Propose a New Method for Extracting Hand using in the Arabic Sign Language Recognition (Arslr) System

Abdelmoty M.Ahmed[1]
Department of Systems and Computer Engineering
Faculty of Engineering /Al-Azhar University
Cairo, Egypt

Reda Abo Alez [2]
Department of Systems and Computer Engineering
Faculty of Engineering /Al-Azhar University
Cairo, Egypt

Muhammad Taha [3]
Department of Mathematics
Faculty of Science /Al-Azhar University
Cairo, Egypt

Gamal Tharwat [4]
Department of Systems and Computer Engineering
Faculty of Engineering /Al-Azhar University
Cairo, Egypt

Abstract—**The pattern recognition is considered of significant systems where applied computer vision techniques, and most important of these systems recognize the hand gestures which are used in the interpretation and translation of Arabic Sign Language(ArSL) for the deaf and dumb to written text software systems.**

Arabic Sign Language Recognition (ArSLR) is used as a popular tool of communication between the hearing disabled and the deaf people throughout the Arab world.

This paper proposes an image based ArSLR system using gesture recognition techniques, which allows the user to interact with the outside world. In particular, the design and development of a system for automatic translation of Arabic sign language to text in Arabic is proposed in this paper. And that includes extraction of the stem Arabic words, distinguish and different meanings of similar words.

The proposed system start with acquiring the image of alphabet letters gestures indicative Secondly to detect the hand of the image and isolate it from the background, the third stage is extraction and conclusion features and characteristics of hand sign pattern, then Classification process through one of the most effective classification algorithms to classify the form of hand gesture in the fourth step .Finally is the interpretation and translation of the letter in the alphabet indicative (sign language) contrast to his character in the alphabet of the Arabic language (Interpretation and Translation).

In this paper we will provide the results of the first stage of a system image capture then offer processing preliminary results of each image taken with the aim detect the hand for each, according to in the second stage of the system, and will be using the results of this paper in the study of other stages of the system and displayed in a separate research papers after the completion of their work.

Keywords—*ArSL, ArSLR, Features Extraction; Pattern recognition; sign language; Hand Gestures; Hand Detection; Classification.*

I. INTRODUCTION

Sign language is the prominent means of communication among the deaf and the hearing-disabled. Two approaches are mainly followed in sign language; they are vision-based and sensor-based. The advantage of vision-based systems over the counterpart is that users do not need to use complex equipments but in the preprocessing stage requires sufficient computations in place of cameras which are used in vision based systems , sensor-based systems use sensor enabled instrumented gloves this paper presents computer vision-based gesture interface that is part of a sign language recognition system, and also explains computerized sign language recognition system for the vocally disabled (deaf and dumb) that uses sign language for communication.

Generally there are three levels of image based ArSLR, they are continuous recognition, alphabet recognition and isolated-word recognition. The input of the vision based methods is a set of images or video sequence of the signs. The signers are asked to have pause between the signs to isolate the signs which is done manually This paper will present research progress and findings on techniques and algorithms for hand detection as it will be used as an input for gesture recognition process.

The translation of ArSL to Arabic text language system using image and pattern recognition technology is presented in Section I. Previous works and literature survey details are described in section II.The relationship of sign language to deaf and dumb with the knowledge of Pattern Recognition is discussed in Section III. Overview of the proposed system is discussed in Section IV. In Section V, the computer simulation results of first stage involved in the system are presented, In Section VI, the computer simulation results of second stage involved in the proposal system, finally future directions and conclusions are summarized in Section VII.

II. PREVIOUS WORKS AND LITERATURE SURVEY

A lot of research efforts has been carried on developing systems for sign languages from around the word that to concentrate mostly in vision based and Gloves sensor based areas.

In Glove-based recognition systems introduced by Sidney and Geoffrey [1] to recognize hand shape or hand gestures of sign language under different illumination changes. But it is necessary to use a glove based input device or motion capture system which restricts user motion and system mobility. Kim et al [2] proposed a gesture recognition system for the Korean sign language by using fuzzy-min-max neural network and data-glove. Lee et al [3] developed a new glove sensor and proposed its application to learn Korean finger spelling by using K-mean method and their developed glove-sensor. A finger spelling recognition method by using distinctive features of hand shape was proposed by Tabata et al [4]. An Arabic sign language translation system on mobile devices was introduced by Halawani [5]. To use input device of wearable computer, a new glove-based input device was proposed by Tsukada et al [6]. Statistical template matching was used to recognize Pakistan sign language based on data glove by Khalid Alvi et al [7]. Arabic sign language recognition system was developed using an instrumented glove proposed by AI-Buraiky et al [8]. x. Zabulisy et al. [9] proposed a vision-based hand gesture Recognition system for Human-Computer Interaction.

In vision-based approaches introduced to overcome these problems. Hamada et al [10] introduces a hand shape estimation approach to overcome occlusion by using multi-ocular images using two cameras. Rogerio Feris et al [11] proposed an approach to exploit depth discontinuities for finger spelling recognition to differentiate between similarities of some signs by using multi flash camera. The Hidden Markov Model (HMM) [12] and Dynamic Programming [13] were used to recognize American sign words. Salleh et al [14] provided a good idea to convert sign language to voice recognition based on feature extraction and HMM from grey scale images. On the other hand, Tanibata et al. [15] provided a prototype approach based on feature extraction to solve hand occlusion problem for Chinese sign language recognition. Mohandes [16], [17] introduced a prototype system to recognize the Arabic sign language based on Support Vector Machine (SVM) and also an automatic Translation system to translate Arabic Text to Arabic Sign Language. Foong et al [18] proposed A Sign to Voice system prototype which is capable of recognizing hand gestures by transforming digitized images of hand sign language to voice using Neural Network approach.

III. THE RELATIONSHIP OF SIGN LANGUAGE TO DEAF AND DUMB WITH THE KNOWLEDGE OF PATTERN RECOGNITION

Sign language is the term given to a means of communication is the voice used by people with special needs acoustically (deaf) or audio (dumb), despite the fact that there are other practices that could be classified as conversational indicative levels such as divers signals and some special signs I have some police forces or military or even between gangsters and other [19]

Sign language has become recognized as the language of global communication between people with special needs, but has become the creators of the Deaf ability to creativity poems and pieces of literature, and translation of the oral poetry to the language-dependent locomotors rhythm of the body, particularly the hands movement, considered to hand a great way to express your fingers and formations, can laugh and cry, rejoice and become angry, and express the desire, and we call edgy.

A. Communication systems for deaf people

There are several systems for communication between people with special needs of the deaf and dumb such as [20]

- Oral method: a deaf education and training without the use of sign language, spelling or fingers, do not use oral communication only reading and writing.

- Hand Gestures help to teach speech: It forms of moving the hands and aims to help teach deaf spoken language, representing put your hands on the mouth, nose or throat or chest, a way to express a particular character of verbal machine outlet.

- Read the lips: and rely attention and understand what the person is saying to monitor the movement of the lips, and exits of the letters of the mouth, tongue and throat, during a speech pronunciation.

- Hint language: It is a handy way to support spoken language, in which the speaker uses a set of hand movements carried out near the mouth with all the voices of speech and these tips provide the reader with the language of the lips and the information that describes what confused in reading this and make the hidden voices visible.

- Alphabetic indicative fingers or misspelling fingers: a technique to communicate rely representation alphabet and are often used in names. Or flags words that have no agreed-upon sign.

- Way pronunciation tuned: based on a set of principles that the most important speech is not limited to out vote abstract way but that speech comprehensive expression which interfere with body movements such as hand gestures ,facial features, rhythm , tone and signs.

- destruction Contact: This means that the use of effectively all possible means available to communicate and integrate all audio, manual, oral ,gestures, signs, movements of hands , fingers , lips, reading and writing systems to facilitate communication and facilitation.

B. Pattern Recognition

Pattern Recognition defined as a classification of input data via extraction important features from a lot of noisy data, Pattern Recognition (PR). It aims to extract information about the image to classify its contents. Inputs are in the form of digitized binary valued 2D images or textures containing the pattern to be classified [21] and Pattern Recognition science are associated closely linked to all the intelligent systems based decision-making.

Computer vision overlaps significantly with the image processing, and pattern recognition, most computer vision algorithms usually assumes a significant amount of image processing has taken place to improve image quality.

Pattern recognition (also called machine learning) studies various mathematical techniques (such as statistical techniques, neural network, support vector machine) to classify different patterns. The input data for pattern recognition can be any data. Pattern recognition techniques are widely used in computer vision. Many vision problems can be formulated as classification problem.

Most of Pattern Recognition Systems first start collects the data to be classified, then the analysis of these data, and then described the important features are extracted as information numerical symbolism (Analysis/Description), Then classification data entered by features derived from this data, according to one of Classification methods (Classification/Recognition), the figure 1 is describe A generic Pattern Recognition System (PRS) scheme.[22]
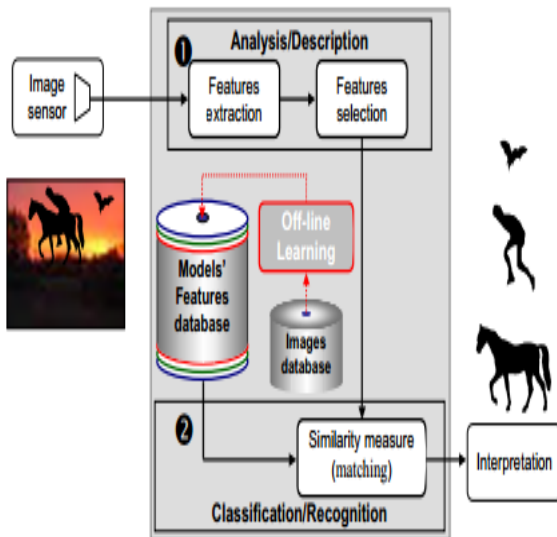

Fig. 1.   A generic pattern Recognition System (PRS) scheme.

There are many methods of classification of which statistical method, which relies on statistical features of the pattern of the entrance, structural pattern recognition method (Syntactic method) which depends on relations between the features to be recognized on the corresponding pattern of pattern entrance, Template matching method is used in image processing in how to identify shapes in an image, In this method one looks for parts in an image which match a template. Artificial neural network (ANN) method is a self-adaptive trainable process that is able to learn to resolve complex problems based on available knowledge. A set of available data is supplied to the system so that it finds the most adapted function among an allowed class of functions that matches the input.

For pattern recognition, various researches & various algorithms have been proposed; the types of pattern recognition are governed by its design cycle. As we know, it consists of basic elements like visual perception; feature extraction and classification, there are various different techniques and algorithms to implement these basic elements.

The Fig. 2 describe pattern Recognition Algorithm scheme. [23]
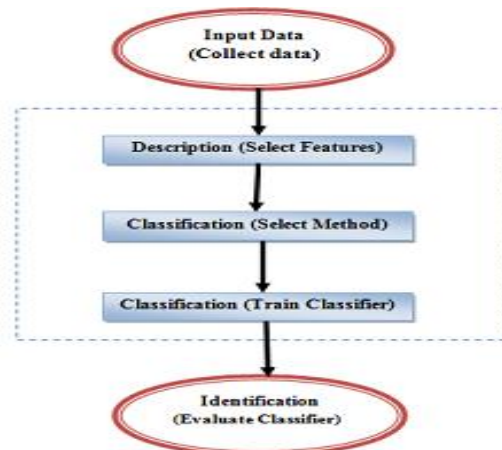

Fig. 2.   Pattern Recognition Algorithm scheme.

Fig.3 represents a systematic outline of the system to identify the image of a character indicative of the Arabic language and how to distinguish it According to Pattern Recognition algorithm .[24]
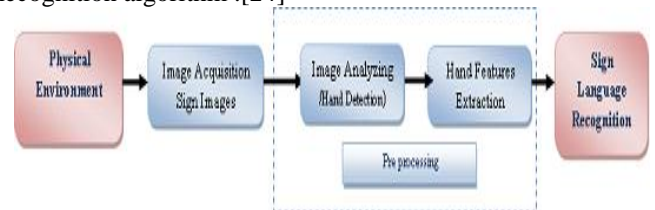

Fig. 3.   Hand Gesture Recognition scheme outlines

IV. OVERVIEW OF THE PROPOSED SYSTEM

In this section an overview of the proposed system is described for the automatic conversion of sign language to Arabic language text. The functional block diagram is given in Fig. 4.
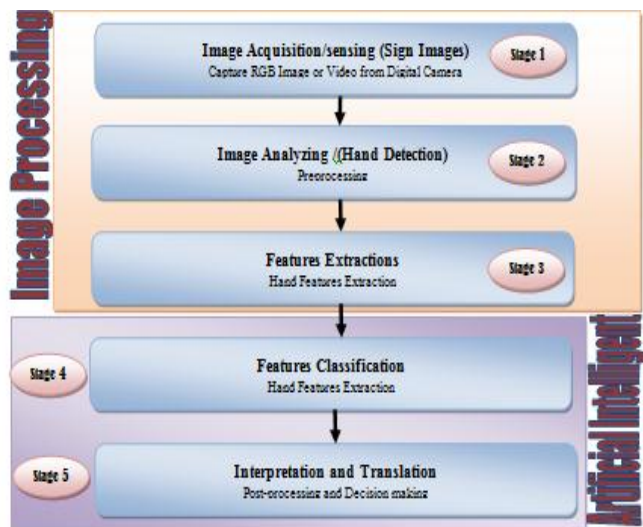

Fig. 4. the basic stages of the proposed ArSLR System

The proposed system for translation from the sign Language to the Arabic language consists of five basic stages shown in the figure 4. It will in this paper research study of the first and second stages, the first stage which is interested in it to capture images to collect input data, the second stage to interested about the image processing and hand detection and isolate it from the background.

A. Image Acquisition and sensing

In pattern recognition system, first the visual data is captured from the environment using input device like camera, Data entered for this stage give rise for gestures carried out by a number of indicators by wearing the glove dark color in different lighting environments with a light background, or without wearing a glove (natural color of the skin) with a dark background or wear a glove Light-colored and black background so that it output of this stage is a set of colored image (RGB) representing the hand gestures corresponding to each one letter of the Arabic sign language [25]

TABLE I.   ArSL DATASET OF ALPHABETS LETTERSACTIONUNITS

B. Image Analyzing and preprocessing

Pre-processing is a process of preparing data for another procedure. This preprocessing step aims to convert the data into a format that can be more easily and effectively processed [26] ,in this paper, the pre-processing steps are built on the basis of several combinations from the following image processing operations: transfer the RGB image to Gray ,Sobel [27] to edge detection, median filtering, histogram equalization, binary image processing (i.e. threesholding) in HSI color space, and de-saturation. These image-processing operations are discussed in more detail in section 6.

This phase depend on the hand detection of the image, we have focused in this step and that by design algorithm for image processing and detection hand from it, this algorithm is applied to the image, by several methods in order to obtain the best sample can be used later in the classification stage.

In this stage the image is in colored of the type RGB, that contain hand gesture and rear therefore apply image processing operations to be hand isolate or discovered from the Image.

C. Features Extractions

In this stage conducted Hand is described in the image caused by hand detection step according to one of the ways described based on the outer frame and the interior forms,, after that select the best features in the description gesture and distinguish them from other gestures, then, feature extraction, which have been selected for the processing of income data, which are then used in the training process or test.

D. Features Classification

In the stage classification, select one of the statistical classification algorithms , neural networks or any pattern recognition methods, to design classified and was educated under the supervision of the training data (database formed in the previous stage of the system) to classify a new gesture (hand shape) is not represented in the training set to one of the existing varieties. The Inputs for this stage is database features of hand shape, and the output is expected the names of new varieties of thoughtful gestures.
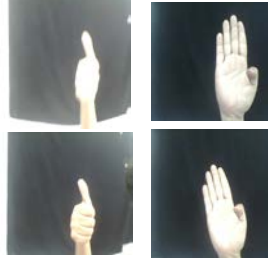
E. Interpretation and Translation

In the translation stage is matched to the expected product was thoughtful gesture name in the stage classification with a letter corresponding to the Arabic language class is expected.

V. THE COMPUTER SIMULATION RESULTS OF IMAGE ACQUISITION

The detail of the first stage is mentioned in this section. In the first stage we work system for capturing images through the laptop integrated webcam and that has enabled us to capture several consecutive images and store, three sets of these images preparation to use as sets in training and testing of the system after processing, the following table shows the models for these sets of images captured from the proposed system that programmed by language Python software tool.

TABLE II.   ArSL DATASET SET TEST AND TRAINING IMAGES

Set Name	Models of Images	Description
1 The first set of test and training Images		A set of gestures to the alphabet of the language of the deaf, it was performed by hand directly without wearing gloves , a white background and lighting in different ways
2 The second set of test and training Images		A set of gestures to the alphabet of the language of the deaf, it was performed by hand directly without wearing gloves and a dark background,
3 The Third set of test and training Images		A set of gestures to the alphabet of the language of the deaf, it was performed by wearing gloves, and a white background,

The following figure shows the part of the user code in the work of the proposed system for capturing images

```python
import numpy as np
import cv2

camera = cv2.VideoCapture(0)
i = 0
while True:
    f,img = camera.read()
    cv2.imshow("webcam",img)
    if (cv2.waitKey(5) != -1):
        break
    cv2.imwrite('{0:05d}.jpg'.format(i),img)
    i += 1
```

Fig. 5.   The part of code by Python to capture images

VI. THE COMPUTER SIMULATION RESULTS IMAGE ANALYZING AND HAND DETECTION

The detail of the first stage is mentioned in this section. At this stage, we process the resulting images of the previous stage of the characters indicative gestures and the captured by different methods, as we explained in the previous table where this stage to implement a number of steps.

A. Image conversion from original color to grayscale

In this step is converting the image matrix with original colors to the data grayscale image, by maintaining the gloss color (luminance) and ignore the color gamut (Hue) and saturation to components of color of the original image, according to the formula used in the color scheme of change from RGB to gray scale I= 0.2989*R+0.5870*G+0.1140*B

```python
filename = 'C:\mywork\ALIF.jpg'
image = cv2.imread(filename)
cv2.imshow('color_image',image)
def rgb2gray(rgb):
    return np.dot(rgb[...,:3], [0.299, 0.587, 0.144])

gray = rgb2gray(image)
plt.imshow(gray, cmap = plt.get_cmap('gray'))
plt.show()
```

Fig. 6.   The part of code by Python to convert images from RGB to gray scale

B. Adjust the contrast of the image

In the second step adjusted the contrast of the image, through noise removal, edge detection and image processing structurally so as to be able to detect hand in the grayscale image. This step was carried out in two different methods; the first method is called Sobel method to detect the edges of the hand in the image, and depend on the use of linear filters to adjust the contrast of the image.

We depend on Sobel detection that refers to computing the gradient magnitude of an image using 3x3 filters [26]. It is detected strongly edges of hand and do not care about the weak edges such as the canny method, as not to lose information from the shape of the hand as in the Laplacian method. The following figure shows the comparison between the results of edge detection in several methods.
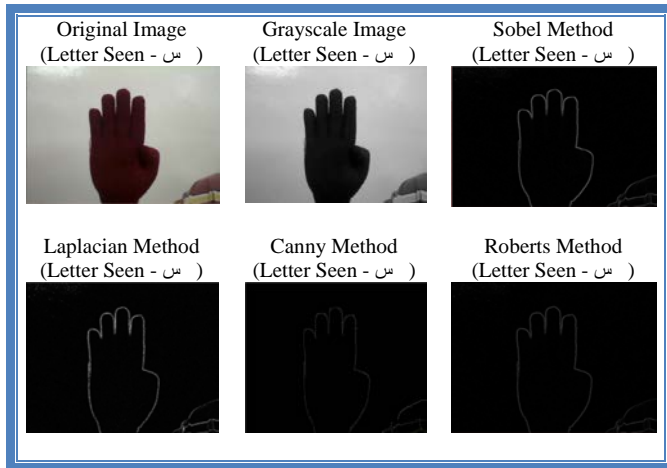
Fig. 7. The comparison between the results of methods to edge detected the letter Seen (حرف السين – س)

The second method relies on the use of experimental threshold to adjust the contrast of the image so that the color scheme pixel that is equal to or less than this threshold equal to the black color values be otherwise be white color while maintaining a grayscale image after that this image(Grayscale) convert to black and white image type and become all matrix values in either black = 0 or White = 1,after that threshold is used as a way to open a structural binary image, to produce a new image free from noise ,the following figure shows the output of this method.
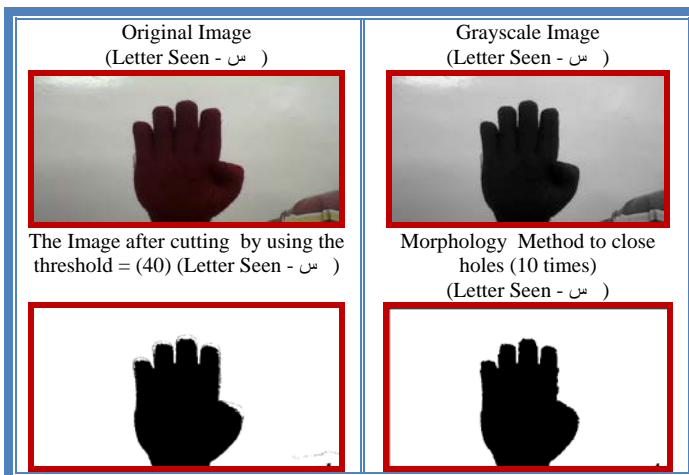

Fig. 8. The results adjust the contrast of the image according to the second method

Was selected threshold (40) on the grounds that the signal performer wearing a dark-colored glove, because the effect of lighting on dark color is less, we can also make the system more flexible programming element visible on the control interface system to control the value of the threshold increase or decrease depending on the color of the glove user, to get a more accurate results,

The levels of gray, ranging from (0) completely black and (255) completely white, And by showing the density of images captured scheme and note the appropriate value that can then separate the image into two different regions, it found that the value of (40) are suitable value experimentally for all the photos.

The third method adjust image contrast density distribution of input image within a new field in the output image values, where the work area with a matching value (0, 0.19) in the income image with the value (0, 1), thus all Pixel in the input image, which is approximately equal intensity $(0.19*255\approx48.5)$ will be equal to a white color density 255. It produces a gray image with pixels according to the following figure then converts the resulting gray image to a binary image.

The figure below illustrates, the stages of implementation of the algorithm detect the edges of the hand, depending on this method.
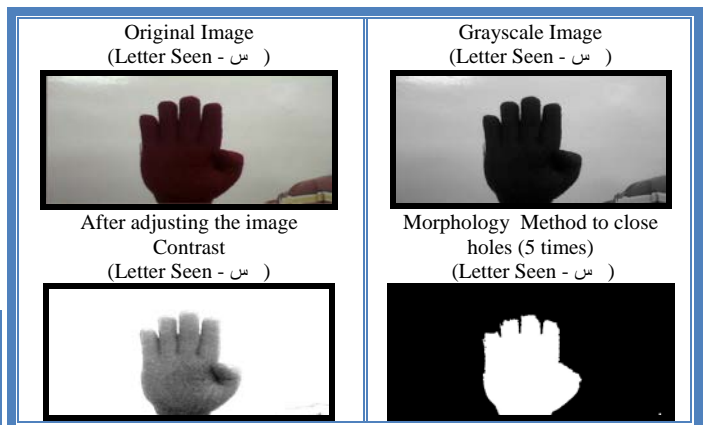

Fig. 9. The results adjust the contrast of the image according to the third method

C. Hand extract element of the image

In this section hand element is extracted from the captured image in the first stage after preprocessing in second stage by using this proposed algorithm that proved effective in the first and second phase of construction of the proposed system, the following figure illustrates the basic steps of the proposed algorithm.
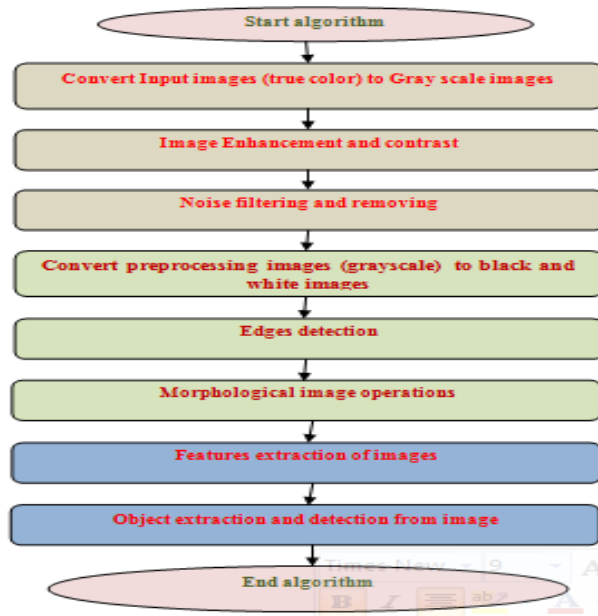
Fig. 10. The main stages of the proposed algorithm to hand extraction in ArSLR system

The researcher explained the results of the first six steps in this algorithm, and will lists the results of the last two steps, which describe the extraction hand. From the image, the detection of the hand area of the image and extracted entitled detected items in the image, the following figure shows the resulting image from the process of extracting the hand element of the original image by the three methods used to detect hand
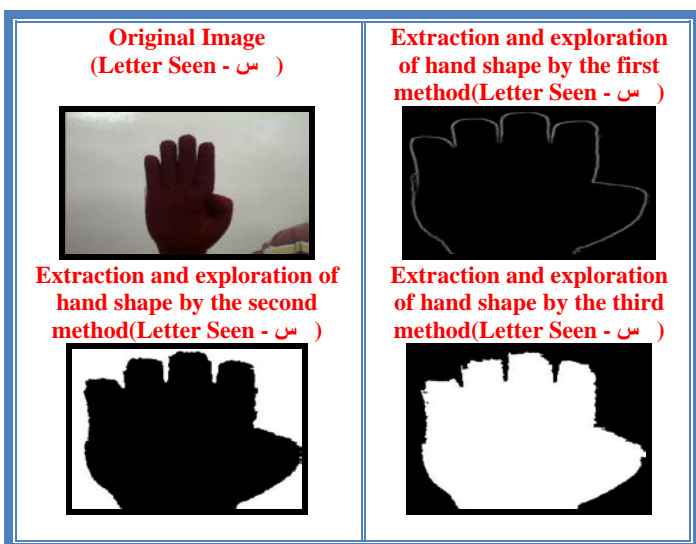

Fig. 11. Hand extract from the background after the detection in the ArSLR system

VII. CONCLUSIONS AND FUTURE DIRECTIONS

A. Conclusions

Simulation work was taking images gestures indicative character system, and processed and analyzed to extract the form of hand Image, the creation of a database digital images of Arabic sign language to be used as data entered and training works base in order to teach her to identify any new gesture does not exist in the training database.

It has been proven that the method of cutting the image using the public threshold color of dark-colored glove is an effective solution to overcome the problem at hand lighting detection of the image.

Conclude that the performance of the gesture does not affect the way the hand detection phase in the proposed system.

Converge the second and third methods in the detection and hand drawn from the image resolution, to the same set of images tested and be in better case detection accuracy, in the case of the second test images and the third group.

B. Future works and directions

The Researcher hopes to complete the study of the stages of building ArSLR system, use it in the education of deaf Bilingual/Bicultural manner language [28], or use an initial model in educational programs for the Deaf by Computer, As well as a means to communicate with the deaf and understand their language.

REFERENCES

[1] Fels S, Hinton G. Glove-Talk: a neural network interface between data-glove and a speech synthesizer. IEEE Transactions on Neural Networks. 1993; 4(1):2-8.

[2] Kim J, Jang W, Bien Z. A dynamic gesture recognition system for the Korean sign language (KSL). IEEE Transactions on Systems, Man, and Cybernetics, Part B. 1996; 26(2):354-359.

[3] Lee C, Bien Z, Park G, Jang W, Kim J, Kim S. Real-time recognition system of Korean sign language based on elementary components. In: Proceedings of the Sixth IEEE International Conference on Fuzzy Systems. 1997:1463-1468.

[4] Tabata Y, Kuroda T. Finger spelling recognition using distinctive features of hand shape. In: 7th ICDVRAT with Art Abilitation. Maia, Portugal: 2008:287-292.

[5] Halawani S. Arabic sign language translation system on mobile devices. IJCSNS. 2008; 8(1):251.

[6] Tsukada K, Yasumura M. Ubi-Finger: a Simple Gesture Input Device for Mobile and Ubiquitous Environment. Journal of Asian Information, Science and Life (AISL). 2004; 2(2): 111-120.

[7] Khalid Alvi A, Azhar M, Usman M, Mumtaz S, Rafiq S, Rehman R, et al. Pakistan Sign Language Recognition Using Statistical Template Matching. In: Proceedings of world academy of science, engineering and technology. Rome, Italy: 2005;3:52-55.

[8] AI-Buraiky S. Arabic sign language recognition using an instrumented glove. Master Thesis, King Fahd University of Petroleum & Minerals, Saudi Arabia. 2004.

[9] Zabulis X, Baltzakis H, Argyros A. Vision-based Hand Gesture Recognition for Human-Computer Interaction. In: Stephanidis C, editor(s). The Universal Access Handbook - Human Factors and Ergonomics Series. Boca Raton, FL, USA: CRC Press; 2009: 1- 30.

[10] Hamada Y, Shimada N, Shirai Y. Hand shape estimation using sequence of multi-ocular images based on transition network. In: Proceedings of the International Conference on Vision Interface. 2002:161-166.

[11] Feris R, Turk M, Raskar R, Tan K, Ohashi G. Exploiting depth discontinuities for vision-based finger spelling recognition. In: Proceedings of the 2004 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'04) Vol 10. IEEE Computer Society; 2004.

[12] Grobel K, Assan M. Isolated sign language recognition using hidden Markov models. In: 1997 IEEE International Conference on Systems, Man, and Cybernetics, 1997.'Computational Cybernetics and Simulation'. 1997:162-167.

[13] Dreuw P, Rybach D, Deselaers T, Zahedi M, Ney H. Speech recognition techniques for a sign language recognition system. In: Interspeech. Antwerp, Belgium: 2007:2513-2516.

[14] Salleh NS, Jais J, Mazalan L, Ismail R, Yussof S, Ahmad A, et al. Sign Language to Voice Recognition: Hand Detection Techniques for Vision-Based Approach. In: Fourth Int. Conf. on Multimedia and ICT in Education, m-ICTE 2006. Seville, Spain: 2006: 967 -972.

[15] Tanibata N, Shimada N, Shirai Y. Extraction of hand features for recognition of sign language words. In: The 15th International Conference on Vision Interface. 2002:391-398.

[16] Mohandes M. Arabic sign language recognition. In Proceedings of the Int. Conf. on Imaging Science. 2001. Las Vegas, Nevada, USA: CSREA PRESS; 2001 ;(1):25-28.

[17] Mohandes M. Automatic Translation of Arabic Text to Arabic Sign Language. ICGST International Journal on Artificial Intelligence and Machine Learning, AIML. 1996; 6(4):15-19.

[18] Foong OM, Low TJ, Wibowo S. Hand Gesture Recognition: Sign to Voice System. International Journal of Electrical, Computer and Systems Engineering (IJECSE). 2009; 3(4):198-202.

[19] https://ar.wikipedia.org/wiki/%D9%84%D8%BA%D8%A9_%D8%A5%D8%B4%D8%A7%D8%B1%D8%A9

[20] Jie Liu, Jigui sun Pattern recognition: An overview, 1st edn. Wang. College of computer science & technology, china. IJCSNS, International Journal of Computer Science & Network Security. VOL.6. NO.6. June 2006.

[21] Navdeep Kaur,Usvir Kaur, Survey of Pattern Recognition Methods, International Journal of Advanced Research in Computer Science and Software Engineering ,Volume 3, Issue 2, February 2013.

[22] Kidiyo Kpalma and Joseph Ronsin, An Overview of Advances of Pattern Recognition Systems in Computer Vision Source: Vision Systems: Segmentation and Pattern Recognition, ISBN 987-3-902613-05-9, IETR, UMRCNRS 6164.

[23] M. Parasher, S. Sharma, A.K Sharma, and J.P Gupta, "Anatomy On Pattern Recognition, "Indian Journal of Computer Science and Engineering (IJCSE), vol. 2, no. 3, Jun-Jul 2011

[24] Sawsan Asjea, Dr. Eng S. Khawatmi and Dr. Eng A.C. Aljundi , Study and Implementation of Sign Language System for Deaf , Aleppo University,Electrical and Electronic ,Faculty Computer Engineering Department, 2010

[25] http://www.3refe.com/vb/showthread.php?t=65830

[26] Hanegan, K., Unpivoting and Pivoting Your Data to Make it Suitable for Analysis, http://spotfire.tibco.com/community/blogs/tips/archive/2010/ 02/19/unpivoting-and-pivoting-your-data-to-make-it-suitableforanalysis. Aspx, (19 February 2010).

[27] G.T. Shrivakshan , Dr.C. Chandrasekar ,A Comparison of various Edge Detection Techniques used in Image Processing IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 5, No 1, September 2012 ISSN (Online): 1694-0814

[28] Andrews, J., Liu, H.T., Liu, C.J., Dacres K. & Gentry, M. 2013. "Adapted little books": An emergent literacy intervention for signing deaf children. The Association of College Educators of Deaf and Hard of Hearing(ACEDHH) Conference, Santa Fe, New Mexico