

Arabic Alphabet and Numbers Sign Language Recognition

Mahmoud Zaki Abdo

Electronic, communication, and computer department
Faculty of engineering, Helwan University
Cairo, Egypt

Sameh Abd El-Rahman Salem

Electronic, communication, and computer department
Faculty of engineering, Helwan University
Cairo, Egypt

Alaa Mahmoud Hamdy

Electronic, communication, and computer department
Faculty of engineering, Helwan University
Cairo, Egypt

Elsayed Mostafa Saad

Electronic, communication, and computer department
Faculty of engineering, Helwan University
Cairo, Egypt

Abstract—This paper introduces an Arabic Alphabet and Numbers Sign Language Recognition (ArANSLR). It facilitates the communication between the deaf and normal people by recognizing the alphabet and numbers signs of Arabic sign language to text or speech. To achieve this target, the system able to visually recognize gestures from hand image input. The proposed algorithm uses hand geometry and the different shape of a hand in each sign for classifying letters shape by using Hidden Markov Model (HMM). Experiments on real-world datasets showed that the proposed algorithm for Arabic alphabet and numbers sign language recognition is suitability and reliability compared with other competitive algorithms. The experiment results show that the increasing of the gesture recognition rate depends on the increasing of the number of zones by dividing the rectangle surrounding the hand.

Keywords—hand gestures; hand geometry; Sign language recognition; image analysis; and HMM

I. INTRODUCTION

The main problem in the language of deaf people makes it difficult to translate thoughts and feelings into words and phrases understandable and aware. The normal people translate ideas into words audible, but the deaf people translate ideas into visual signs through the hands movement.

Over the years used the deaf and dumb signs among themselves. It became the different sign language of each community in the world. These signs are only one thing for the deaf and dumb to communicate with each other and the outside world to them [1]. There has been a growing interest in the recognition of human hand movements. Normally, there is no problem when deaf persons communicate with each other by using their common sign language. The problem appears when a Deaf people want to communicate with a non-deaf people. Usually, both will be disgruntled in a very short time [2].

A Sign language is different from country to another country; the researchers attempt to unify the sign language in each country separately have been carried out such as Jordan, Saudi Arabia, and Egypt to help persons of the deaf and dumb for each country. The researchers are working on hand gestures in different sign languages as the Australian Sign Language

(Auslan) [3], the Chinese Sign Language (CSL) [4], the American Sign Language (ASL) [5], and the Dutch Sign Language. The Arabic Sign Language has not received attention in researchers [7].

All Previous researches on sign languages depend on glove or vision based methods [6]. In the glove based method, the user wears special devices, like special gloves or markers, the system related with data on the hand shape and motion. While in the vision method, the system recognizes the gestures by using image processing techniques without putting any limitation on the user [7].

The work [2] created an automatic translation system for gestures of manual alphabets in the Arabic sign language recognition. It does not rely on using any visual markings or gloves. The extracted features depend on two stages, feature-vector-creation stage and edge detection stage. It used multilayer perceptron (MLP) classifier and minimum distance classifier (MDC) to detect 15 characters only of 28 characters.

The research work in [7] investigated appearance-based features for the deaf person- vision-based on sign language recognition. It does not depend on a segmentation of the input images, and he used the image as a feature. The system used a combination of features including PCA, hand trajectory, hand position, and hand velocity. The grey scale image with a reduced frame size 195x165 pixels and downscaling to 32x32 pixels used **rwth-boston-104** database.

A system of the recognition and translation of the numbers was designed [8]. The system is consist of four main phases; pre-processing phase, feature extraction phase, interpolation phase and the classification phase. The extracted features are scale invariant and make the system more flexible. The experimental result revealed that the system was able to recognize a representing numbers from one to nine based on the minimum Euclidean distance between the numbers.

The research work in [9] introduced new two features for American Sign Language recognition: those are kurtosis position and principal component analysis PCA. Principal component analysis is used as a descriptor to provide a

measure for hand orientation and hand configuration. PCA has been used before in sign language as a dimensionality reduction. As a local feature for measuring edges and reflecting the position, Kurtosis position is used. It used motion chain code that represents the movement of hand as a feature. The system input is a sign from RWTH-BOSTON-50 database, and the recognition error rate of the output is 10.90%.

A system for the recognition and translation of the Arabic letters was designed [10]. The system depends on the inner circle position on the hand contour and divides the rectangle surrounding by the hand shape into 16 zones. The extracted features are scale invariant. Experiments revealed that the system was able to recognize Arabic letters based on the hand geometry. The experiment results shown that the different signs gesture recognition rate of Arabic alphabet for were 81.6 %.

The research work in [11] used Adaptive Neuro-Fuzzy Inference system (ANFIS). The system used 30 Arabic sign language alphabets visually. The recognition rate of the system was 93.55%. The research work in [12] built an ArSL system and measures the performance of ArSL data collected. The system based on Polynomial Classifiers. It collected a 30 letter of ArSL. It collected the data by using gloves marked with six different colours at different regions as shown in Fig. 1 [12]. The recognition rate is 93.41 %

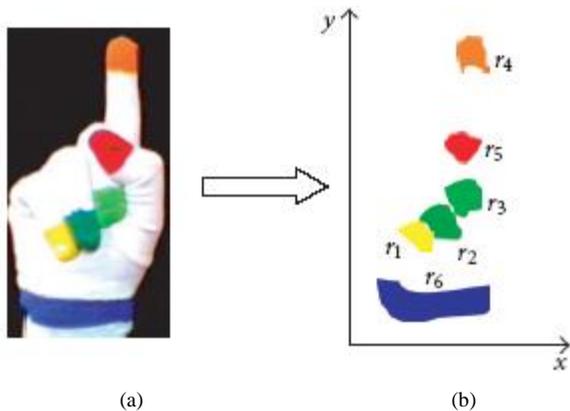


Fig. 1. (a) Colored gloves and (b) output image segmentation

This paper is organized as follows. Section two explains HMM classifier. Section three presents the proposed system. Section four shows the experimental data. Section five explains the experimental results. Section six presents the conclusions.

II. HMM CLASSIFIER

HMM is used as a classifier for speech [13] and used in sign language recognition systems. In HMM-based approaches, the information of each sign is modelled by a different HMM. The model that gives the highest likelihood is selected as the best model and the test sign is classified as the sign of that model [14]. It consists of a set of N states where the transition from each state to another state. It is denoted by Eq. 1:

$$\lambda = (A, B, \pi) \quad (1)$$

- The state transition probability distribution $A = \{a_{ij}\}$ where its elements represent the transition probability

from each state to another state. State transition coefficients having the properties Eq. 2 and Eq. 3.

$$a_{ij} \geq 0 \quad (2)$$

$$i \geq 1, j \leq N$$

$$\sum_{j=1}^N a_{ij} = 1 \quad (3)$$

- The observation symbol probability distribution in state j, $B = \{b_j(k)\}$ where its elements represent the probability of certain observation to occur at a particular state $\{1 \leq j \leq N, 1 \leq k \leq M\}$, where M is a number of observation sequence $O_1 O_2 \dots O_M$
- The initial state distribution $\pi = \{\pi_i\}, 1 \leq i \leq N$

III. PROPOSED SYSTEM

The proposed system, as shown in Fig. 2, consists of five phases, skin detection, removing background, face and hands isolating, Observation detection, and Hidden Markov Model HMM classifier. The Maximum recognition probability $W_i, i = 1 \dots N$, where N is a number of letters, is corresponding to letter detection.

The system components described in the following subsections: Sub section (A) presents skin detection and removing background. Sub section (B) presents face and hands isolating. Sub section (C) presents observation detection and HMM.

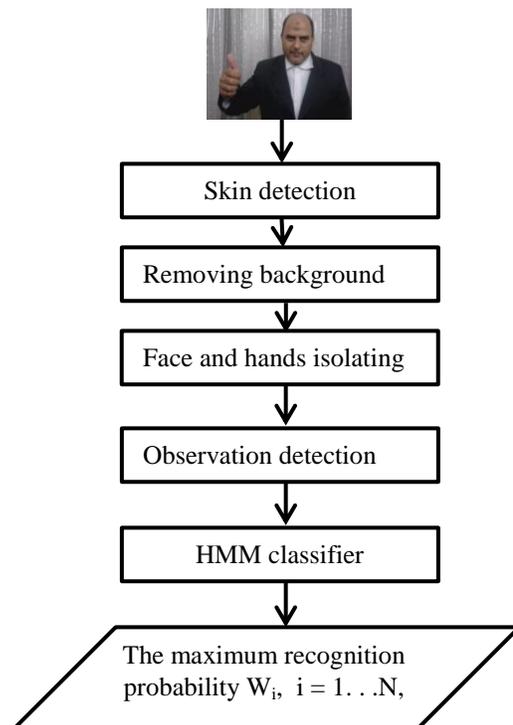


Fig. 2. Proposed system architecture

A. Skin Detection and Background Removal

The algorithm uses skin detection [15]. The algorithm adopts skin colour detection as the first step. Due to YCbCr color space transform, YCbCr is faster than other approaches

[16, 17]. The algorithm calculates the average luminance Y_{avg} of the input image as given in Eq.4. $Y_{avg} = \sum y_{i,j}$ (4)

Where $y_{i,j} = 0.3 R + 0.6 G + 0.1 B$ is normalized to the range {0 to 255}, where i, j are the indices of the pixel in the image.

According to Y_{avg} , the algorithm can calculate the compensated image $C_{i,j}$ by the following equations Eq.5 and Eq.6 [15]:

$$\begin{aligned} R'_{i,j} &= (R_{i,j})^\tau \\ G'_{i,j} &= (G_{i,j})^\tau \end{aligned} \quad (5)$$

$$C_{i,j} = \{R'_{i,j}, G'_{i,j}, B_{i,j}\}$$

Where $\tau =$

$$\begin{cases} 1.4, & Y_{avg} < 64 \\ 0.6, & Y_{avg} > 192 \\ 1, & \text{otherwise.} \end{cases} \quad (6)$$

It should be noted that the algorithm compensates the colour of R and G to reduce computation. Due to chrominance (Cr) which can well represent human skin, the algorithm only consider Cr factor for colour space transform to reduce the computation. Cr is defined as follows Eq. 7 [17]: $Cr=0.5R'-0.419G'-0.081B$ (7)

Accordingly, the human skin binary matrix can be obtained as follows:

$$S_{ij} = \begin{cases} 0, & 10 < Cr < 45 \\ 1, & \text{otherwise} \end{cases} \quad (8)$$

Where '0' is the white point and '1' is the black point. The algorithm implements a filtration by a 5×5 mask.

First, the algorithm segments S_{ij} into 5×5 blocks, and calculate show many white points in a block. Then, every point of a 5×5 block is set to white point when the number of white points is greater than half the number of total points. Otherwise, if the number of black points is more than a half, this 5×5 block is modified to a complete black block, as shown in Fig. 3 [16].

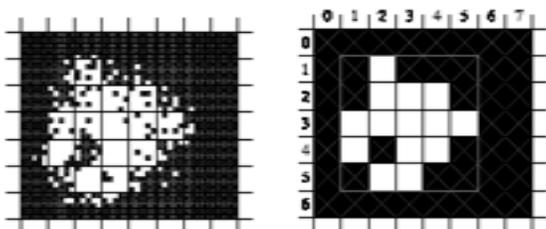


Fig. 3. (a) An example of S_{ij} (b) Noise removal by the 5×5 filter

Fig. 4 shows the resultant image shapes after skin detection and removing the background [18] of image.

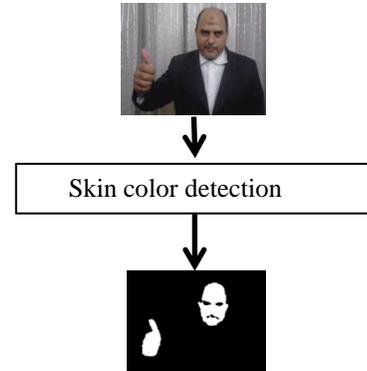


Fig. 4. Skin colour detection and removing background

B. Face and Hand Isolating

The algorithm tracks the objects in each image. The algorithm neglected the small objects, and then detects the largest objects as hands and the face. The algorithm isolates the hand and face as in Fig. 4. After detecting the skin colour and removing background the position of the face and hands can be isolate and detected as Fig.5.

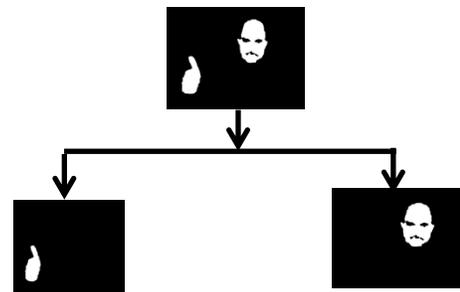


Fig. 5. Isolating the face and hands

Figure 4 shows the detected skin with background removal. The image contains a right hand and a face. The algorithm detects the hand and a face by the position and shape of each. Fig.5 shows isolating the face and hands, then isolate the right hand to detect the letter.

C. Observation Detection and HMM

The proposed algorithm divides the rectangle surrounding by the hand shape in Fig. 5 into 16 zones as in Fig. 6.



Fig. 6. Hand contour detecting

Col 4	Col 3	Col 2	Col 1	
4	3	2	1	Row 1
8	7	6	5	Row 2
12	11	10	9	Row 3
16	15	14	13	Row 4

Fig. 7. 16 Zones hand shape

Fig. 8 summarize a creation of observation sequence and compute the HMM $\lambda = (A, B, \pi)$ for each letter and test a letter to get maximum $P(O|\lambda)$.

<p>Step 1: Divide the rectangle surrounding by the hand into 16 zones as Fig. 6 and Fig. 7.</p> <p>Step 2: Count the number of white pixel in each zone.</p> <p>Step 3: Sort the zone numbers in ascending order depends into a number of white pixels in each zone.</p> <p>Step 4: The observation vector of the letter is a vector of a sorted zone numbers, numbers between 1 and 16.</p> <p>Step 5: Train the HMM for each letter $\lambda = (A, B, \pi)$ to maximize $P(O \lambda)$</p> <p>Step 6: To test a letter: Given the observation sequence $O=O_1O_2 \dots O_{16}$ and a model $\lambda = (A, B, \pi)$ for this letter, then compute $P(O \lambda)$ for each letter. The target letter is the maximum $P(O \lambda)$.</p>

Fig. 8. The proposed algorithm of calculating the observation vector and using HMM to train and test the letters

IV. EXPERIMENTAL DATA

To tune and test the proposed system, Arabic Alphabet Sign Language Recognition ArASLR database (ArASLRDB) is generated as follows.

The ArASLRDB corpus consists of 29 alphabet Arabic letters and numbers from 0 to nine as shown table 1. Clothes of signers are differently and the brightness of their clothes is completely different from person to other.

The image of the database is 640 x 480 pixels saved in jpg file format. The implementation is using the following as table 1:

- Number of signs: 38.
- Number of images: 357.
- Number of training images: 253.
- Number of testing images: 104.
- Average images per sign: 9.4.
- Average training images per sign: 6.7.
- An average test image per sign: 2.7.
- Percentage of training images per sign: 70.87%.
- Average testing images per sign: 29.13%.

The programme is implemented using a Windows based MATLAB (R2013a).

TABLE I. ARABIC ALPHABET DATABASE

	Arabic Sign name	English Sign Name	Number of images		Arabic Sign name	English Sign Name	Number of images
1	أ	Alef	11	20	ف	fa3	7
2	ب	Ba3	8	21	ق	Kaaf	11
3	ت	Ta3	10	22	ك	Kaf	9
4	ث	Tha3	7	23	ل	Laam	11
5	ج	Geem	12	24	م	Meem	10
6	ح	Ha3	10	25	ن	Noon	14
7	خ	Kha3	7	26	هـ	ha3	6
8	د	Dal	8	27	و	Waw	10
9	ذ	Thal	10	28	ى	Ya3	11
10	ر	Ra3	8	29	لا	Laa	10
11	ز	Zay	10	30	1	1	11
12	س	Seen	10	31	2	2	10
13	ش	Sheen	6	32	3	3	10
14	ص	Sad	8	33	4	4	6
15	ض	Dad	10	34	5	5	14
16	ط	Da3	9	35	6	6	9
17	ظ	Thaa3	8	36	7	7	11
18	ع	Aien	6	37	8	8	8
19	غ	Ghain	10	38	9	9	11

V. EXPERIMENTAL RESULT

In this research, the HMM is applied on the ArASLRDB with 29 Arabic alphabet sign language. The recognition system is tested when dividing the rectangle surrounding by the hand shape in Fig. 5 into 4, 9, 16, and 25 zones.

- At 16 zones: the recognition rate changes with changing the number of states until arrive to 100 % at 19 states, as shown in Fig.9.
- At 4 and 9 zones: the recognition rate cannot arrive to 100%, as shown in Fig. 9.

- At 4 zones: the recognition rate is very poor and cannot exceed 40%.
- At 9 zones: the recognition rate cannot exceed 97% whatever increases the number of states.
- At 25 zones: the recognition rate changes with changing the number of states until arriving 100 % at 18 states, as shown in Fig.9.
- The average time to execute the proposed algorithm to arrive of 100% recognition rate is shown in table 2.

TABLE II. RESULT OF THE PROPOSED SYSTEM FOR DIFFERENT NUMBER OF ZONES

Zone number	State number	Time (Sec)	Recognition Ratio
4 zones	30	0.0332	37.5 %
9 zones	30	0.0708	96.93 %
16 zones	19	0.0902	100 %
25 zones	18	0.1065	100 %
36 zones	13	0.1358	100 %
49 zones	12	0.1531	100 %
64 zones	11	0.1872	100 %

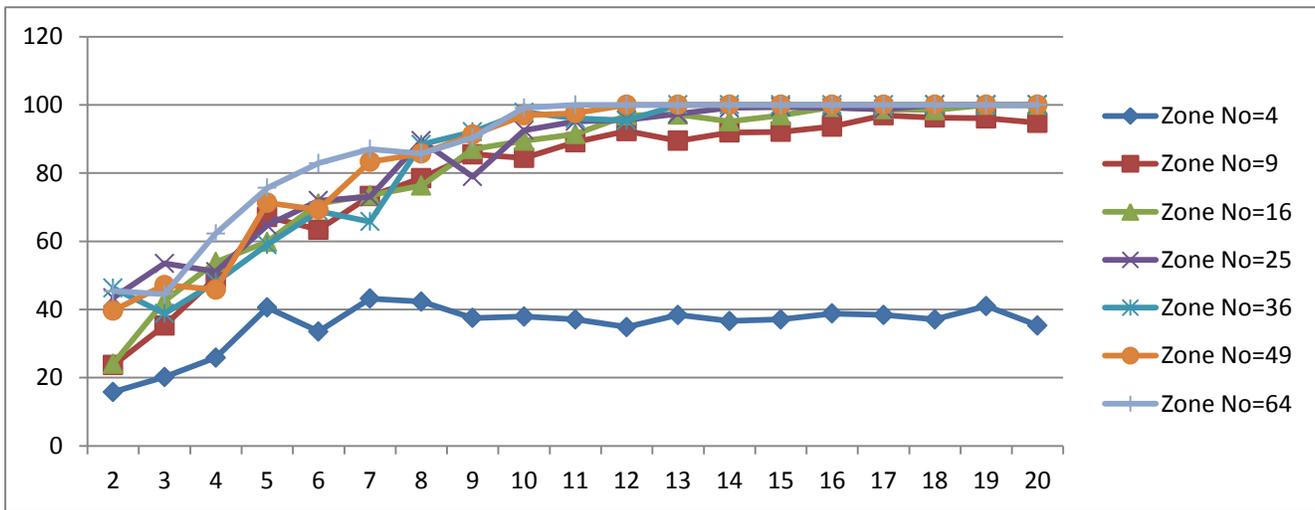


Fig. 9. The recognition rate verse us the change of the number of states for 4, 9, 16, and 25 zones

Finally, the best number of zones=16 with 19 states to recognition Arabic alphabet of sign language. The algorithm can achieve to 100% recognition rate with increase the zone number more than or equal 16 zones but more time is required.

As shown in table 3, Reference [2] used minimum distance classifier (MDC) and also used multilayer perceptron (MLP) classifier to detect 15 characters only of letters with recognition rate 91.7 % and 83.7 % respectively. Reference [11]

recognized Arabic letters based on the hand geometry and the recognition rate of Arabic alphabet for different signs was 81.6 %. This system can reach a 100 % recognition rate with increasing number of zones and number of states. Reference [13] used Gloves marked with six different colour, the system used polynomial classifiers to recognize 30 letters with recognition rate of 93.41 %. Reference [12] did not use gloves and used ANFIS to recognize 30 letters by recognition rate of 93.55 %.

TABLE III. COMPARISON WITH ARSL ALPHABET RECOGNITION

	Instruments used	Number of Letters	Classifier	Recognition Rate
ArASLR	None: Free Hands	29 Letters and 9 Numbers	HMM	100 % on more than 16 zones
EL-Bendary etl. [2]	None: Free Hands	15 Letter	MDC MLP	91.3% 83.7 %
Jarrah, etl [12]	None: Free Hands	30 Letters	ANFIS	93.55 %
Assaleh, etl [13]	Gloves marked with six different colour	30 Letters	polynomial classifiers	93.41%
ArSLAT [11]	None: Free Hands	29 Letter	Outer of the inner circle zones	83.16

VI. CONCLUSIONS

In this paper, a new feature is used to recognize the Arabic Alphabet sign language via HMM. The proposed system is demonstrated experimentally. The phases of the proposed algorithm consists of skin detection, background exclusion, face and hands extraction, feature extraction, and also classification using Hidden Markov Model (HMM). The proposed algorithm isolates the hand from the image to recognize the letter. The proposed algorithm divides the rectangle surrounding by the hand shape into zones. The best number of zones is 16 zones. The observation of HMM is created by sorting zone numbers in ascending order depending on the number of white pixels in each zone. Experimental results show that the proposed algorithm achieves 100% recognition rate with minimum execution time at 16 zones with 19 states.

REFERENCES

- [1] El-Bendary, Nashwa, et al. "ArSLAT: Arabic Sign Language Alphabets Translator." 2010 International Conference on Computer Information Systems and Industrial Management Applications (CISIM), 2010.
- [2] A. A.Youssif, Amal Elsayed Aboutabl, Heba Hamdy Ali,"Arabic Sign Language (ArSL) recognition system using HMM ", IJACSA, vol. 2, no. 11, 2011.
- [3] E.J. Holden, G. Lee and R. Owens, Automatic recognition of colloquial Australian sign language, IEEE Workshop on Motion and Video Computing 2, pp.183–188, Dec. 2005.
- [4] F. Chen, C. Fu and C., "Hand gesture recognition using a real time tracking method and hidden Markov models", Huang, Image and Vision Computing 21, pp. 745–758, Mar. 2003.
- [5] Starner, "Visual Recognition of American Sign Language Using Hidden Markov Models", Doctoral dissertation, Massachusetts Institute of Technology, 1995.
- [6] M. Al-Rousan, O. Al-Jarrah, and M. Al-Hammouri, "Recognition of dynamic gestures in Arabic sign language using two stages hierarchical scheme", The International Journal of Intelligent and Knowledge Based Engineering Systems, vol. 14, no. 3, 2010.
- [7] Rybach, D., "Appearance-based features for automatic continuous sign language recognition", Diploma Thesis, RWTH Aachen University, Germany, 2006.
- [8] Mahmoud Zaki Abdo, Alaa Mahmoud Hamdy, Sameh Abd El-rahman Salem, El-sayed Mostafa Saad. "An interpolation based technique for sign language recognition." Radio Science Conference (NRSC), 30th National. IEEE, 2013.
- [9] Zaki M. M., Shaheen S. I., "Sign language recognition using a combination of new vision based features, Pattern Recognition Letters", vol. 32, Issue 4, pp. 572-577, 2011.
- [10] Mahmoud Zaki Abdo, Alaa Mahmoud Hamdy, Sameh Abd El-rahman Salem, El-sayed Mostafa Saad," Arabic sign language recognition", International Journal of Computer Applications, pp. 0975 – 8887, vol. 89, no. 20, Mar. 2014.
- [11] Al-Jarrah, Omar, and Alaa Halawani. "Recognition of gestures in Arabic sign language using neuro-fuzzy systems." Artificial Intelligence 133.1-2, pp. 117-138, 2001.
- [12] Assaleh, Khaled, and M. Al-Rousan., "Recognition of Arabic sign language alphabet using polynomial classifiers", Journal on Applied Signal Processing , pp. 2136-2145, 2005
- [13] Rabiner, Lawrence R., and Biing-Hwang Juang. "An introduction to hidden Markov models." ASSP Magazine, IEEE 3.1, pp. 4-16, 1986.
- [14] Oya Aran, "Vision based sign language recognition: modelling and recognizing isolated signs with manual and non-manual components", 2008, Phd thesis.
- [15] Pai, Y. T., Ruan, S. J., Shie, M. C., & Liu, Y. C. A simple and accurate color face detection algorithm in complex background. In Multimedia and Expo, IEEE International Conference on IEEE, pp. 1545-1548, 2006.
- [16] S. Gundimada, Li Tao, and v. Asari, "Face detection technique based on intensity and skin color distribution", in 2004 International Conference on Image Processing, vol. 2, pp. 1413–1416, Oct. 2004.
- [17] K. P. Seng, A. Suwandy, and L.-M. Ang, "Improved automatic face detection technique in color images", TENCON 2004, vol. 1, pp. 459–462, Nov. 2004.
- [18] Khaled, H., Sayed, S. G., Saad, E. S. M., & Ali, H. "Hand gesture recognition using modified 1\$ and background subtraction algorithms". Mathematical Problems in Engineering, 2015.