Skin-Tone Segmentation in Real-Time Vision of Basic Hand Sign Through Pattern Recognition Using Hidden Markov

Raphael Baligod, Leis Alemania, Alaiza P. Dabatos, Adomar L. Ilao
Malayan Colleges Laguna, Cabuyao City, Laguna 4025 Philippines.
alilao@mcl.edu.ph

***Abstract*—The study applied computer vision to perform hand sign recognition commonly affected by skin color during image segmentation. Through YCbCr color algorithm, image color which is based on RGB color space was converted to pure color space which separates luminance and chrominance components resolving skin tone issue of extracting hand contour from real-time right-hand sign. A prototype was developed to recognized hand sign captured by a low-cost web camera. The captured hand sign image passed through canny edge detection and absolute difference threshold technique to construct 2D hand contour in conjunction with Haar classifier and Hidden Markov Model in providing a real-time American Sign Language (ASL) interpretation of the demonstrated hand sign.**

***Index Terms*—Hand Sign; Skin Color; Pattern Recognition; Hidden Markov Model; Haar Classifier.**

I. INTRODUCTION

Disability is a physical or mental condition that limits movements, senses or activity of a person. It is a condition of being unable to do things in the normal way. One form of disability is deafness where according to World Federation of the Deaf [1], there are about 70 million deaf people who use sign language as their first language or mother tongue.

Nowadays, technology is being challenged in inventing new technological tools to improve the lives of people with deaf disability or help society to learn techniques that bridge the gap of communication. There are several scientific contributions on hand gestures in activating application [2][3][4], hand sign recognition in different native languages and numbers [5][6][7].

Hand sign recognition is still affected by skin tone during image segmentation which leads to inaccurate hand sign interpretation [8][9]. After hand segmentation, pattern recognition technique will identify the segmented image by contrasting image characteristics from training set comprised of positive and negative images.

The study used a low-cost web camera as source of real-time hand sign without fancy glove attached to a computer terminal. The real-time hand sign interpretation requires prototype development, which will be tested to different parameters in securing the quality of segmented hand image before analyzing image features to identify the equivalent letter from the American alphabet A to Z

II. EXPERIMENTAL

A. Research Design

The study was designed to verify the real-time hand sign gesture of a basic hand sign learners whether the hand sign gesture has been properly executed without expert participation.

Initially, the study captured real-time hand sign image from an attached web-camera to a computer terminal. The device was manually configured to perform frame-slice per second on the said image feeds. The said image was processed through YCbCr color segmentation algorithm [10] to classify skin-colors by identifying the red color of the skin regardless of the lighting conditions. After image enhancement, image segmentation was performed by canny edge algorithm combined with absolute difference threshold technique generating hand contour.

Normally, sign language is represented by static and dynamic letters. Dynamic letters required hand movement like J and Z.

Static and dynamic hand sign was segmented and analyzed based on image characteristics through pattern recognition techniques using Haar Classifier [7][11] and Hidden Markov algorithm [5][15][16][17][18]. Such was taken from an observation sequence stored in a training set composed of positive and negative images. Through pattern recognition algorithms it was able to identify the equivalent alphabet letter shown on Figure 1.

Prototype was developed using OpenCV library integrated to C# available in Microsoft Studio 2010.

Figure 1: Sign Language Alphabet

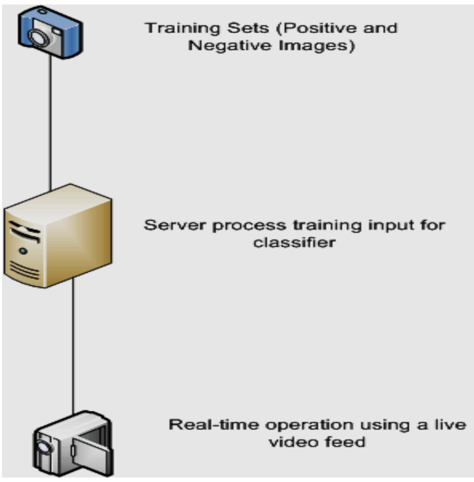

Figure 2: Implemented Research Setup

B. Methodology

The study applied quantitative experimental research design in determining prototype performance and defining appropriate environmental parameters to improve the quality of image segmentation, which is normally affected by different environmental constraints.

Hand sign recognition requires integration of complex algorithms such as image segmentation algorithm and pattern recognition algorithm. However, in order to perform such process, it needs a minimum hardware specification listed in Table 1.

Table 1
Computer Hardware Specification

Parameters	Value/Description
Operating System	Windows 7
Processor	Dual Core from Intel or AMD at 2.4GHz
Memory	4GB
Graphics	nVidia GeForce 8600/9600GT, ATI/AMD Radeon HD 2600/3600
Hard drive Capacity	200 MB available space

During the experiment, low cost web camera was sufficient to clearly seize a hand image either static or dynamic hand gesture. Table 2 shows the specifications of the web camera used in this study.

Table 2
Specifications of the Web Camera

Parameters	Value/Description
Resolution	1300K Pixels in hardware
FPS	30fps/sec
Lens	2-layer glass lens, shimmering inducing
Maximum Supported Video Resolution	1280x960
Video Format	Microsoft AVI
Video Stream Rate	30fps (CIF) and 30Fps (VGA)
Color Depth	24Bit True color
Operating System	Vista/XP/2000/Windows 7/8

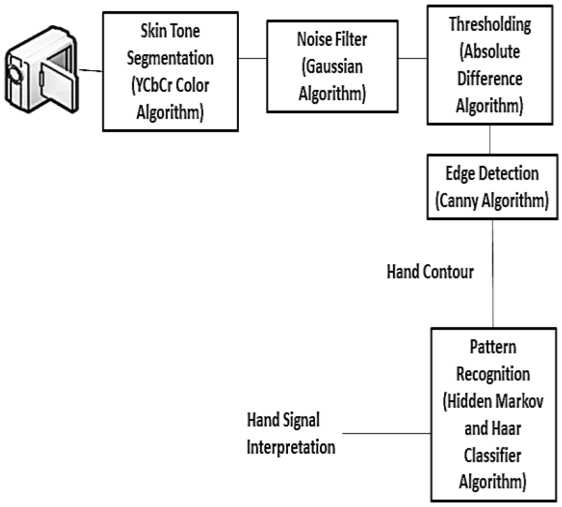

Figure 3: Hand Sign Interpretation Architecture

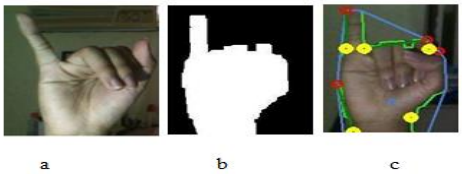

Figure 4: Image segmentation (a) Captured hand sign (b) Hand contour (c) Finger position using Hidden Markov Model

Figure 3 shows data flow of the hand sign recognition. After hand sign acquisition as shown in Figure 4(a), the image was analyzed based on the structure of the signs through shape, position and movement of the hand. Skin detection, based on Region V using YCbCr color space, identifies different frame from captured image after which Gaussian Filter was integrated for image noise reduction. Captured image was converted to grayscale to perform edge detection using canny edge algorithm segment image which leads to hand sign shape. After image reprocessing, inner region of the segmented hand was filled with white color to obtain hand contour by Absolute Difference thresholding technique as shown in Figure 4(b), which is vital in identifying static and dynamic hand sign. Using the hand contour, finger position was identified by Hidden Markov Model shown in Figure 4(c). Hidden Markov Model gesture recognition technique generates gesture modeling, gesture analysis and gesture recognition [10][16][17][18].

Pattern recognition through supervised learning requires training set to associate the different characteristics of a hand sign. The training set stockpiled 30 right-hand of the identified respondents and other available right-hand found under Region V of Fitzpatrick's Skin Color Classification[12] for Haar Classifier and Hidden Markov Model to map hand contour and provide real-time American Sign Language (ASL) interpretation of the demonstrated hand sign.

Haar Classifier and Hidden Markov Model was trained using XML files consisting of 520 positive hand sign images under Region V of Fitzpatrick's Skin Color Classification with different hand position angle over the different background colors and two light conditions and 520 negative images for a total of 1,040.

III. RESULTS AND DISCUSSION

The study needs to identify different environment parameters where the said prototype will be able to recognize hand sign based on distance, hand angle, lighting condition (dim light / normal light) and background color.

In order to acquire the proper environmental parameters, several experiments were conducted. The detection of the hand sign was presented in 0's and 1's, in which 0 and 1represent a miss and a hit, respectively. Classification of hit from miss detection was determined by the prototype based on its ability to identify the equivalent letter from a real-time hand sign.

The highest accuracy percentage resulted from test trails based on the lists of web camera configuration, image background colors and lighting conditions are the appropriate parameters to lessen computer vision constraints, which normally reduces the efficiency of image segmentation algorithm.

The experiment shows ideal parameters to any computer vision application when a low cost web camera was utilized as the acquisition medium to capture images during real-time setting. For hand position, hand must be strictly 90 degrees to properly generate a hand sign interpretation. Any hand sign performed more or less than 90 degrees would interfere with hand sign interpretations.

The test trials on hand distance from the web camera started from 10 cm up to 100 cm resulting to a minimum distance of 60cm up to a maximum distance of 90 cm.

Trials on background colors with normal lighting condition were conducted using white, blue, light blue, green, light green and black. It shows that YCbCr skin-tone model proved to be an effective algorithm. However, dim light condition resulted to another image constraint. Dim lighting condition generated shadow which interfered with Y-space (*luminance space*) generating background noise which distorted the hand segmentation. It contradicted YCbCr skin-tone model as stated, it should eliminate the need for an appropriate light condition [13]. However, white background with normal light condition has been utilized as the background color of the real-time hand sign experiment.

The ideal environmental settings resulted from the experiment such as distance, hand angle position, light condition and background color was utilized to determine the prototype performance in terms of accuracy percentage of each letters of the American alphabet.

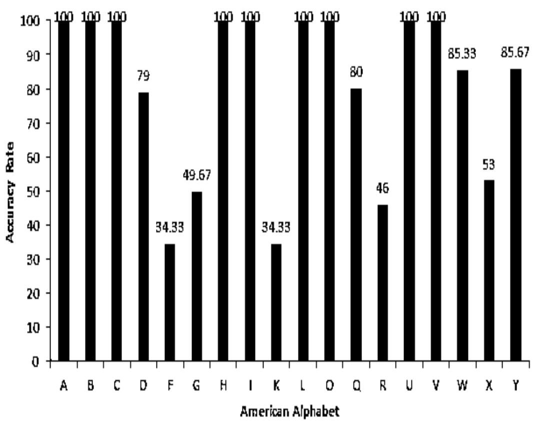

Figure 5: Prototype's Accuracy Percentages of A to Z

Figure 5 shows the accuracy rate of hand sign recognition. Accuracy rate of static letters ranged from 34.33 up to 100%. They were affected by the depth of certain hand signs. Dynamic letters such as J and Z produced false recognition due to the reasons that the initial position of *J* is the position of *I*. Similar findings to letter *Z,* initial position is similar to static position of *D*. Letters such as *E, M, N, S, T* also produced false recognition. These letters required depth hand dorsal analysis where 2D environment as applied by Mitra (2007) [14] failed to identify the depth of the hand dorsal. Other letters such as *F, G, K, R, X* also need minimal amount of depth analysis.

IV. CONCLUSION

The YCbCr color model was an effective algorithm for properly detecting a skin-tone color under Region V of Fitzpatrick's Skin Color Classification. There were no color conversions between the Luminance components of the image, therefore the threshold set for the Region V skin-tone region was not altered despite lighting conditions. Hand dorsal becomes a major problem for a 2D image even after implementing Hidden Markov model where several hand contours having minor variations, which makes it significantly difficult to differentiate them.

ACKNOWLEDGEMENT

Authors would like to express our greatest gratitude to the support given by the College of Computer and Information Science, Malayan Colleges Laguna, Cabuyao City, Laguna, Philippines.

REFERENCES

[1] Access Date: July 10, 2014. http://wfdeaf.org.

[2] Alsheakhali, M., Skaik, A., Aldahdouh, M., and Alhelou, M. 2011. Hand Gesture Recognition System. *Information & Communication Systems*.132.

[3] Rajesh, J. R., Nagarjunan, D., Arunachalam, R. M., and Aarthi, R.. 2012. Distance Transform Based Hand Gestures Recognition For Power Point Presentation Navigation. *Adv. Comput.* 3:41-48.

[4] Zaletelj, J. Comparison of Video and Depth-based Hand Gesture Interfaces.

[5] Lee, H. K., and Kim, J. H. An HMM-based threshold Model Approach For Gesture Recognition. *Pattern Analysis and Machine Intelligence, IEEE* .199;21(10):961-973

[6] Vanco, M., Minarik, I., and Rozinaj, G. 2014. Evaluation of static Hand Gesture algorithms. *Systems, Signals and Image Processing (IWSSIP)*. International Conference on IEEE:83-86

[7] Bui, T. T. T., Phan, N. H., and Spitsyn, V. G. Face and Hand Gesture Recognition Algorithm Based On Wavelet Transforms And Principal Component Analysis. *Strategic Technology (IFOST)*.2012;2012 7th International Forum on IEEE:1-4.

[8] Kelly, D., McDonald, J., and Markham, C. 2008. A System For Teaching Sign Language Using Live Gesture Feedback. In Automatic Face & Gesture Recognition. *8th IEEE International Conference.* ;1-2.

[9] Nadgeri, S. M., Sawarkar, S. D., and Gawande, A. D.. 2010. Hand Gesture Recognition Using CAMSHIFT Algorithm. *Emerging Trends in Engineering and Technology (ICETET)*. 37-41.

[10] Rumyantsev, O., Merati, M., and Ramachandran V. 2012. Hand Sign recognition through palm gesture and movement. *Image Processing*.

[11] Dabre, K., and Dholay, S. 2014. Machine Learning Model For Sign Language Interpretation Using Webcam Images. *Circuits, Systems, Communication and Information Technology Applications (CSCITA)*. International Conference on IEEE:317-321

[12] Phung, S. L., Bouzerdoum, A., and Chai Sr, D. Skin Segmentation Using Color Pixel Classification: Analysis And Comparison. *Pattern Analysis and Machine Intelligence, IEEE* .2005;27(1):148-154;

[13] Kakumanu, P., Makrogiannis, S. and Bourbakis, N. A Survey Of Skin-Color Modeling And Detection Methods. *Pattern Recognition*.March 2007; 40(3):1106–1122.

[14] Mitra, N. J., Guibas, L. J., and Pauly, M. . Symmetrization. *ACM Transactions on Graphics (TOG)* .2007;26(3): 63.

[15] Hornegger, J., Niemann, H., Paulus, D., and Schlottke, G. (1994). Object Recognition Using Hidden Markov Models. *Pattern Recognition in Practice IV: Multiple Paradigms, Comparative Studies and Hybrid Systems*, *16*: 37-44.

[16] Chen, F. S., Fu, C. M., and Huang, C. L. 2003. Hand Gesture Recognition Using A Real-Time Tracking Method And Hidden Markov Models. *Image and vision computing*, *21*(8): 745-758.

[17] Vogler, C., and Metaxas, D. 2001. A Framework For Recognizing The Simultaneous Aspects Of American Sign Language. *Computer Vision and Image Understanding*, *81*(3): 358-384.

[18] Starner, T. E. 1995. *Visual Recognition of American Sign Language Using Hidden Markov Models*. Massachusetts Inst Of Tech Cambridge Dept Of Brain And Cognitive Sciences.