
IMPLEMENTATION OF VISION BASED SIGN LANGUAGE FOR EFFECTIVE COMMUNICATION: A SOLUTION TO THE PROBLEM OF RECOGNIZING SIGN LANGUAGE

J.P.Banusha Lakmal Mahinda and J.C. Balasuriya

Abstract

This paper addresses some of the issues faced by the differently abled, especially people incapable of speech (the mute), who are involved in Open and Distance Learning (ODL), and proposes a solution by using popular technologies such as computer vision, image processing and machine learning.

People without speech are often discriminated in society due to problems of communication, and therefore, are unable to contribute to the development of society. With regard to open and distance learning (ODL), proper communication is a must. A solution can be found by implementing the Real Time Sign Language Recognition System.

Instead of a hardware based approach, a vision based approach for Real Time Sign Language Recognition System has been used. A vision based approach is a cheaper approach than the hardware based approach. Real Time Sign Language Recognition system gets the input from a normal web camera in real time. It first recognizes the region of the hand and afterwards the correct hand posture from the geometric parameters of the hand (width height ratio and angle of the hand). Later it gives output in the form of text/voice. In today's world, popular technologies such as computer vision technology, image processing technology and machine learning technology are the core technologies that are used to implement this system. The above mentioned technologies interact with each other to produce a proper system output.

Further, the communication barrier between the differently abled and the "normal" can be reduced to a certain extent by implementing these kinds of systems, and the differently abled will be better able to contribute to the development of society.

KEY WORDS – *Open Distance Learning, Sign Language, Recognition, Real Time, Computer Vision, Image processing, Machine Learning*

Introduction

Communication is fundamental for human beings. It is vital for understanding people and their behavior. People use communication for sharing and interacting with others and to gain important information about places, peoples, personal feelings, things, and events. Communication is absolutely essential for building relationships with people in order to satisfy most needs. A significant portion of

the day is spent establishing and maintaining a rapport with people. In fact, communication is one of the most important factors in learning. Communication is defined as sending, giving, or exchanging information, and ideas which are often expressed non-verbally and verbally. Non-verbal communication is the act of saying what's on your mind without speech.

There are people who are unable to express their ideas verbally. 'Deaf-mute' is a term which was used historically, by "normal" people, to identify a person who is both deaf and mute from their birth. The term continues to be used to refer to the deaf, mainly within a historical context, to indicate people without speech, or who have some degree of speaking ability, but who do not try to speak or due to the lack of attention for typical voices. But the modern world does not accept this term to identify a person who is both deaf and mute, as they use standard sign language for communication. Sign language is a solution to a certain extent. But it is not an appropriate solution for the above problem.

People without speech use sign language to express their ideas. But the issue is that they still have a difficulty in communicating, both among themselves and with others. In fact, it is often more difficult to conduct a simple conversation between a "normal" individual and one that is deaf or mute, because the so-called "normal" individual is usually not familiar with a Sign Language. So these types of people have a problem expressing their ideas effectively, and are therefore, discriminated in society.

A sign language (also signed language) is a language which, instead of acoustically conveyed sound patterns, uses visually transmitted sign patterns (manual communication, body language and lip patterns) to convey meaning, simultaneously combining hand shapes, orientation and movement of the hands, arms or body, and facial expressions to fluidly express a speaker's thoughts. Sign languages commonly develop in deaf communities, which can include interpreters, friends and families of the deaf as well as people who are deaf or hard of hearing themselves.

Real time sign language to recognition system is the proposed solution for communication problems among people without speech and "normal" people who do not know sign language. The above system will have capability of detecting the sign language in real time using a video camera and converting it in to voice out put. This system consisted of three modules: Video Capturing module, Hand detection module and Hand recognition module.

Objectives

The main objective of the system is to find a solution for reducing the communication gap between people who use sign language for their communication and people who verbally communicate. In addition, one needs to consider how computer vision techniques are applicable in solving these kinds

of problems and how these techniques can be used efficiently and effectively. To this end, I first used OpenCV (Open Source Computer Vision) library with C++, next I designed a system for solving the above mentioned problem, and finally I evaluated the proposed solution.

Research issues and Study Framework

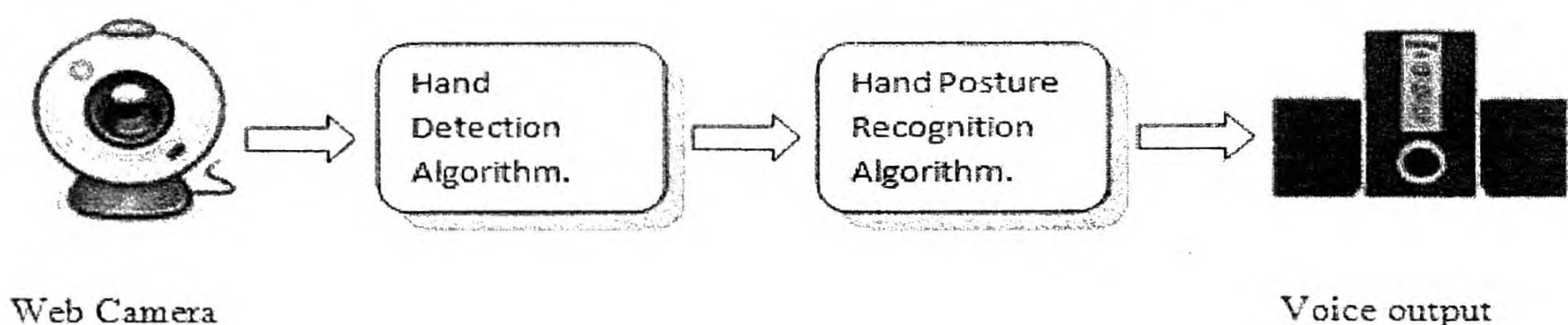
This paper will focus on the design which has been made to cater to both the functional and nonfunctional requirements. This included the overall design, designs of the individual components of the system, etc.

The authors' first approach was to identify the problem domain and identify the difficulties faced by people without speech when they communicate with each other. Through web research and interviews, the authors were able to locate the hand postures and the meanings of the English alphabet and the numbering systems. The system was implemented according to the American Sign Language (ASL).

The functional requirement is to see how the system is somewhat performing transformation of a set of input/s to a corresponding output/s. Those transformations would be the functions of the system. In Real Time Sign Language Recognition System, the basic functional requirement is to give the voice output according to the identified real time hand posture. Non functional requirements are the qualities of a system which are the characteristics of the system that cannot be expressed in functional requirements. Those are, the system should be accurate, productive and reliable, the system should be fast and should send quick responds, and the GUI should be simple, attractive and efficient. Before going to implement the system, I clearly identified functional and non-functional requirements.

At the very top level, the system is divided in to two categories. Those are Hand detection module and Hand recognition module. The very top level diagram is shown in Figure 3.1.

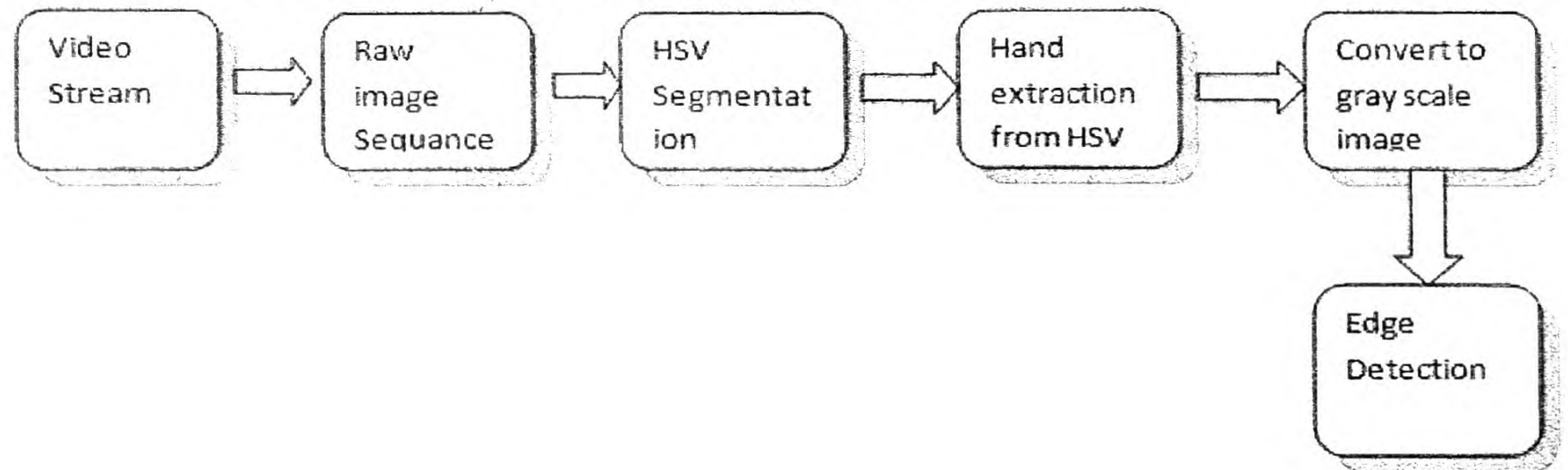
Figure 3.1 Top level architecture of the Real time sign language recognition system



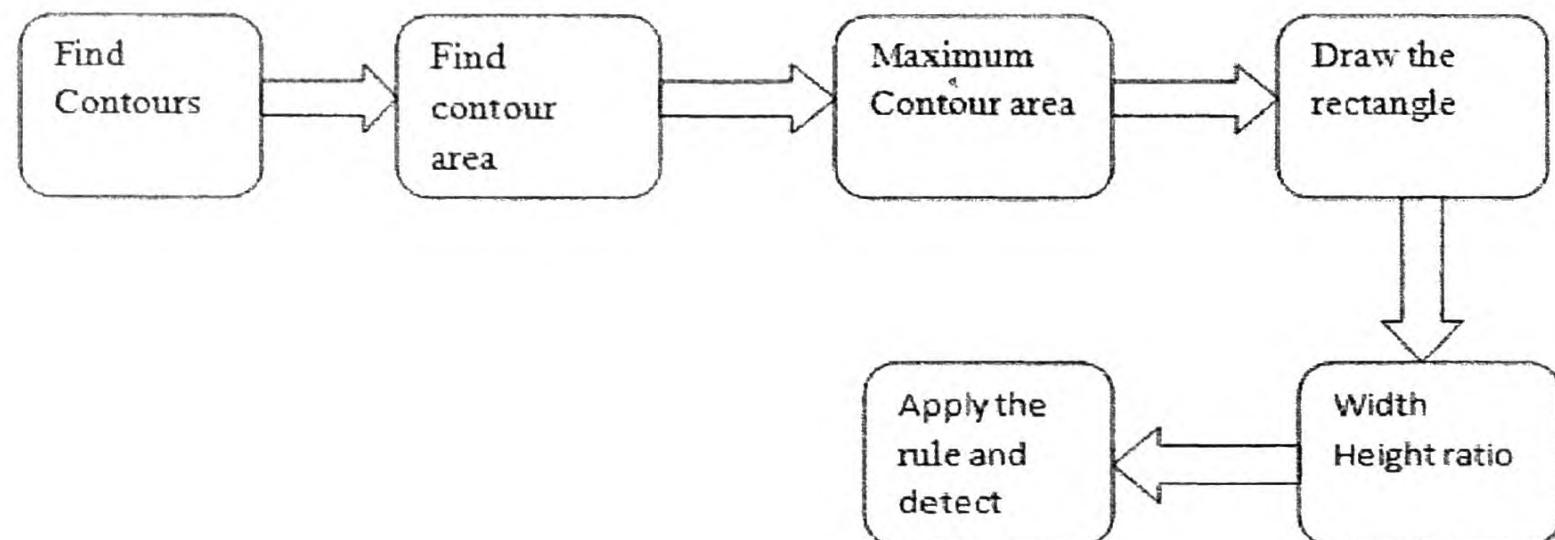
Hand detection module is the most critical part of the system. There are various methods available for hand detection. Here, the skin color based hand detection algorithm is used. The reason for selecting the method is that skin color is unique to every person. The advantage of the method is that it is dependent on distance

and background, as Real Time Sign Language Recognition System adopts different types of backgrounds and different types of lighting conditions. This method is useful for working in these types of complex backgrounds. The algorithm of the hand detection module is shown below Figure 3.2.

Figure 3.2 – Skin Color based hand detection algorithm



After detecting the hand, the system should have the ability to recognize hand posture. Here, the authors used shape parameters of the hand (width height ratio and angle of the hand). First the system was trained for shape parameters (width height ratio and angle of the hand) to recognize each and every posture. When the system is at work, it should identify the correct hand posture using pre-trained postures parameter set. How this hand posture recognition module works is shown in figure 3.3.



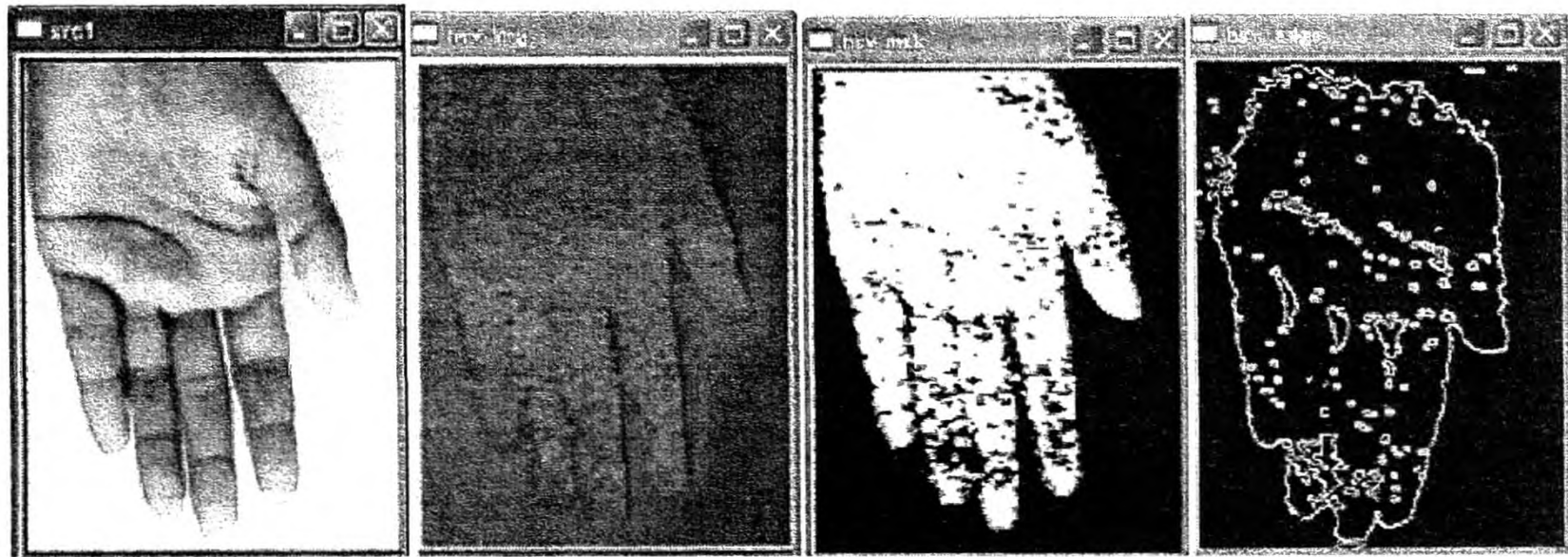
Methodology

To implement this system, modern techniques of computer vision and machine learning are used. The paper, provide a detailed description of how these techniques are used to implement the above mentioned system.

For the hand tracking step, authors have used the HSV color due to the characteristic color values of human skin. Using color information that is similar to a skin color of the candidate region which has high value, only the hand region from the background image is extracted.

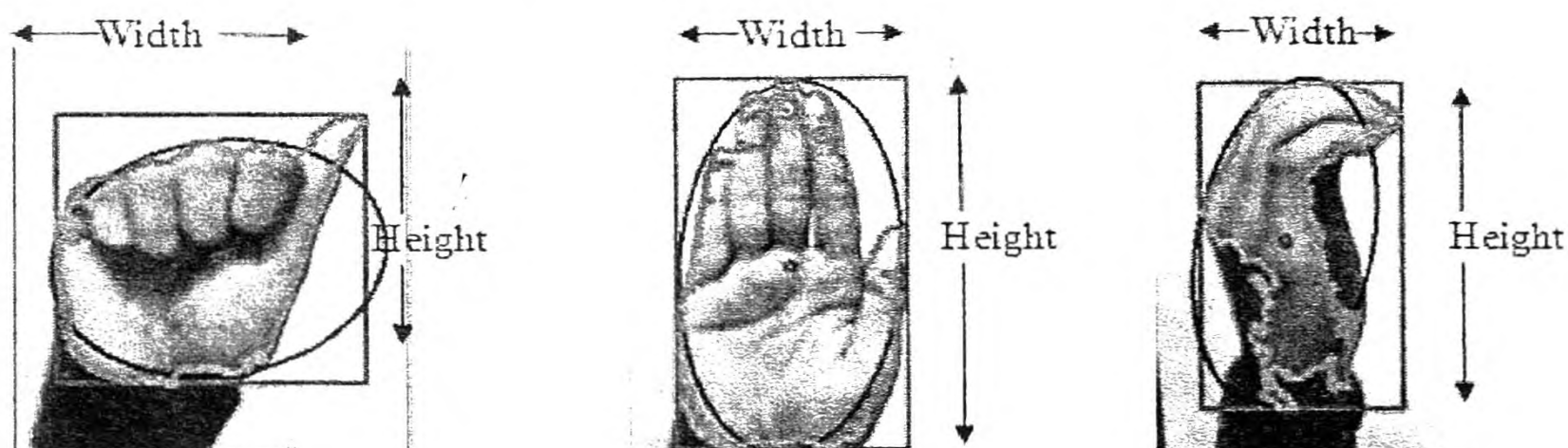
In this process, first, real time video has been converted in to image sequence with 20frames/sec. Then these images have been converted in to HSV image sequence. Afterwards, the writers defined upper (max) and lower (min) HSV values that were to be extracted from the image. (In this case max and min HSV values of the human skin color). For skin color, Hue in between 0 and 20, Saturation in between 30 and 150, value in between 80 to 255. After the application of this filter for HSV image sequence, the writers have extracted the hand region by using skin color values of the HSV images .Afterwards HSV image was converted into gray scale image. Gray scale image is a binary image and here it represents hand region from bit value 1 (white color) and other background area from bit value 0 (black color). Black color to white color has a huge gradient. After that, hand region edges have been extracted using the technique called canny edge detection. Figure 4.1 shows how the image transformation happens.

Figure 4.1 – Image transformation



After extracting the hand region by using the method of canny edge detection, the boundary of the hand has been extracted by using the running length coding contours representation. This identifies the contour area and finds the maximum contour area. The maximum contour area is the hand boundary region and a rectangle can be drawn surrounding the hand region which extracts the width to height ratio of that hand surrounding region. This is essential for the identification of the hand location; By using the contour representations, one can get the current angle of the hand. By using both width to height ratio of the hand region, rectangle and current angle of the hand, it was possible to identify the current hand posture; and from variance of both area rectangle ratio and the angle of the hand, one can identify the hand gesture using the pre defined rules. Pre defined rules are rules that are defined in the training process. Then the system identifies the correct hand posture and the gesture and gives voice output meaning of the gesture and the posture. Figure 4.2 shows the different width height ratios of the different hand postures.

Figure 4.2 – Different width height ratios of different hand postures.



Findings and Discussion

Real time sign language recognition system is a kind of interface for the differently abled.

Some attempt to find a solution through a variety of ways. They use a glove based technique instead of the vision based techniques. In the glove based technique, they mostly deal with hardware technologies; users of the system need to buy the hardware. This method is not natural, nor cheap. So I try to address this issue using the vision based approach. At present, most research follows hand gesture recognition. Most people are engaged in experiments on hand gesture recognition for human computer interaction (HCI) and robot control. I use vision based techniques such as hand gesture recognition for sign language recognition. This system captures the real time video from a web camera and extracts the hand region and recognizes the hand posture; the output is in text/voice format. This is more natural than using hardware gloves.

The real time sign language recognition system has certain limitations. The main problem is the condition of lighting. When lighting conditions are changed, it affects the final output of the system to some extent. Another factor is the background. Changing the background also affects the final out put of the system. Therefore, I have considered these two problems and plan to further explore how to minimize the effect of the condition of lighting and the background changes.

This program has been tested using 2.4 GHz, 1GB RAM PC with visual C++ programming language using the OpenCV libraries. A low-priced USB color camera was used as image input device, and resolution size of images was 320×240. For performance evaluation, the system has been tested on a set of 5 users. Each user has performed a predefined set of 6 postures and therefore, we have 60 postures to evaluate the application results. Table 5.1 shows the performance evaluation results. In here, the recognition rate is calculated by applying the bellow equation.

$$rate = \frac{\alpha}{\beta} \times 100\%$$

α = Frames Recognize posture correctly

β = Total number of frames

Table 5.1 - Recognition percentages of hand postures

Person	Posture							Total
	A	B	C	D	E	F		
#1	98%	99%	99%	85%	80%	90%	91.8%	
#2	99%	99%	99%	82%	84%	93%	92.6%	
#3	97%	98%	98%	84%	84%	91%	92%	
#4	98%	97%	99%	82%	85%	92%	92.1%	
#5	99%	99%	99%	81%	83%	93%	92.3%	

There are lots of approaches available for these kinds of systems. Some for them are Ada-boost Haar classification, neural network based approach and appearance based approach. By comparison, those methods have certain kinds of limitations. If we use Haar training for the above system, we need 1000-2000 positive images for each posture and 1000 negative images. The training time for a particular posture is nearly a week. If we use this method, we need to switch on the computer the whole week when the training completes. So this method is very inefficient as we need lots of resources. The neural network based approach has its limitations as well as it is a real time system and produces an image each 0.0005 sec; it is unable to properly handle that sequence and training.

Conclusion

Communication is the most essential factor for the human being. A specific segment of society has been neglected by the rest of the society because of difficulty in communication, which has motivated the research, to create the real time sign language recognition system for people without speech.

I have developed a gesture and posture recognition system which can recognize sign language and give the voice output. This system is fully automatic and works in real time. I used Open CV, (Open Source Computer Vision Library) with visual C++ for implementation. In this system, the current frame rate is 20 frames/sec.

There are several advantages of this system: one need not take into consideration the background and one need not wear gloves. Currently this system recognizes 6 signs with 90-95% accuracy at a good lighting condition.

However, one can also observe some limitations of this system. When the lighting condition becomes dull, the recognition rate rapidly goes down. This problem can occur when skin-like colors are present in the background.

The Real Time Sign Language Recognition System recognizes only postures. But sign languages have both postures and gestures. In this current system, there is no mechanism for the recognition of postures which is an area that needs to be looked at in future.

References

- David J. Sturman, David Zeltzer (1994), *A Survey of Glove-based Input*, pp. 30-39
- Hong Shan Neoh, Asher Hazanchuk, *Adaptive Edge Detection for Real-Time Video Processing using FPGAs*, San Jose, CA 95134
- OpenCV haartraining (Rapid Object Detection with a Cascade of Boosted Classifiers Based on Haar-like Features), <<http://note.sonots.com/SciSoftware/haartraining.html>>
- Pakistan Sign Language Recognition Using Statistical Template Matching, (2005), Aleem Khalid Alvi, M. Yousuf Bin Azhar, Mehmood Usman, Suleman Mumtaz, Sameer Rafiq, Razi Ur Rehman, Israr Ahmed, viewed 30 March 2010, <<http://www.waset.org/journals/waset/v3/v3-13.pdf>>
- Takafumi Miyatake, Hitoshi Matsushima and Masakazu Ejiri, (1997), *Contour representation of binary images using run-type direction codes*, *Machine Vision and Applications*, pp 193-200.
- V. A. OLIVEIRA, A. CONCI, *Skin Detection using HSV color space*, Computation Institute - Universidade Federal Fluminense - UFF - Niterói, Brazil.
- Xiaoming Yin, Ming Xie (August 2007), *Finger identification and hand posture recognition for human-robot interaction*, *Image and Vision Computing*, Volume 25 , Issue 8, Pages: 1291-1300.

CONTRIBUTORS

- Abeyssekera N, Senior Lecturer , Department of Management Studies, Faculty of Humanities and Social Sciences, The Open University of Sri Lanka
nabey@ou.ac.lk
- Balasuriya, J.C. (Dr.) Senior Lecturer, Faculty of Information Technology, University of Moratuwa, Sri Lanka *jcbala@uom.lk*
- Bulankulame I, Lecturer, Department of Social Studies, Faculty of Humanities and Social Sciences, The Open University of Sri Lanka
- Fernando T.A. (Dr.) Senior Lecturer, Department of Social Studies, Faculty of Humanities and Social Sciences, The Open University of Sri Lanka
- Jayatilleke B.G.(Dr.) Senior Lecturer, Educational Technology Division, The Open University of Sri Lanka, *bgjay@ou.ac.lk*
- Kandasamy, Maheswari (Professor) Open University Malaysia
- Karunanayaka S, (Dr.) Senior Lecturer, Secondary and Tertiary Education Faculty of Education, Open University of Sri Lanka *spkar@ou.ac.lk*
- Kulasekera G. U Lecturer, Educational Technology Division, The Open University of Sri Lanka *dgkul@ou.ac.lk*
- Lekamge D (Professor), Dean, Faculty of Education, Open University of Sri Lanka, *gdlek@ou.ac.lk*
- Liyanage B. C.(Dr.) Senior Lecturer , Department of Civil Engineering, Faculty of Engineering Technology, The Open University of Sri Lanka *bcli@ou.ac.lk*
- Mahinda J.P.B. L , Faculty of Information Technology, University of Moratuwa, Sri Lanka *banushalakmal86@gmail.com*
- Ratwatte H. (Dr.) Senior Lecturer, Department of Language Studies, Faculty of Humanities and Social Sciences The Open University of Sri Lanka
hratw@ou.ac.lk
- Settinayake, W.A.R.P. University of Moratuwa, Sri Lanka *rakee677@gmail.com*
- Silva N de, (Emeritus Professor) Faculty of Medicine, University of Kelaniya.
ndsil@ou.ac.lk , desilva@slt.lk
- Thanaraj T. Senior Lecturer, Secondary and Tertiary Education, Faculty of Education, Open University of Sri Lanka, *thanaraj2006@gmail.com*
- Vidanapathirane, U (Professor) Vice Chancellor of the Open University of Sri Lanka
- Weeratunga, L. Secretary to the President of Sri Lanka
- Wickramasinghe T. Lecturer, Department of Social Studies, Faculty of Humanities and Social Sciences, The Open University of Sri Lanka
thiwankee_w@yahoo.com
- Wickramasinghe V (Dr.) Senior Lecturer, Department of Management of Technology, Faculty of Engineering, University of Moratuwa, Sri Lanka
vathsala@mot.mrt.ac.lk

National Digitization Project

National Science Foundation

Institute : National Science Foundation

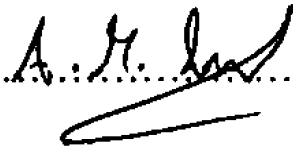
1. Place of Scanning : Sanje (Private) Ltd, Hokandara

2. Date Scanned :02/06/2017.....

3. Name of Digitizing Company : Sanje (Private) Ltd, No 435/16, Kottawa Rd,
Hokandara North, Arangala, Hokandara

4. Scanning Officer

Name :Angelo Melvin Luwis.....

Signature :.....

Certification of Scanning

I hereby certify that the scanning of this document was carried out under my supervision, according to the norms and standards of digital scanning accurately, also keeping with the originality of the original document to be accepted in a court of law.

Certifying Officer

Designation :Information Officer.....

Name :Renuka Sugathadasa.....

Signature :.....

Date :02/06/2017.....

"This document/publication was digitized under National Digitization Project of the National Science Foundation, Sri Lanka"