Sign Language Gesture Recognition through Computer Vision

2 authors:


Ruth Diko Wario
University of the Free State


Casam Njagi Nyaga
University of the Free State

IST-Africa 2018 Conference Proceedings
Paul Cunningham and Miriam Cunningham (Eds)
IIMC International Information Management Corporation, 2018
ISBN: 978-1-905824-59-5

Sign Language Gesture Recognition through Computer Vision

Casam Njagi NYAGA, Ruth Diko WARIO
Department of Computer Science and Informatics, University of the Free State, Private Bag X13, Kestell 9866, Republic of South Africa
Tel: +27716640550, Email: 2016185122@ufs4life.ac.za, warioRD@ufs.ac.za

Abstract: This paper presents a study that was conducted as a usability test on an existing gesture recognition system. Computer vision gesture recognition can offer hope in creation of a real time interpreter system that can solve the communication barrier that exists between the deaf and the hearing who don't understand sign language. The objectives of this study were to determine the effectiveness of the system, to determine the efficiency of the system and to determine the satisfaction of the deaf participants on the use of the system. In the study, 7 deaf participants evaluated the usability of the gesture recognition system. The researcher employed observation data collection technique followed by an interview which was facilitated by a sign language teacher. The participants found the system to be effective and efficient. All the participants also appeared to be interested, desire to be involved and were motivated by the system.

Keywords: computer vision, deaf, gesture recognition, sign language, and usability

1. Introduction

Research in computer vision hand gesture recognition has demonstrated to be interesting and challenging[1][2]. Hand gesture research can be grouped in two categories. The first category is computer vision based method, which enables development of visual based gesture recognition systems. Gestures involve movement of the hands, face, or other parts of the body in order to convey messages [3].The second category is glove based method that enables the measurement of different parameters like hand position and location of the fingertips[4]. This study was based on the second category which is vision based. The system that was tested utilizes the web camera to capture the South African Sign Language (SASL) gestures from the deaf participants and then display the respective finger spelling numerical characters. The sign language is the main communication method for people with hearing impairment. The deaf face communication challenges when interacting with the hearing who don't understand the sign language. To addressed this challenge computer vision can be applied to Human Computer Interaction HCI to create sign language interpretation and learning systems[5].

2. Objectives

The primary objective of the study was to establish the level of usability of the gesture recognition system. The following objectives were pursued in order to achieve the above primary objective:

1. To determine the effectiveness of the system.
2. To determine the efficiency of the system.
3. To determine the satisfaction of the deaf participants on the use of the system.

3. Methodology

The research instruments that were used in this study included observation and interview.

3.1 Observation

The researcher used an observation checklist, which contained the usability indicators used to measure the effectiveness, efficiency and satisfaction of the system. During the testing of the system, observations were made**.** User testing, involves the users completing tasks using the system or a prototype system in a controlled environment [6]. There are several methods used to collect data during a user test and they include the think-aloud protocol, co-discovery or co-participation, question-asking protocol, performance measurement, teaching methods and remote testing. This study used the performance measurement technique of user testing whereby data about the user's performance is captured while the users perform predetermined tasks on the system to be evaluated [7].

3.2 Interviews

An interview is a face-to-face social communication between the assessor (e.g. researcher) and the interviewee (e.g. user). There are four different kinds of interview techniques, namely unstructured, structured, semi-structured and focus group interviews [8]. In this study, a structured interview was conducted. The researcher conducted interviews after the user testing sessions through the assistance of a sign language teacher as an interpreter. The teacher asked the interview questions by signing to the students then verbally explained to the researcher who recorded the answers. The questions that were asked are as below.

1. Based on your interaction with the system, what improvements would you recommend?
2. What challenges and problems did you face while using the system?
3. Based on your interaction with the system, what did you like about it?
4. Do you think the system has the ability to facilitate training for deaf participants?

3.3 Population and Sampling

A population represents all conceivable elements, subjects or observations relating to a particular area of interest to the researcher [9]. Kumar (2015) defined sampling as the process of selecting a few (a sample) from a bigger group (the sampling population) to become the basis for estimating or predicting the prevalence of an unknown piece of information, situation or outcome regarding the bigger group. The target population for this study was all the 10 students in grade 10 of a school for the deaf that is situated near the University of the Free State. The school was identified purposively because it is a special school for the deaf while grade 10 students were identified as the target population because they are senior in age and are more experienced with computer use than the lower grade students. As stated above purposive sampling was applied in this research. In purposive sampling the researchers select their sample on the basis of their own knowledge of the population, its elements, and the nature of the research aims. In short, it is based on their own judgment or the purpose of the study [11].

In this study, the researcher purposefully identified a sample of 7 participants from grade 10. This sample size is well representative of the population because early studies indicated that 5 users in a usability test would uncover approximately 80% of the usability problems in a product [12],[13],[14].

4. Technology Description

The system was developed using Open Source Computer Vision (OpenCV) and python computer programming language. OpenCV is a cross-platform and free library of programming functions mostly targeted at real-time computer vision [15]. Python is an open source high-level programming language designed to be easy to read and simple to implement [16]. The system was run on an Elite Book HP laptop running on windows 7 operating system.

4.1 Procedure of Data Collection

The researcher obtained an ethical clearance from the University. The researcher also got permission from the school management and used an informed consent form that was signed by the teacher after she explained well to the students about the research study. The participation in the study was on voluntary basis. The research was conducted over a period of one month in November 2017. The most time was spent on preparation and analysis of the results of the usability study but the main study took one day since the number of participants was small.

4.2 User Testing Tasks

In this study each student got a chance to test the system. The students were required to perform 4 tasks with the system the tasks are summarised in the table below

Table 1: Task Details

S/N	Task Description	Task Details
1	Task 1	Finger spelling for number 1
2	Task 2	Finger spelling for number 2
3	Task 3	Finger spelling for number 3
4	Task 4	Finger spelling for number 4

Based on the above tasks the functionality of the system was determined with the indicators used to measure effectiveness, efficiency and satisfaction as explained in the results section.

4.3 Flow Chart of the System

Below is a flow chart to representing the steps followed by the system as it executes.

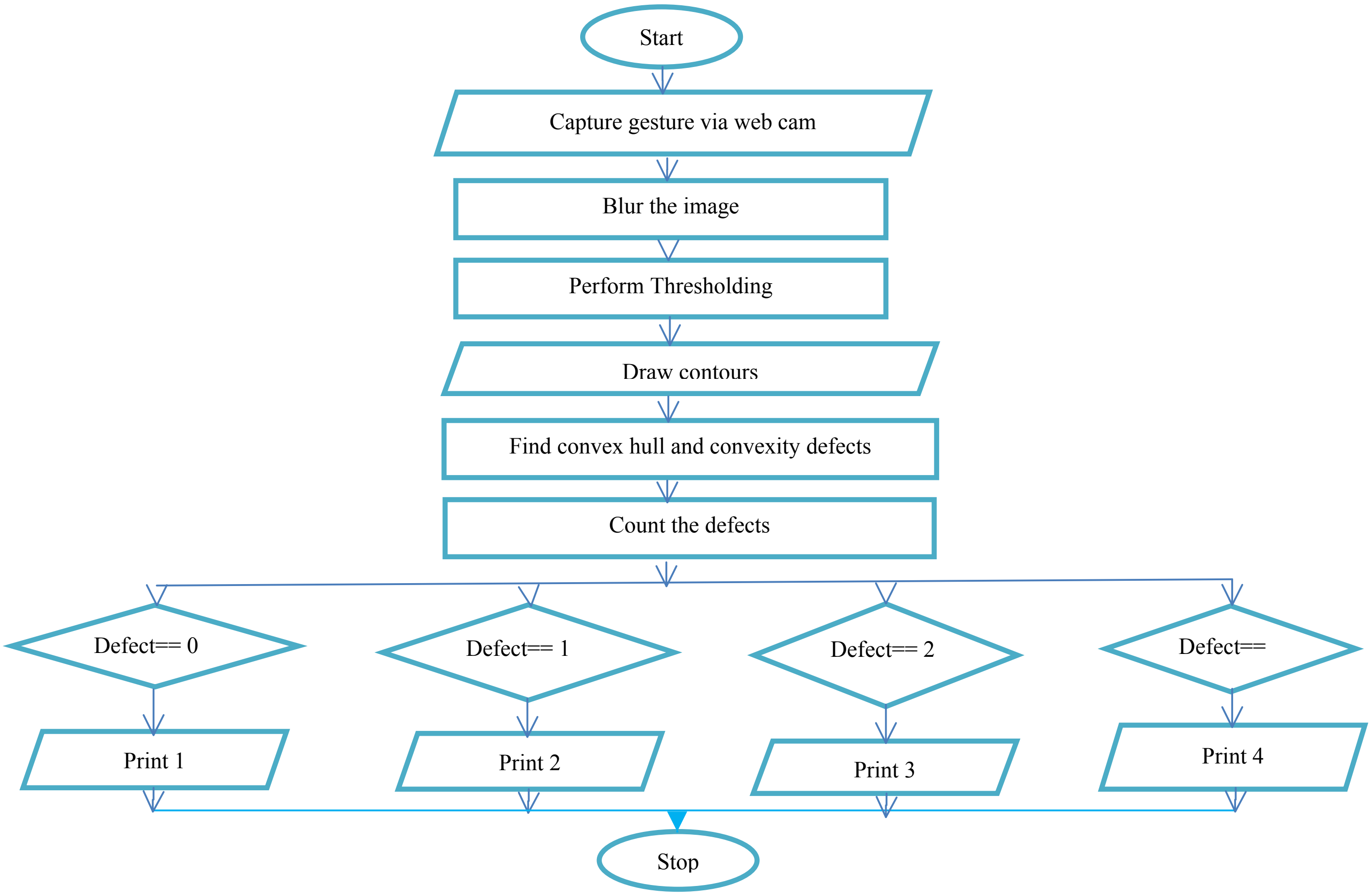

Figure 1: Flow chart vision based gesture recognition system

Below is a short description of the system process represented by the above flow chart.

- Capture the video frames

The first step in the program is to capture the video through the web camera. The captured video stream is a frame of images. The images are converted from RGB to grayscale. With the grayscale image it is easy to identify the pixels that are of interest and those that are not because the skin pixels tend to be similar in intensity.

- Blur the image

The second step is to blur the image in order to make it smooth and to reduce the noise. This is done because the interest is not on the details of the image but the shape of the hand which is to be tracked. In this program Gaussian Blurring was used on the original image.

- Thresholding

The third step is to use thresholding for image segmentation, to create binary images from grayscale images. In thresholding a low pass filter is applied by allowing only particular color ranges to be highlighted as white while the other color are suppressed by showing them as black. In this program Otsu's Binarization method was applied.

- Draw contours

The fourth step is to draw contours along the boundaries of the black and white image. The region is then enclosed with an envelope that is referred to as the convex hull.

- Find convex hull and convexity defects

The fifth step is to find the convex points and convexity defects. The convex points are the tips of the finger while the convexity defects are the deepest points of deviation on the contour.

5. Results

Table 2: Participant's Demographics

Gender	Frequency	Percentage %
Male	4	57
Female	3	43
Total	**7**	**100**

5.1 Observation Results

Study participants were observed while they were performing their user testing tasks. The main purpose of the observation was to watch the participants in terms of their behaviour in order to gauge their satisfaction level. A total of 7 participants were observed. The results are as follows:

5.2 Effectiveness

In respect to this study, effectiveness refers to being able to successfully complete a task [17][18] the metrics that were used to measure effectiveness included:

1. Task completion rate – which is the total number of completed tasks divided by the number of participants.
2. Number of errors – for the purpose of this study an error is any unsuccessful attempt to make a correct gesture.

5.2.1 Task Completion Rate

Table 3: Task Completion Rate

Task	Completed task	Incomplete task	number of participants	Task completion rate %
Task 1	7	0	7	100%
Task 2	6	1	7	86%
Task 3	7	0	7	100%

Task 4	5	2	7	71%

As indicated table 2 above, the task completion rate for task 1 and task 3 was 100% which means that all the participants were able to complete those tasks with ease. For task two one of the participants was not able to complete the task and for task 4, two participants were not able to complete the task within the first attempt.

5.2.2 *Number of Errors*

The errors were experienced due to the error descriptions indicated in table 3 that is each participant was required to place his hand within a rectangular box that is considered as the region of interest. The region of interest is used in order to reduce the background noise. It proved challenging for some participants to place their hands within the region of interest and then perform the required task. It was also noted that some participants placed their head and hand within the region of interest which caused an error.

Table 4: Number of Errors

Task	Number of errors	Error description
Task 1	0	
Task 2	1	Hand out of region of interest
Task 3	0	
Task 4	2	Hand out of region of interest Head and hand within region of interest
Total	3	

From table 4 above it can be seen that 100% of the participants appeared focused, involved and interested. The results also show us that 86 % of the participants appeared motivated while 71% of the participants appeared not impatient. On the other hand the results show us that 57% of the participants appeared not to be nervous while 43% of the participants were calm.

Table 5: Observation Summary

Behaviour	Category	Frequency	Percentage %
Appears interested	Yes	7	100
	No	0	0
Appears impatient	Yes	2	29
	No	5	71
Appears calm	Yes	3	43
	No	4	57
Appears nervous	Yes	3	43
	No	4	57
Appears involved	Yes	7	100
	No	0	0
Appears motivated	Yes	6	86
	No	1	14
Appears focused	Yes	7	100
	No	0	0

5.3 *Efficiency*

Efficiency is the amount of effort required to complete a task [17][18] the metric used to measure efficiency included:

1. Task completion time – total time spent to complete the four gestures by the deaf participant.
2. Number of completed task – total number of completed gestures

5.3.1 Task Completion Time

In this study the task completion time was taken to be the total time spent to complete the four tasks by the participants

Table 6: Task Completion Time

Participant ID	Task 1Time (sec)	Task 2 Time (sec)	Task 3 Time (sec)	Task 4 Time (sec)
Participant 1	1.38	2.08	1.79	3.82
Participant 2	4.95	2.61	4.58	3.15
Participant 3	3.12	1.94	2.42	4.70
Participant 4	1.67	1.35	1.2	1.34
Participant 5	1.23	1.52	2.71	3.79
Participant 6	1.41	2.29	2.43	3.61
Participant 7	3.49	1.23	2.54	1.46

Table 7: Task Completion Time summary

Task ID	Mean Time taken (sec)	Min time taken (sec)	Max time taken (sec)
Task 1	2.464286	1.23	4.95
Task 2	1.86	1.23	2.61
Task 3	2.524	1.2	4.58
Task 4	3.124	1.34	4.7

Efficiency can be measured by (ISO /IEC 25062:2006) as the ratio of the task completion rate to the mean time per task.

Table 8: Task Efficiency

Task	Task Compilation rate %	Task Time (min)	Efficiency
Task 1	100%	1.23	123
Task 2	86%	1.23	106
Task 3	100%	1.2	120
Task 4	71%	1.34	95

From table 7 above, it is easy to see that for task 1 has the highest efficiency level followed by task 3 and then task 2, while task 4 has the least efficiency.

5.4 Interview Questions and the Responses

The results of the interview revealed the following:

Question one addressed the improvements to the system where majority of the participants indicated that no improvements were needed. One respondent indicated that they had no idea on what to improve. Similarly one participant indicated that the system needed improvements while one participant indicated that the system should include more tasks that is, it should identify more numbers. Question two addressed the issue of challenges and problems faced by the participants. From table 8 below, it is easy to see that three participants faced no challenges while 3 indicated that it was difficult to place the hand on the rectangular box that captures the gesture. One participant indicated that the system was slow. Question three addressed what the participants liked about the system. As indicated in table 8, two participants liked the system, one participant was excited by the system, two participants found it useful because it can detect their hands while two participants had no idea. Question four addressed if the system has the ability to facilitate training for the deaf. Two participants responded that the system will be very helpful while five responded that the system will be helpful.

Table 9: Interview Summary

Interview question	Number of responses			
	No Idea what to improve	No improvements	Include more tasks	There should be improvements
Question 1	1	4	1	1
	No challenges in usage	Difficult to put hand in region of interest	System was slow	
Question 2	3	3	1	
	Likes the System	Excited	Find it useful (can detect hand)	No Idea
Question 3	2	1	2	2
	Very helpful	Helpful		
Question 4	2	5		

6. Business Benefits

The findings of the study will form guidelines that will be important in designing future vision-based gesture recognition systems. The findings will also bring insight into how sign language recognition could be utilised for the indexing task, providing the required basis for sign language search engines. Another envisaged application is the development of a sign language tutor to be used in the areas of both education and rehabilitation. It could support patients suffering from hearing loss, deaf people with sign language deficiencies, as well as interested hearing people, in learning sign language. This study may also be of benefit to researchers or teachers who might be willing to conduct similar studies in the future.

7. Conclusions

Sign language gesture recognition as the study found out will be helpful to deaf people in assisting them to interact with computers. The system was found to be highly effective and efficient by the participants. All the participants also appeared to be interested, involved and motivated by the system. However, it was not easy for most of the participants to place their hand on the rectangular box, which was considered to be the region of recognition. Hence it is appropriate to recommend that a bigger area is required so that the participants will not have to struggle to place the hand in the region before they start performing the gestures. The system needs to be improved to be able to recognize all the South Africa alphabet and numerical characters. For the system to be effectively applied in a controlled environment like a deaf classroom then a white background can be placed behind every seat of the computer user to reduce the challenge of complex background noise and improve the background subtraction.

References

[1] Y. Zhou, G. Jiang, and Y. Lin, "A novel finger and hand pose estimation technique for real-time hand gesture recognition," *Pattern Recognit.*, vol. 49, pp. 102–114, 2016.

[2] K. Liu and N. Kehtarnavaz, "Real-time robust vision-based hand gesture recognition using stereo images," *J. Real-Time Image Process.*, vol. 11, no. 1, pp. 201–209, Jan. 2016.

[3] & R. Mahesh, "Gesture recognition: A survey of Gesture Recognition techniques using Neural networks," *Glob. J. Comput. Sci. Technol. Neural Artif. Intell. Vol. 13 Issue 3 Version 1.0.*, 2013.

[4] L. & Kulkarni, "Appearance Based Recognition of American Sign Language Using Gesture Segmentation IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 03, 2010, 560-565," 2010.

[5] E. & N. Kyriakos, "A Dynamic Gesture and Posture Recognition System. DOI 10.1007/s10846-013-9983-7," 2014.

[6] M. Nektarios, K., Stavrinoudis, D., Sokoli, S. & Xenos, "Combining experimental and inquiry methods in software usability evaluation: the paradigm of LvSeducational software," *J. Syst. Inf. Technol.*, vol. 12, no. 2, pp. 120–139, 2010.

[7] J. Nielsen, “Usability 101: Introduction to Usability,” 2012. [Online]. Available: http://www.nngroup.com/articles/usability-101-introduction-to usability.
[8] J. H. Fontana, A. & Frey, *"Interviewing: the art of science". In: Denzin, NY. L. eds. Handbook of Qualitative Research.* Los Angeles: Sage Publications, 1994.
[9] O. & Tijani, *Essentials of Business Research Methods*. Lagos: CSS Bookshops Limited, 2011.
[10] Kumar, *Research Methodology: A step- by- step guide for beginners*. London: Sage publications, 2015.
[11] & B. Mouton, *The practice of social research*. Cape Town: Oxford University Press, 2004.
[12] Virzi, “Refining the test phase of usability evaluation: How many subjects is enough? Human Factors,” no. 34, pp. 457–468, 1992.
[13] J. Nielsen, “Usability engineering. Boston: AP Professional,” 1993.
[14] L. Faulkner, “Beyond the Five-User Assumption: Benefits of Increased Sample Sizes in Usability Testing,” *Behav. Res. Methods, Instruments, Comput.*, vol. 35, p. 3, 2003.
[15] V. Pulli, Kari; Baksheev, Anatoly; Kornyakov, Kirill; Eruhimov, “‘Realtime Computer Vision with OpenCV’. Queue. pp. 40:40–40:56. doi:10.1145/2181796.2206309,” 2012.
[16] P. Christensson, “Python Definition. Retrieved 2017, Dec 2, from https://techterms.com,” 2010.
[17] Albert & Tullis, “Measuring the User Experience: Collecting, Analysing, and Presenting Usability Metrics. Morgan Kaufmann Publishers Inc. San Francisco, CA, USA,” 2008.
[18] Z. Lv, A. Halawani, S. Feng, S. ur Réhman, and H. Li, “Touch-less interactive augmented reality game on vision-based wearable device,” *Pers. Ubiquitous Comput.*, vol. 19, no. 3, pp. 551–567, Jul. 2015.