HAND SIGNAL DETECTION USING CAMERA

FAIZZATI BINTI MOHD ZAID

This report is submitted in partial fulfillment of the requirement for the award of
Bachelor of Electronic Engineering (Industrial Electronic) With Honors

Faculty of Electronic and Computer Engineering
Universiti Teknikal Malaysia Melaka

June 2016

UNIVERSTI TEKNIKAL MALAYSIA MELAKA
FAKULTI KEJURUTERAAN ELEKTRONIK DAN KEJURUTERAAN KOMPUTER

BORANG PENGESAHAN STATUS LAPORAN
PROJEK SARJANA MUDA II

Tajuk Projek : .**HAND SIGNAL DETECTION USING CAMERA**

Sesi Pengajian : 1 2 / 1 3

Saya **FAIZZATI BINTI MOHD ZAID** mengaku membenarkan Laporan Projek Sarjana Muda ini disimpan di Perpustakaan dengan syarat-syarat kegunaan seperti berikut:

1. Laporan adalah hakmilik Universiti Teknikal Malaysia Melaka.
2. Perpustakaan dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan dibenarkan membuat salinan laporan ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. Sila tandakan (√) :

☐	**SULIT***	*(Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972)
☐	**TERHAD****	**(Mengandungi maklumat terhad yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)
☐	**TIDAK TERHAD**	

Disahkan oleh:

_____ (TANDATANGAN PENULIS)

Tarikh: ……..

_____ (COP DAN TANDATANGAN PENYELIA)

Tarikh: ……..

"I hereby declare that the work in this project is my own except for summaries and quotations which have been duly acknowledge."

Signature : .....

Author : FAIZZATI BINTI MOHD ZAID

Date : .....

"I acknowledge that I have read this report and in my opinion this report is sufficient in term of scope and quality for the award of Bachelor of Electronic Engineering (Industrial Electronics/ Computer Engineering/ Electronic Telecommunication/ Wireless Communication)* with Honours."

Signature : .....

Supervisor"s Name : EN. ZULKARNAIN BIN ZAINUDIN

Date : .....

Specially dedicated to

my beloved parents and siblings who have encouraged, guided and inspired me

throughout my journey of education.

ACKNOWLEDGEMENT

Praise to Allah S.W.T that with His blessings I‟m able to finish this project in time successfully. This project was completed after I have working for it for two semesters. I have learnt and gained a lot of experience and my knowledge was improved within these two semesters.

First of all, I would like to deliver my sincere appreciation to my supervisor, Encik Zulkarnain Bin Zainudin. Encik Zulkarnain has tought me from the beginning of my project starts. He guides me back to the corect path whenever I lost. Besides, he teaches me in a manner that easy to understand by giving some examples. Futhermore, I really appriciate Encik Zulkarnain for giving me guidance to improve my softskill, too.

I would like to express my gratitude to my friends who share with me the knowlegde in solving computing technical problem. My friends were helped me a lot by sharing the experience, listen to my problems and give explanation on the problems I faced.

Furthermore, I would like to express my deepest appreciation to my beloved family members who always support me morally. They willing to listen to me and encourage me all the times. I am grateful for everyone‟s constant support and help.

ABSTRACT

Hand signal detection (HSD) using camera is a way to create a useful and highly adaptive interface between machines and their users. The aim of the project is to develop a system that able to detect our hand gestures in order to perform designated tasks. It can be done with signal processing techniques to recognize images of hand gestures from camera. This technique can assist a disabled person to perform specific tasks such as switching on and off light, fan or air-condition. The recognition of gesture is difficult because gesture exihibit human variability. It have a varius types of systems and methods available for hand detection recognition. This thesis focuses on hand signal detection using camera by proposing a combinational algorithm to detect, tracking, and recognize for interaction with an application via hand signal. There are three major parts in HSD; first are the input, second is processing data, and lastly the result of hand signal. Each section has divided into subsections which are proposed to achieve their own goals.

To be more specific, this project is done especially for disabled people who cannot walk, and cannot perform any physical task as mentioned before. The only way for them to communicate is through sign language or gestures. This project is done by using template matching method by Matlab software. Finally, a complete hand signal detection has been successfully develope with the spesific algorithm approach. All the objectives have been successfully achieved.

ABSTRAK

Pengesan isyarat tangan (HSD) menggunakan kinect adalah satu cara yang sangat berguna untuk mewujudkan perantaraan di antara mesin dan pengguna. Tujuan projek ini adalah untuk mewujudkan satu sistem yang dapat mengesan isyarat tangan untuk melaksanakan tugas-tugas yang ditetapkan. Ia boleh dilakukan dengan teknik pemprosesan isyarat untuk mengenali imej isyarat tangan daripada kamera. Teknik ini boleh membantu orang yang kurang upaya untuk melaksanakan tugas-tugas tertentu seperti menghidupkan dan mematikan lampu, kipas atau penghawa dingin. Mengenal pasti isyarat adalah sukar kerana manusia mempamerkan pelbagai isyarat. Ia mempunyai pelbagai jenis sistem dan kaedah yang ada bagi mengenal pasti pengesanan tangan. Tesis ini memberi tumpuan kepada pengesanan isyarat tangan menggunakan kamera dengan mencadangkan gabungan algoritma untuk mengesan, menjejaki dan mengenali untuk berinteraksi dengan aplikasi melalui isyarat tangan. Terdapat tiga bahagian utama di dalam HSD; pertama adalah input, kedua pemprosessan data, dan akhir sekali keputusan isyarat tangan. Setiap bahagian telah dibahagikan kepada subseksyen yang dicadangkan untuk mencapai matlamat mereka sendiri.

Untuk lebih spesifik, projek ini dijalankan terutamanya bagi orang-orang kurang upaya yang tidak boleh berjalan, dan tidak boleh melaksanakan apa-apa tugas fizikal seperti yang dinyatakan sebelum ini. Satu-satunya cara untuk mereka berkomunikasi melalui bahasa isyarat atau isyarat tangan. Projek ini dijalankan dengan menggunakan kaedah padanan template oleh perisian Matlab. Akhir sekali, pengesan isyarat tangan yang lengkap telah berjaya diwujudkan dengan pendekatan algoritma spesifik. Semua objektif telah berjaya dicapai.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

This chapter is mainly discussed the overview of human motion detection system and its possible applications. A briefly explanation about the system operation, the design, program and also the scope of the project can be found in the following chapter.

1.1 Project Overview

The reason for hand signal detection project been design by aiming basic shapes made by human bare hand according to image subtraction from the background and at the same time suspiciously detect hand detection. This technique can assist a disabled person to perform specific tasks such as switching on and off light, fan or air-condition.

Communication in our daily life is generally in voice, but hand signal has its own significance and sometimes it play an important role in order to conveying the information. Furthermore, to develop gesture recognition requires high quality cameras and complicated computer vision algorithms to recognize and analyze hand signals that made by user.

As we know, web camera is a simplest tool, which is user just need to plug in the web camera to a personal computer and run the software to start monitoring. By using this tool, there are many advantages such as easy for real-time user interaction. Besides that, it also allows the development of algorithms that classify and perform recognition of the image data.

Additionally, a reason why this hand signal detection using camera can became very important because it also can help people to switch off or on their appliance without touching the switch. It also can prevent people from get electrical shock when switch off or on something because most off the time people really don‟t care about their safety by switch off or on using wet hand. It also can help people who are deaf and dumb, that are why hand signal plays an important role in people daily life. Gestures are the very first form of communication. So this area influenced a very much to carry on the further work related to hand signal detection recognition.

1.2 Project Objectives

The objectives of this project is to design and build a prototype for hand signal detection and to gather all the knowledge had been gained and put them into an applications by building this model. In order to develop this system, the basic concept is required especially in software programming. The main objectives of this project are as follows:

i. System is not limited to any persons to assist them from move to other place.
ii. To develop a smart silent communication.
iii. To create a natural user interface gestures for the system.

1.3 Problem Statement

Nowadays technology development between human and computer keep evolving from time to time. This is due to the fact that human will not satisfy with what they get and they always wanted more. However, to make human-computer interaction easier going is not too easy.

There is several Implementations have been created in order to achieve smooth interaction between human and computer. The goal for this project is to implement by using web camera based hand signal detection system that is capable to perform a specific task.

1.4 Scope of Project

There are three major parts in hand signal detection using camera sensor. First is the input, processing data, and lastly output. Each section has divided into subsections which are proposed to achieve their own goals.

For input part, a camera been used as a RGB camera. The purpose using this tool as input because it can detect and capture the object or movement and have their own build in data. In this project, the reference movement is user hand movement.

For processing data, MATLAB software had been used to process the data and do the image processing. This part is an important part, where the software needs to be developing in order to process the input signal data. Lastly is the output, where specific task will be the output of this project. The specific task of the project will be decide by the processing data where it will fully depend on the input that created by the people.

1.5 Project Methodology

1.5.1 Project Process Flow

i. Choose the project title

ii. Analysis the project

iii. Do the literature review, project objectives, problem statement, and methodology

iv. Design the concept and development

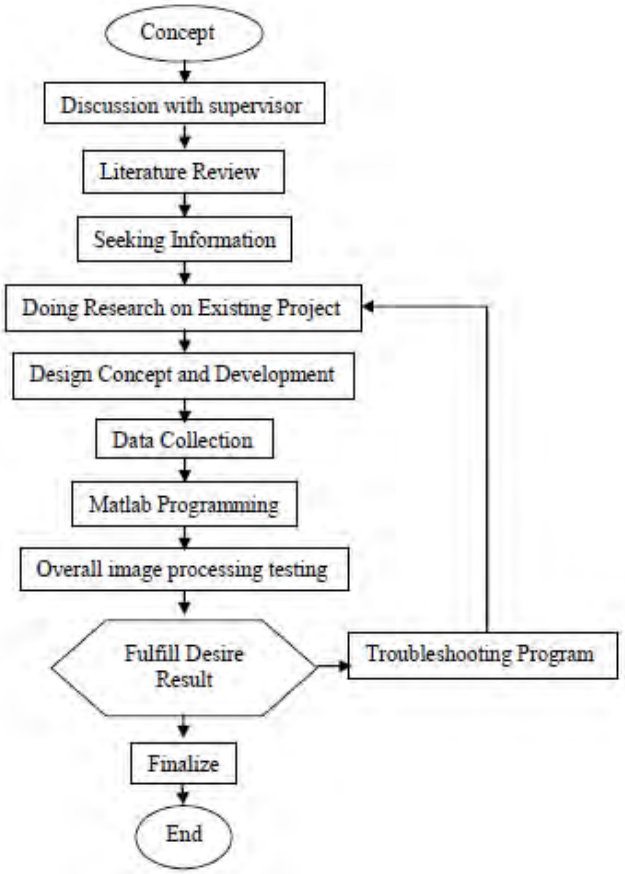
v. Collecting the images (data)

vi. Prepare the software

vii. Troubleshooting and analysis

viii. Final presentation

1.5.2 Flowchart



1.6 Report Structure

The report consists of five chapters. Chapter 1 discusses about the introduction of the project which includes the objectives of the project, problem statements, and scope of the project, project significance and report structure.

Chapter 2 discuss about the Literature review. This chapter included literature review on hand signal detection and its application.

The next chapter, chapter 3 discusses about project methodology of the project. The methodology involved system analysis, system design, system development and system testing.

Then, chapter 4 discusses about Result and Discussion of the project. It is about result full decision after the completion of this project.

The final chapter is chapter 5 that explains about conclusion and recommendation for future work related to this project.

CHAPTER 2

LITERATURE REVIEW

This chapter will be some literature review that provides a simple summarization together with the synthesis obtain from the references studies for the project purpose. It is important to give an overview about the knowledge on security system and gain more detail about web camera, hand and signal detection with some specific area. All the information collected is important to ensure this project achieved its own objectives.

2.1 Introduction

Recently, there has been a surge of interest on hand detection, tracking, and gesture recognition. While a comprehensive literature review of this field would be a daunting challenge because of the huge number of publications and concentrate in this chapter about a review of the most representative developments that are pertinent to the thesis research topics.

Human hands are the parts that people use the most and with the most ease to interact with the world. Hands are highly articulated structures with some 27 degrees of freedom (DOF). Because of these high DOF of the human hand, hand gesture recognition is indeed an extremely challenging task. Even though hand postures and gestures are frequently considered as being identical, there are actually differences. While the hand posture is a static motionless pose, such as making a palm posture and holding it in a certain position, the hand gesture is a dynamic process consisting of a sequence of changing hand postures over a short duration, as for instance waving the hand.

Because of the inherent complexity of the hand gesture, its recognition can be divided into two stages: the low stage hand posture detection and the high stage hand gesture recognition. For a computer to recognize a hand gesture, first the hand should be detected from the captured image, and then the recognition of the gesture made by that hand should be done in a similar way as humans do. However, this is a more challenging technique to be implemented due to the limitations of such a natural system. Segmentation of the hand image is the first stage of all these systems.

Gloves or finger marks have been used to extract the hand posture information from the image and facilitate the hand segmentation process by removing the varying skin color issue of the problem. This approach permits the system to detect hands easily, and it is invariant against lightning conditions which may change. Another way for solving the problem is by implementing a hand gesture recognition system with a uniform background such as a black curtain.

On the other hand, vision-based methods are more natural and useful for real time applications. Besides, computer-vision-based interfaces provide many outstanding advantages:

i. Computer vision is nonintrusive.

ii. Sensing is passive and silent.

iii. Installed camera systems can be used for other jobs.

iv. Sensing and processing hardware is commercially available at low cost.

On the other hand, vision based systems need application-specific image processing algorithms, programming, and machine learning. The need for all these to work in a variety of environments causes several issues because these systems require user and camera independent and invariant against the background and lighting changes to attain real-time performance.

A reliable set of characteristic features and relevant information of how they correlate in representing hand gestures are needed in order to successfully recognize hand gestures. It provides a characterization of gestures‟ spatial properties that divides gesture interpretation into two group‟s 3D hand model-based and appearance-based.

Appearance-based methods were developed for hand posture and gesture recognition directly from images using visual features such as hand contours and fingertips positions. Other methods, called 3D model-based, provide a geometrical representation of the hand configuration using the joint angles of a 3D hand‟s articulated structure recovered from the sequence of images.

2.2 Edge detection

Edge detection is an important image processing technique with wide range of applications. Several edge detection algorithms have been developed in the past few decades, however no single algorithm is suitable for all types of applications. One of the main applications of edge detection techniques is in the process of image segmentation and object detection.[6] Edge maps help in representing faces as a

single unit. An edge detection system has wide variety of applications, as shown in Figure 2.2.

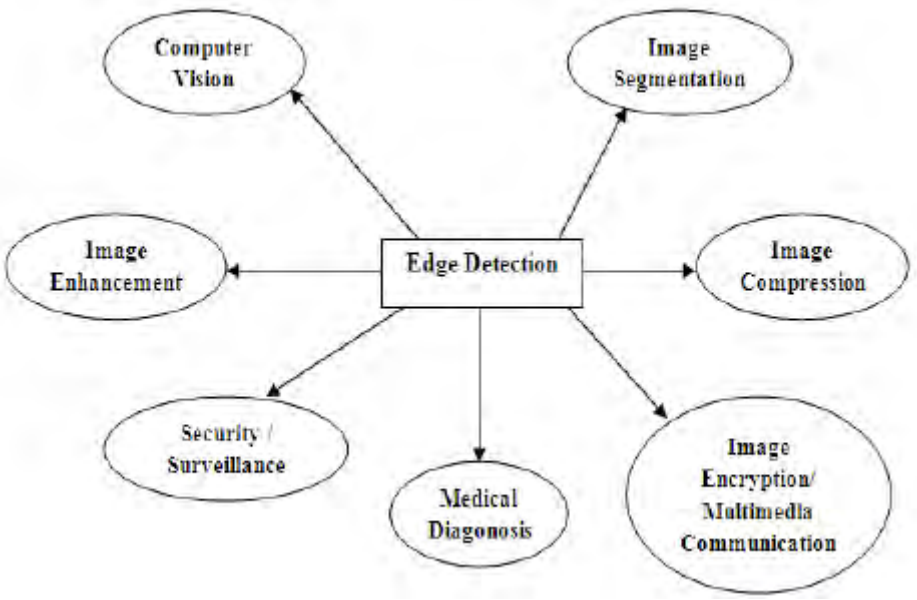


Figure 2.2: Applications of edge detection system

Generally, Edge detection contains three steps namely Filtering, Enhancement and Detection. Filtering is some major classical edge detectors work fine with high quality images, but often are not good enough for noisy pictures because they cannot distinguish edges of different significance. Noise is unpredictable contamination on the original image. There are various types of noise, but the most broadly studied two kinds are white noise and „salt and pepper" noise. During salt and pepper noise, pixels in the image are very different in color or intensity from their surrounding pixels; the defining characteristic is that the value of a noisy pixel bears no relation to the color of surrounding pixels. In general this type of noise will only affect a small number of image pixels. When analyzed, the image contains dark and white dots, hence the word salt and pepper noise.

Meanwhile, enhancement means digital image enhancement techniques are concerned with improving the quality of the digital image. The principal objective of enhancement techniques is to produce an image which is better and more suitable than the original image for a specific application. Linear filters have been used to solve many image enhancement problems. [6] Not all image sharpening problems can be satisfactorily addressed through the use of linear filters.

Lastly, detection consider as many points in an image have a nonzero value for the gradient, and all of these points are not considered as edge for a particular application. Some methods should be used to determine which points are edge points or not.

2.3 SIFT Features

A method called Scale Invariant Feature Transform (SIFT) [5] was introduced to extract distinctive invariant features from images that can be used to carry out reliable matching between various views of an object or a scene. The SIFT features utilize scale-space extreme in the difference-of-Gaussian (DoG) function convolved with the image as the key points which are invariant to image scaling. By assigning a consistent orientation for every key point depending on local image features, the key point descriptor can be represented in relation to this rotation and thus attain invariance to image rotation.

In addition to the invariance against image scaling and rotation, the SIFT features are also partially invariant against changes in lighting conditions and 3D camera viewpoint. The recognition is achieved by comparing individual features to a database of features from known objects using a fast nearest-neighbor method, followed by a Hough transform to discover clusters belonging to a single object, and finally executing verification by using least-squares solution for consistent pose parameters.

This recognition method can robustly recognize objects with cluttered backgrounds and occlusion, and the method can achieve real-time performance for low resolution images.[2] Therefore, the SIFT is adopted in this thesis for the bare hand gesture recognition. The four detection stages for SIFT features are:

i. Scale-space extrema detection.
ii. Keypoint localization.
iii. Orientation assignment.
iv. Generation of keypoint descriptors.

Detection stages for SIFT features were described as the following:

Interest points for SIFT features stand for local extreme of difference-of-Gaussian filters at different scales. Given a Gaussian-blurred image:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \qquad (2.1)$$

Where, $$G(x, y, \sigma) = \frac{1}{2\Pi\sigma^2} e^{-(x^2+y^2)} / 2\sigma^2 \qquad (2.2)$$

is a variable scale Gaussian, the result of convolving an image with a difference-of-Gaussian (DoG) filter: $G(x, y, k\sigma) - G(x, y, \sigma)$ is given by:

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \qquad (2.3)$$

Which there is the difference of the Gaussian-blurred images at scales σ and $k\sigma$.

The first stage to detect interest points is the convolution of the image with Gaussian filters at different scales, and the production of difference-of- Gaussian (DoG) images from the difference of adjacent blurred images.

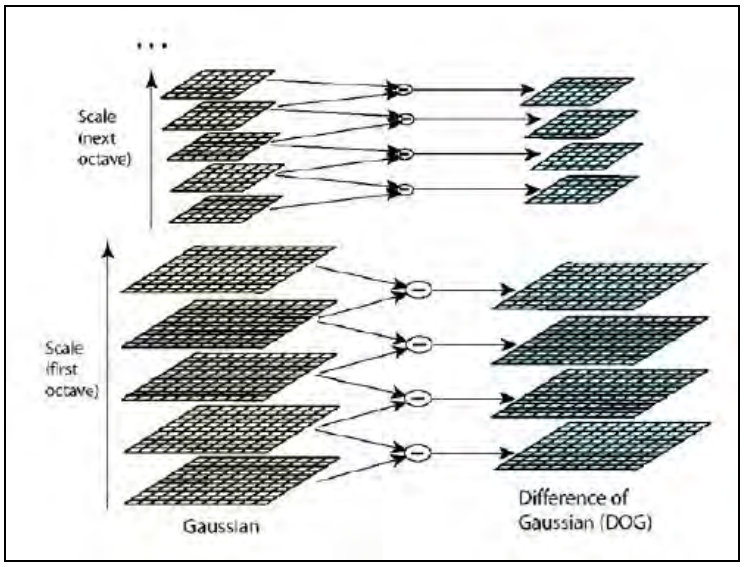


Figure 2.3.1: The blurred images at different scales and the computation of the difference-of-Gaussian (DoG) images.

An efficient approach to construction of $D(x, y, \sigma)$ is shown in Figure above. The convolved images were grouped by octaves (an octave stands for doubling the value of σ), and the value of it was chosen so that a fixed number of blurred images per octave was obtained. This also ensures that the same number of difference-of-Gaussian images per octave was obtained.

Note, the difference-of-Gaussian (DoG) filter gives an approximation to the scale- normalized Laplacian of Gaussian $\sigma^2 V^2$G. The difference-of-Gaussian filter is in effect a tunable band pass filter.

Interest points (called key points in the SIFT framework) are recognized as local maxima or minima of the difference-of-Gaussian (DoG) images across scales. Every pixel in the DoG images is compared to its 8 neighbors at the same scale, plus the 9 corresponding neighbors at neighboring scales. If the pixel is a local maximum or minimum, it is chosen as a candidate key point. For every candidate key point:

i. Interpolation of nearby data is used to determine its position accurately.
ii. Keypoints with low contrast are eliminated.
iii. Responses along edges are removed.
iv. The keypoint is assigned an orientation.

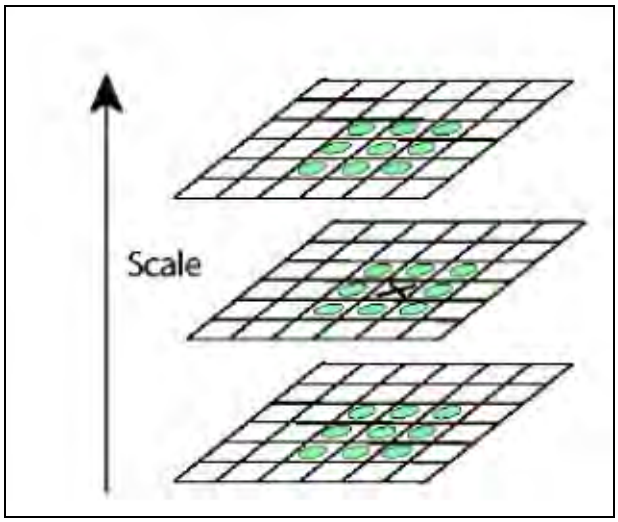


Figure 2.3.2: Local extreme detection, the pixel marked it is compared against its 26 neighbors.

Meanwhile, next figure, (Figure 2.4.3) illustrates the computation of the keypoint descriptor. In order to determine the keypoint orientation, a gradient orientation histogram is computed in the neighborhood of the keypoint (using the Gaussian image at the closest scale to the keypoint's scale). The contribution of each neighboring pixel is weighted by the gradient magnitude and a Gaussian window with a σ that is 1:5 times the scale of the keypoint.