 2010b; Omkar and Monisha, 2011).

The fingertips skeleton image may contain a few noise pixels, which can be taken out using regional connectivity analysis. Figure 6c illustrates the fingertips skeleton result. Figure 7 shows that for fingertip detection, the fingertip position must obviously be one of the finger skeleton’s endpoints.

The following steps are performed so as to find out whether the endpoints represented the wrist or the fingertips. First, we looked for the endpoint’s convex

hull. We then search for the distance between the neighbouring endpoints. The endpoint that has the biggest distance will be tagged as the wrist while the rest of the endpoints will be tagged as fingertips.

Experimental Results

This experiment is conducted using a PC with Intel Core i7, 2450M, 8G, 1.97GHz processor. MATLAB software is utilized with the Windows 8.1 operating system. The dataset used for the experiment is composed of one single-handed gesture images with different illumination conditions and backgrounds. The image resolution used is 320×240.

The dataset is sourced from the Static Hand Posture Database ‘<http://www.idiap.ch/resource/gestures>’ and ‘<http://www.lifeprint.com/asl101/pages-layout/handshapes.htm>’.

We consider the segmented hand region, its centroid or centre given by (X_C, Y_C) is the centroid of the hand image is calculated. In this research work, the silhouette is calculated through the use of the following formula:

$$X_C = \frac{\sum x_i}{M} \quad (7)$$

where, x_i represents the X-coordinate of each boundary pixel of the image shown in Equation 7:

$$Y_C = \frac{\sum y_i}{M} \quad (8)$$

where, y_i represents the Y-coordinate of each boundary pixel of the image shown in “Equation 8”.

x_i and y_i are the x and y coordinates of the i -th pixel in the hand region, M represents the total number of pixels in the hand region. Lastly, we compute the Δx and Δy centroid of the hand image.

Figure 8 shows the detected fingertips specifically for hand image with a complex background. The results show that the fingertips for all images that have a complex background are successfully detected when all the fingers are open.

Figure 9 shows the detected fingertips for the ASL alphabet datasets. This research’s method was able to detect the fingertips of all the images. However, our method has this limitation: At present, it could not detect fingertips that are closed or those that are bent into the palm. This could be the possible direction of our future work. Figure 9 also shows the estimated centroid along with the detected fingertips.

Figure 10 shows the binary image of hands together with the histogram of detected fingertips. The number of bars represents the number of fingers detected in each image.

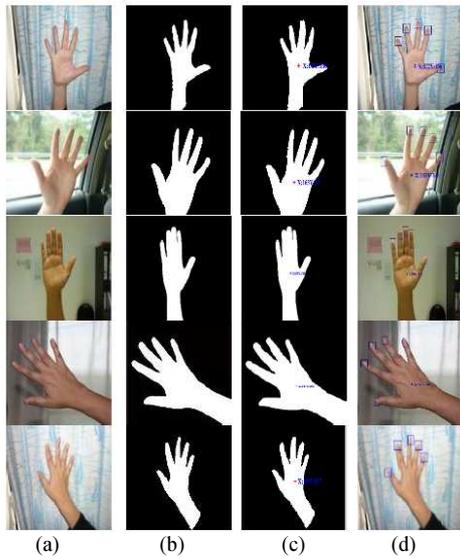


Fig. 8. (a) Original image (b) segmented image (c) centroid detection (d) Fingertips detection in the original image

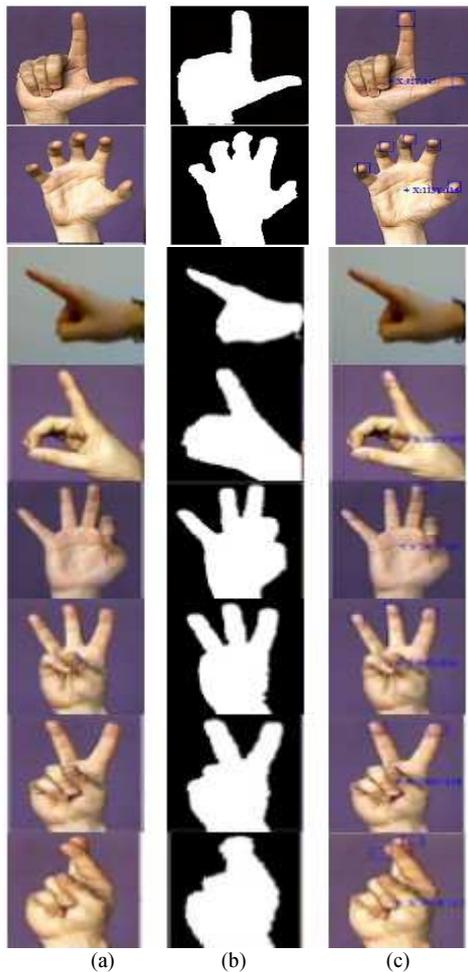


Fig. 9.(a) Original image (b) segmented image (c) Fingertips detection in the original image

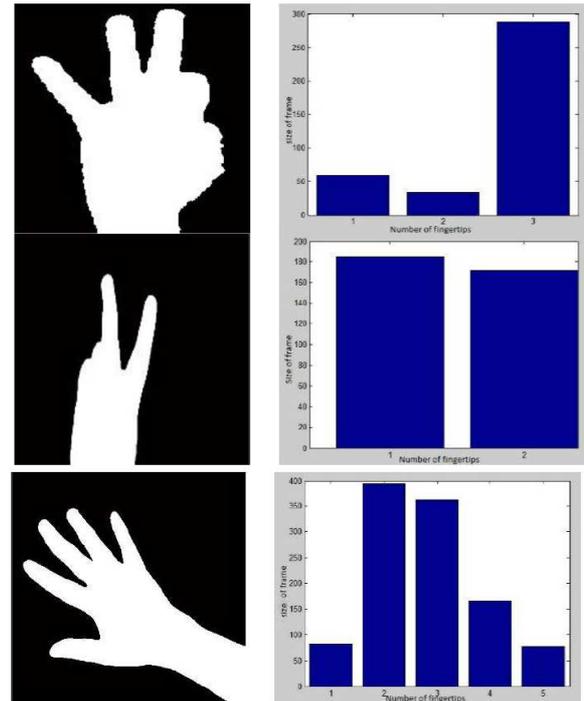


Fig. 10. Result of image binarization with the corresponding histogram for different gestures

Once the fingers' end points are obtained, a calculation of the distances between them is made. The particular sign is recognized through the use of a rule-based system which involves the number of fingertips, the number of branches and the distance between the neighbouring fingertips, after which, the limits for each sign are specified. The proposed method's efficiency is then calculated for three different backgrounds, as shown in Table 1. The processing time is recorded afterwards. Each of these calculations contains all the American Sign Language's signs for alphabets from A to Z. A correct result (CN) is obtained whenever a sign is correctly recognized while an Incorrect Result (IN) is obtained whenever the sign is unidentifiable or recognized as incorrect. The computation of efficiency uses the following formula. Alphabets J and Z are considered as dynamic gestures; thus, the researchers can only capture the final position. This approach's efficiency is illustrated in "Equation 9". The different background for ASL is shown in Fig. 11:

$$\text{Efficiency} = \frac{CN}{(IN + CN)} \times 100\% = 72 / (6 + 72) \times 100\% = 92\% \quad (9)$$

The confusion matrix for binary classification has four categories, where T_p stands for true positive, F_N stands for false negative, F_p stands for true positive and T_N stands for true negative counts. One plot the False Positive Rate (FPR) on the x-axis and the True Positive Rate (TPR) on the y-axis. Therefore, one plot Recall on the x-axis and

Precision on the y-axis, The Recall indicates the True Positive Rate, whereas the Precision indicates the Positive Predictive. The recall is the same as TPR. Accuracy (ACC) is calculated as shown in “Equation 10 to 12”:

$$\text{Recall } TPR = \frac{T_p}{T_p + F_N} \quad (10)$$

$$\text{Precision } PPV = \frac{T_p}{T_p + F_p} \quad (11)$$

$$ACC = \frac{T_p + T_N}{T_N + F_p + T_p + F_N} \quad (12)$$

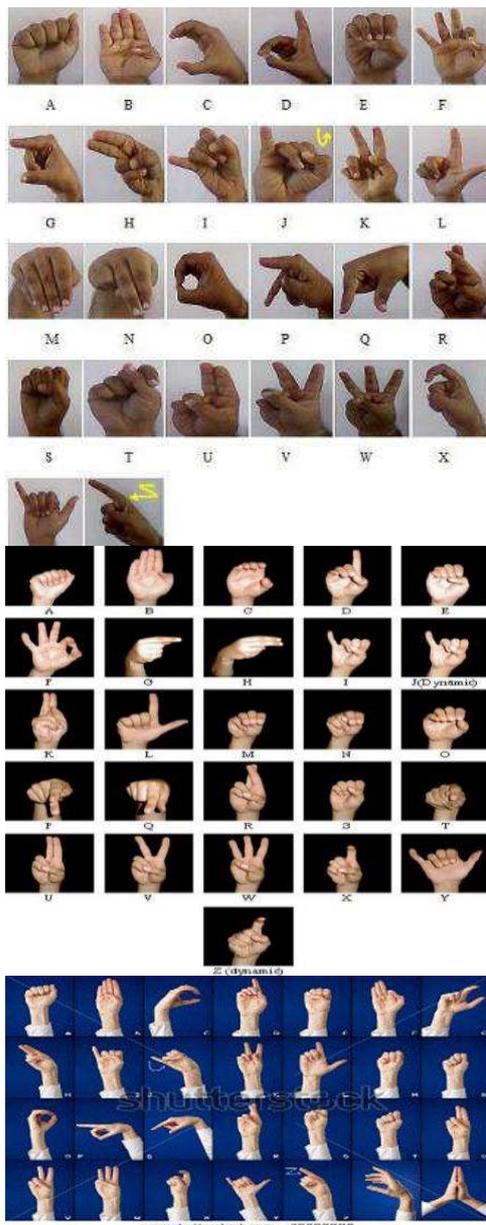


Fig. 11. ASL alphabets in different backgrounds

Table 2 shows the comparison of our proposed method with existing works. As shown in the table, we managed to get good accuracy, precision and recall which is 96.3, 92.46 and 97.33% compared to other methods.

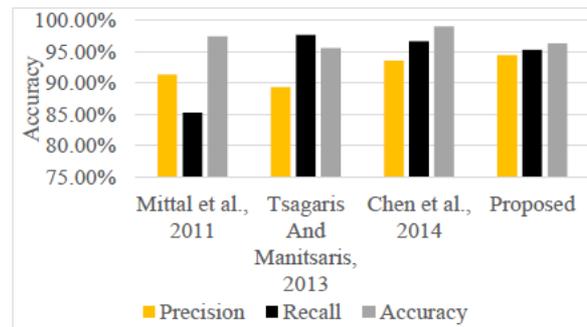


Fig.12. The result of Average detection performance (Accuracy, Recall and Precision) for four methods

Table 1. Process time for each gesture

Sign	Three different background White, Black, Blue			Detected fingertips	Time process (mm)
	CN	IN			
A	2	1	0	0	266,278,288
B	3	0	4	4	666,668,678
C	3	0	5	5	704,702,706
D	3	0	1	1	327,325,330
E	2	1	0	0	203,201,207
F	3	0	3	3	510,504,502
G	3	0	2	2	430,427,429
H	3	0	2	2	477,479,474
I	3	0	1	1	328,327,330
J	3	0	1	1	329,331,332
K	3	0	3	3	504,502,508
L	3	0	2	2	480,484,487
M	2	1	0	0	202,206,207
N	2	1	0	0	285,281,287
O	2	1	0	0	281,280,284
P	3	0	3	3	530,532,535
Q	3	0	2	2	402,405,409
R	3	0	2	2	479,482,483
S	2	1	0	0	277,275,279
T	3	0	1	1	336,334,339
U	3	0	2	2	470,474,477
V	3	0	2	2	488,490,493
W	3	0	3	3	508,510,512
X	3	0	1	1	333,337,339
Y	3	0	2	2	405,408,411
Z	3	0	1	1	364,367,369

Table 2. Method comparison

Method	Precision (%)	Recall (%)	Accuracy (%)
Mittal <i>et al.</i> (2011)	91.37	85.30	97.44
Tsagaris and Manitsaris (2013)	97.72	95.56	89.41
Chen <i>et al.</i> (2014)	93.56	96.69	99.00
Proposed	92.46	97.33	96.30

Conclusion and Future Work

This research paper presented a new method that detects the fingertips positions based on skeletonization as well as a few morphological operations. This proposed method is quite effective in recognizing static hand gestures. This work can be further applied for the purpose of sign language recognition as well as for the HCI domain.

Towards getting highly accurate hand detection, this work employed the free-form skin colour model, as described in (Dawod *et al.*, 2010a), along with Morphological filtering. Based on the conducted experiments, this proposed method succeeded in detecting all the open fingertips from images that have different complex backgrounds, as well as those images from the ASL alphabet dataset which performs various gestures.

As a future work, the method that we used will undergo improvement towards recognizing dynamic hand gestures and closed fingertips. Moreover, a possible extension of this work can be applied to the performance of static gestures recognition.

Acknowledgement

I wanted to take this opportunity to thank Dr. Md. Jan and Dr. Junaidi for their constant support and guide me for doing very good research work also, they provided to me a chance to be the corresponding author. Indeed, it is a big responsible to me, therefore we are a very effective team for the research.

Funding Information

The researchers wish to thank Universiti Kebangsaan Malaysia (UKM) and Ministry of Education, Malaysia for supporting this work by research grant DIP-2014-018 and FRGS/1/2014/ICT07/UKM/02/2.

Author's Contributions

Ahmad Yahya Dawod: Design the research work and prepare the workflow. Undertake the required experiments and analysed the obtained results.

Md Jan Nordin: Research adviser and Review the article and proposed algorithm.

Junaidi Abdullah: Development of the approach coordinator and evaluation of the obtained results. .

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript. Authors confirm that there is no conflict of interest.

References

- Al Tairi, Z.H. and M.I. Saripan, 2014. Skin segmentation using YUV and RGB color spaces. *J. Inform. Process. Syst.*, 10: 283-299.
DOI: 10.3745/JIPS.02.0002
- Ayala-Ramirez, V., S.A. Mota-Gutierrez, U.H. Hernandez-Belmonte, R.E. Sanchez-Yanez and C.S.V. Km *et al.*, 2011. A hand gesture recognition system based on geometric features and color information for human computer interaction tasks. *J. Patt. Recognit. Society*, 1: 54-59.
- Bhuyan, M.K., K.F. MacDorman, M.K. Kar, D.R. Neog and B.C. Lovell *et al.*, 2015. Hand pose recognition from monocular images by geometrical and texture analysis. *J. Visual Langu. Comput.*, 28: 39-55.
DOI: 10.1016/j.jvlc.2014.12.001.
- Bhuyan, M.K., D.R. Neog and M.K. Kar, 2012. Fingertip detection for hand pose recognition. *Int. J. Comput. Sci. Eng.*, 4: 501-511.
- Chen, Z.H., J.T. Kim, J. Liang, J. Zhang and Y.B. Yuan, 2014. Real-time hand gesture recognition using finger segmentation. *Sci. World J.*
DOI: 10.1155/2014/267872
- Cooper, H., E.J. Ong, N. Pugeault and R. Bowden, 2012. Sign language recognition using sub-units. *J. Machine Learning Res.*, 13: 2205-2231.
- Dawod, A.Y., J. Abdullah and M.J. Alam, 2010. A new method for hand segmentation using free-form skin color model. *Proceedings of the 3rd International Conference on Advanced Computer Theory and Engineering*, Aug. 20-22, IEEE Xplore Press, Chengdu, pp: V2-562-V2-566.
DOI: 10.1109/ICACTE.2010.5579466
- Dawod, A.Y., J. Abdullah and M.J. Alam, 2010. Fingertips detection from color image with complex background.
- Dung, L. and M. Mizukawa, 2009. Fast hand feature extraction based on connected component labeling, distance transform and Hough transform. *J. Robotics Mechatronics*, 21: 726-729.
- Hasan, H. and S. Abdul-Kareem, 2014. Static hand gesture recognition using neural networks. *Artificial Intellig. Rev.*, 41: 147-181.
DOI: 10.1007/s10462-011-9303-1
- Hasan, M.M. and P.K. Mishra, 2012. Real time fingers and palm locating using dynamic circle templates. *Int. J. Comput. Applic.*, 41: 33-43.
DOI: 10.5120/5547-7615
- Ibraheem, N.A., R.Z. Khan and M.M. Hasan, 2013. Comparative study of skin color based segmentation techniques. *Int. J. Applied Inform. Syst.*, 5: 24-38.
- Joshi, G. and R. Vig, 2015. Histograms of orientation gradient investigation for static hand gestures. *Proceedings of the International Conference on Computing, Communication and Automation*, May 15-16, IEEE Xplore Press, Noida, pp: 1100-1103.
DOI: 10.1109/CCA.2015.7148539

- Kaur, A. and B. Kranthi, 2012. Comparison between YCbCr color space and CIELab color space for skin color segmentation. *Int. J. Account. Inform. Syst.*, 3: 30-33.
- Khan, A. and S. Ravi, 2013. Image segmentation methods: A comparative study. *Int. J. Soft Comput. Eng.*, 3: 84-92.
- Mittal, A., A. Zisserman and P.H. Torr, 2011. Hand detection using multiple proposals. *Proceedings of the British Machine Vision Conference, (MVC' 11)*, BMVA Press, pp: 1-11. DOI: 10.5244/C.25.75
- Nguyen, T.N., H.H. Huynh and J. Meunier, 2013. Static hand gesture recognition using artificial neural network. *J. Image Graphics*, 1: 34-38.
- Omkar, S. and M. Monisha, 2011. Sign language recognition using thinning algorithm. *Ictact J. Image Video Process.*, 2: 241-245.
- Qiu-yu, Z., L. Jun-chi, Z. Mo-yi, D. Hong-xiang and L. Lu, 2015. Hand gesture segmentation method based on YCbCr color space and k-means clustering. *Interaction*, 8: 106-116.
- Raheja, J.L., K. Das and A. Chaudhary, 2012. Fingertip detection: A fast method with natural hand. *Int. J. Embedded Syst. Comput. Eng.*, 0134: 85-88.
- Rautaray, S.S. and A. Agrawal, 2015. Vision based hand gesture recognition for human computer interaction: A survey. *Artificial Intellig. Rev.*, 43: 1-54. DOI: 10.1007/s10462-012-9356-9
- Shakti, S., 2013. Comparative study of various image segmentation methods. *Int. J. Multidisciplinary Acad. Res.*, 2: 1-12.
- Shen, X., G. Hua, L. Williams and Y. Wu, 2012. Dynamic hand gesture recognition: An exemplar-based approach from motion divergence fields. *Image Vision Comput.*, 30: 227-235. DOI: 10.1016/j.imavis.2011.11.003
- Tsagaris, A. and S. Manitsaris, 2013. Colour space comparison for skin detection in finger gesture recognition. *Int. J. Adv. Eng. Technol.*, 6: 1431-1441.
- Vieriu, R.L., I. Mironica and B.T. Goras, 2013. Background invariant static hand gesture recognition based on Hidden Markov models. *Proceedings of the International Symposium on Signals, Circuits and Systems*, Jul. 11-12, IEEE Xplore Press, Iasi, pp: 1-4. DOI: 10.1109/ISSCS.2013.6651245
- Wenkai, X. and J.E. Lee, 2012. Continuous gesture trajectory recognition system based on computer vision. *Applied Math, Inform. Sci.*, 6: 339S-346S.