UNIVERSITI TEKNIKAL MALAYSIA MELAKA

FAKULTI KEJURUTERAAN ELEKTRIK

FINAL YEAR PROJECT REPORT

PERFORMANCE COMPARISON OF OUT-OF-PLANE FACIAL DETECTION USING SPEEDED UP ROBUST FEATURES (SURF) AND SCALE INVARIANT FEATURE TRANSFORM (SIFT)

Nor Ain Zuzila Binti Zolkifly

Bachelor of Electrical Engineering

June 2015

SUPERVISOR ENDORSEMENT

"I hereby declare that I have read through this report entitle "Performance Comparison of Out-Of-Plane Facial Detection Using Speeded up Robust Features (SURF) and Scale Invariant Feature Transform (SIFT)" and found that it has comply the partial fulfillment for awarding the degree of Bachelor of Electrical Engineering (Control, Instrumentation & Automation)""

Signature	:	…….....
Supervisor"s Name	:	Madam Norhaslinda Binti Hasim
Date	:	24/06/2015

PERFORMANCE COMPARISON OF OUT-OF-PLANE FACIAL DETECTION USING SPEEDED UP ROBUST FEATURES (SURF) AND SCALE INVARIANT FEATURE TRANSFORM (SIFT)

NOR AIN ZUZILA BINTI ZOLKIFLY

A report submitted in partial fulfillment of the requirements for the degree of

Electrical Engineering, (Control, Instrumentation & Automation)

Faculty of Electrical Engineering

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2015

STUDENT DECLARATION

I declare that this report entitle "*Performance Comparison of Out-Of-Plane Facial Detection Using Speeded up Robust Features (SURF) and Scale Invariant Feature Transform (SIFT)*" is the result of my own research except as cited in the references. The report has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature	:	…….
Name	:	Nor Ain Zuzila Binti Zolkifly
Date	:	24/06/2015

To my beloved mother and father

ACKNOWLEDGEMENT

Alhamdulillah, thanks, to the Almighty Allah Most Gracious, Most Merciful for the blessing and grace, I have been able to implement undergraduate project which is Final Year Project (FYP). I would like to take this opportunity to express appreciation to all those who have helped me along the implementation of this project especially my supervisor, Madam Norhaslinda Binti Hasim on giving me a lot of advices, ideas and support me. My gratitude also goes to my fellow friends for all the helps and ideas provided during this project. Last but not least to my parents and family for their patience, understanding and support during the whole period of the project.

Thank you.

ABSTRACT

Nowadays, SURF and SIFT become a popular method for facial detection due to its advantages in detecting and recognizing of features. The aim of this research is to implement an algorithm for face detection using SURF and SIFT as well as to test an algorithm for face detection using SURF and SIFT. As SURF and SIFT has their own steps, the different algorithm is used to evaluate the performance of time and the number of feature point detection using SURF and SIFT. The research is started with implement the algorithm of both techniques. Then, the facial image has been captured by using camera with different pose that indicate the out-of plane rotation. After that, the facial image has been tested through MATLAB software. Finally, the performance comparison of out-of-plane facial detection using SURF and SIFT has been evaluated. In general, SURF and SIFT are scale invariant. SURF uses Haar wavelet and it is the fastest feature detector and descriptor. Meanwhile, SIFT uses Difference of Gaussian over various scales of an image and it is the most accurate feature detector and descriptor. It is expected that SURF is better in terms of speed, while SIFT is the best for accuracy. The parameters that have been evaluated from both techniques are time and the number of feature point detection. As a conclusion the performance of technique chosen between SURF and SIFT will be compared.

ABSTRAK

Pada masa kini, SURF dan SIFT telah menjadi kaedah yang popular untuk pengesanan muka berikutan daripada kelebihannya dalam mengesan dan mengenal pasti ciri-ciri muka. Tujuan kajian ini adalah untuk melaksanakan algoritma untuk mengesan muka menggunakan SURF dan SIFT serta untuk menguji algorithm untuk mengesan muka menggunakan SURF dan SIFT. SURF dan SIFT mempunyai langkah-langkah yang tersendiri, algoritma yang berbeza akan digunakan untuk menilai prestasi masa dan bilangan pengesanan titik ciri menggunakan SURF dan SIFT. Kajian ini akan bermula dengan melaksanakan algoritma bagi kedua-dua teknik. Kemudian, imej muka telah ditangkap menggunakan kamera dengan gaya yang berbeza yang menunjukkan imej keluar-putaran satah permukaan. Selepas itu, imej muka telah diuji melalui perisian MATLAB. Akhir sekali, perbandingan prestasi di luar putaran satah pengesanan muka yang menggunakan kaedah SURF dan SIFT telah dinilai. Secara umum, SURF dan SIFT merupakan skala yang tidak berubah. SURF menggunakan *Haar wavelet* dan ianya adalah pengesan ciri dan penghurai muka yang paling cepat. Sementara itu, SIFT menggunakan Perbezaan *Gaussian* yang lebih kepada skala imej yang pelbagai dan ianya adalah pengesan ciri dan penghurai yang paling tepat untuk imej. Hasil kajian dijangka bahawa SURF adalah lebih baik dari segi kelajuan, manakala SIFT adalah yang terbaik untuk ketepatan. Parameter yang telah dinilai dari kedua-dua teknik ini adalah masa dan bilangan pengesanan titik ciri. Kesimpulannya prestasi teknik yang dipilih antara SURF dan SIFT akan dibandingkan.

TABLE OF CONTENTS

CHAPTER	TITLE	PAGE
	SUPERVISOR ENDORSEMENT	**i**
	PROJECT TITLE	**ii**
	STUDENT DECLARATION	**iii**
	DEDICATION	**iv**
	ACKNOWLEDGEMENT	**v**
	ABSTRACT	**vi**
	ABSTRAK	**vii**
	TABLE OF CONTENTS	**viii**
	LIST OF TABLES	**x**
	LIST OF FIGURES	**xi**
	LIST OF ABBREVIATIONS	**xii**
1	**INTRODUCTION**	**1**
	1.1 Research background (Motivation and significance of research)	1
	1.2 Problem Statement	3
	1.3 Objectives	3
	1.4 Scope of work	3
	1.5 Expected project outcome	4
	1.6 Thesis outline	4

2	**LITERATURE REVIEW**	**5**
	2.1 Theory and basic principles	5
	2.1.1 Digital Image Processing	6
	2.1.2 MATLAB to Process an Image	7
	2.1.3 Face Detection	9
	2.1.4 SURF and SIFT for Facial Detection	11
	2.1.5 Formula for Facial Detection using SURF and SIFT	12
	2.2 Review of previous related works	13
	2.3 Summary and discussion of the review	14
3	**RESEARCH METHODOLOGY**	**16**
	3.1 Principles of the methods or techniques used in the previous work	16
	3.2 Overview	17
	3.3 Modelling and Simulation Implementation	18
	3.4 Data Collection	18
4	**RESULTS AND DISCUSSION**	**23**
	4.1 Simulation Result	23
	4.2 Performance Evaluation	30
5	**CONCLUSION AND RECOMMENDATION**	**33**
	REFERENCES	**34**
	APPENDIX A : Project Gantt chart and key milestones	**36**
	APPENDIX B : Sample for SURF technique	**37**
	APPENDIX C : Sample for SIFT technique	**46**

LIST OF TABLES

TABLE	TITLE	PAGE
2.1	Comparison of MATLAB and C Code for simple Matrix Operation	8
2.2	Summary of previous research	15
4.1	Sample 6 for SURF technique	24
4.2	Sample 6 using SIFT technique	25
4.3	Sample 5 using SIFT technique	27
4.4	Sample 5 using SURF technique	28

LIST OF FIGURES

FIGURE	TITLE	PAGE
1.1	Several frontal and in-plane rotated face detection results [3]	2
1.2	Several out-of-plane rotated face detection results [3]	2
2.1	Active pixel representation [8]	9
2.2 (a)	Cr values around the eye [8]	10
2.2 (b)	Cr values around the mouth region [8]	10
2.3	Box filters used by Fast Hessian as approximations to second order derivatives of Gaussians [2]	12
2.4	A 2x2 array of descriptor determined from an 8x8 set of samples [2]	13
3.1	Overview of the proposed facial detection method	18
3.2	The training image	19
3.3	The size of circles at different scales	20
3.4	The 2 × 2 subregions (the direction along the orientation of interest points) [12]	21
3.5	Feature point in diverse course	21
3.6	The influence in pixel‟s neighbourhood [12]	22
4.1	The results for performance in term of time of feature point detection	30
4.2	The results for performance in term of number of feature point detection	31

LIST OF ABBREVIATIONS

SURF	-	Speeded-Up Features Transform
SIFT	-	Scale-invariant Features Transform
SUSAN	-	Smallest Univalue Segment Assimilating Nucleus
3D	-	Three Dimensional
2D	-	Two Dimensional
ROI		regions of interest
CCTV	-	Closed Circuit Television
MATLAB	-	Matrix Laboratory
SNoW	-	Sparse Network of Winnows
DoG	-	Difference of Gaussians
RGB	-	Red Green Blue
FKE	-	Fakulti Kejuruteraan Elektrik

CHAPTER 1

INTRODUCTION

This chapter will give an overview of the project such as project introduction, project objective, project scope, project methodology and summary of this project. This chapter will explain briefly about the work from the beginning until the project is implemented.

1.1 Research background (Motivation and significance of research)

Facial detection becomes a quandary in the way to analyse the identity of a person. Nowadays, there are many technique and method that are utilized extensively for in-plane facial detection such as Harris corner, SUSAN, Hessian-Laplace and others [1]. As for Speeded-Up Features Transform (SURF) and Scale-invariant Features Transform (SIFT) [2] it is consider to be a popular method for facial detection.

This study focuses on out-of-plane facial detection. The method that will be used are SURF and SIFT while comparing their performance on out-of-plane facial detection. Based on experiment on out-of-plane facial detection, the parameter settings utilized in both algorithms are combined, adjusted, and developed in this system.

By referring to the journal written by (Yasir and Amelia, 2012). The point predicated delegacy of patch feature cause the out-of-plane detecting is possible happen. For more precise espial that represents the object of interest surface, the features must be mapped to a 3D model and then they can be rotated to the desired angle of sentiment. However, the features can be mapped to a flat plane for simplification of the out-of–plane rotation detection [3].

In the other hand, SURF algorithm and SIFT algorithm are including in representative feature extraction and matching algorithms. SIFT is invariant to scale, but with high computation involution SIFT has more preponderant performance in features number and precision. Furthermore it has modicum of calculation with great speed and the number of features is withal felicitous. SIFT has been widely applied and researched [4].

The Figure 1.1 shows the several of the normal and in-plane rotation detection results. Most of the researchers only focus on in-plane rotation and the scale invariance. However, when the situation has considerable out-of-plane rotation [5] as shown in Figure 1.2 it is become a problem to them. This will cause the accuracy to decrease. Hence, it is important to compare the two method between SURF and SIFT to show which is better in their performance of out-of-plane rotation facial detection.

Figure 1.1: Several frontal and in-plane rotated face detection results [3]

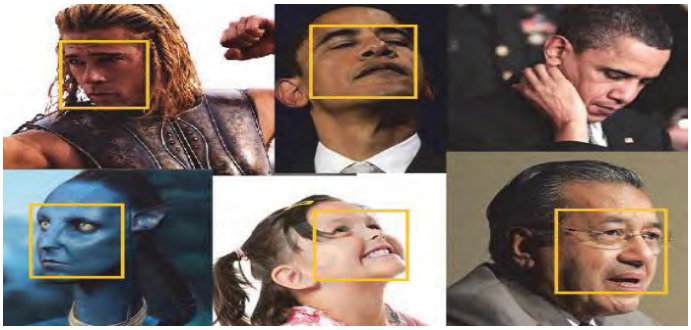

Figure 1.2: Several out-of-plane rotated face detection results [3]

1.2 Problem Statement

In 3D object position, the captured images which are coming from the CCTV can be directly matched to the out-of-plane rotated detection with the flat plane. Hence, face movement position detection has become a challenge in determining their image details. As the face position is uncontrollable, the desired powerful algorithm is highly required to determine the match of someone identity with different facial positioning. As the rotated images are normally harder to detect compared to static images, the level of accuracy in the rotated image is decrement [3]. Hence, the SURF and SIFT algorithms are developed in detecting an object of interest and their performance are compared.

1.3 Objectives

1. To implement an algorithm for face detection using SURF and SIFT
2. To test an algorithm for face detection using SURF and SIFT
3. To evaluate the performance of time and the number of feature point detection using SURF and SIFT

1.4 Scope of work

1. Focus on research about face detection in SURF and SIFT methods by using MATLAB 2013 programming.
2. To use image capture by using camera which is input is the known out-of-plane face image and output is feature point detecting result. Analyses the facial image by

comparing their performance of time of feature point detect and the number of feature point detect. The matching of feature point will not cover in this project.

1.5 Expected project outcome

The expected project outcomes of this project are:

1. New Finding / Knowledge
    a. Image detection algorithm with ability to accurately out-of-plane rotated image detection
2. Research Publication in conferences and journal.

1.6 Thesis outline

This thesis consists of five chapters that have been outlined in the table of content. Chapter 1 will cover on the introduction of the project. A bit of explanation will be done in order to complete this project. It also presents the readers with objectives, problem statements, scope of works, and the thesis outlines of the project. Chapter 2 discusses about literature review. This section reviews on previous researchers work about SURF and SIFT from Institute of Electrical and Electronics Engineers (IEEE) journal, articles, book, technical paper and others. Chapters 3 cover the methodology of the project and describe the flow chart of project activities. The result and discussions will be reviewed in Chapter 4. This chapter will highlight the initial results achieved from data collection and simulation. Finally, the conclusion and recommendations on future research will be enlightened in Chapter 5.

CHAPTER 2

LITERATURE REVIEW

The literature review is some sort of review discusses on information in a specified area and sometimes a detail on a particular topic. A literature review will be a simple summarizing from the references, however generally has an organizational pattern and includes both summarizing together with the synthesis. A synopsis has been usually a recap from the importance of the references, although the synthesis is known as a re-organization, of researcher‟s own idea. It might produce a brand new interpretation of past research material. (The Writing Center, 2012)

2.1 Theory and basic principles

In this section of this chapter, the literature review used in developing the facial detection will be discussed.

2.1.1 Digital Image Processing

A computer readable binary format consisting of logical 0s and 1s is converted from the pictures that are coming from digital images. Customarily, a video evolves with time and generally contains moving and/or transmuting objects, whereas by an image we mean a still picture that does not transmute with time. Digital format, albeit "direct digital"""" systems is converted from continuous signals are becoming more prevalent to obtain the digital images. Diverse exhibit media is utilized to view the digital images that included digital printers, computer monitors, and digital projection contrivances [6].

In the other hand, a digital image is a 2D signal in essence, and is the digital version of the 2D manifestation of the authentic-world 3D scene. Although the world"s picture and image are quite synonymous, the subsidiary distinction that „"picture"""" is the analogue version of „"image"""" is made. A function of two authentic variable is an image, for example, *a*(x,y) with *a* as the amplitude (e.g., effulgence) of the image at the authentic coordinate position (x,y). Sub images are considered contain an image, which is sometimes referred to as regions or regions of interest (ROIs). The amplitudes of a provided for picture will essentially dependably a chance to be whichever legitimate numbers or integers. The last will be generally those outcome of a quantization transform that proselytes a culprit extend (verbally express, the middle of 0 Furthermore 100%) with an discrete number about calibers [7].

2.1.2 MATLAB to Process an Image

An picture will be exactly a set about values sorted out in the manifestation of a grid. Since MATLAB is ideal for grid operations, it makes a plenty from claiming sense will use MATLAB for image-cognate operations. To a large portion of the preparing needs, those MATLAB‟s image transforming toolmaker (images) will be used. Other than authentic-time procurement about images, this toolmaker need every last one of necessary executes on perform sundry geometric, arithmetic, logical, and also different higher gauge transformations on the pictures. The toolmaker will be also fit for taking care of both colours Furthermore grayscale pictures. However, a large portion of the image transforming will be focused looking into a grayscale pictures. Considerably for the shade images, the preparing may be completed on the changed over image, which may be gotten by mapping those RGB colour space under grayscale space [7].

In spite of the fact that structural similarity may be an incredible inspiration on use MATLAB for picture processing, there would a few different purposes behind completing along these lines. Practically scientists in the region of image transforming use MATLAB Similarly as their principle stage for product implementation, which therefore provides for a common dialect with look at distinctive calculations furthermore outlines. Withal, those speeds about sundry calculations might moreover make compared once a common platform, which might make laborious if distinctive individuals were using distinctive modifying dialects that fluctuate extensively as far as pace about operations. Another intriguing reason with use MATLAB, the majority fascinating with the individuals building understudies who don't relish plenitude about coding, may be the curtness for code previously in MATLAB. Table 2.1 compares a percentage samples from claiming rudimental operations in MATLAB what's more C similarly as a reference. Particular case cam wood envision from the examination how sensational those distinction will make to involutes operations for example, such that convolution, filtering, Furthermore grid reversal. Convolution may be the heart from claiming basically every one of the sifting what's more time furthermore recurrence web-domain transformations done picture processing, what's more must make finished as speedy likewise conceivable. Those MATLAB capacity conv() may be withal a standout amongst the

well-kept privileged insights of MathWorks and is those heart about these operations. Those work need been optimized to the grid operations and, hence, operations to MATLAB ended up additional speedy and more effective over coding to different dialects. [7].

Table 2.1: Comparison of MATLAB and C Code for simple Matrix Operation [7]

Operation	Part of Code	MATLAB statements

Addition of two matrices A and B	for (i= =1, i<=M, i++)
	{
	for (j= =1, j<=N, j++)
	{
	D [i] [j] = A [i] [j] + B [i] [j] ;
	{
	}
		D = A + B ;

Multiplication of two matrices A and B	for (i= =1, i<=M, i++)
	{
	for (j= =1, j<=N, j++)
	{
	for (c= =1, c<=N, c++)
	{
	for (r= =1, r<=M, r++)
	{
	D [i] [j] +=A [i] [c]*B [j] [r] ;
	}
	}
	}
	}
		D = A * B ;

There is a restraint of the convenience about MATLAB. In spite of the fact that it will be wondrous to algorithm testing, it is not altogether lucky for ongoing imaging requisitions because of the gradualness about transforming. This gradualness emanates starting with a greater amount levels for accumulation Furthermore translation contrasted with alternate languages, and also from iterative methods the place loops are used every now and again over authentic-world requisitions. The best result might make on test those rationales Also calculations clinched alongside MATLAB initially since this will support the introductory

advancement pace. Once those algorithms may be finalized, it ought to a chance to be translated under C's et cetera aggregated under an executable document for authentic-time alternately close authentic-time requisitions.

2.1.3 Face Detection

Face identification will be performed predicated on the idea of meager organize for Winnows (SnoW). SnoW may be an taking in building design what's more will be used to consigning different classes. It may be concretely customized to enormously titanic scale taking in errands also for domains for which the number for characteristics taking piece in choices may be prodigiously what's more sizably voluminous. SnoW need been utilized prosperously on an assortment for sizably voluminous scale taking in errands clinched alongside domains similar to those common dialect processing, visual transforming furthermore numerous others. For routine SnoW predicated face identification the characteristics used to transfer need aid the animated pixel intensities inside the hopeful district. Pixel power toward every area (x, y) may be communicated done boolean qualities concerning illustration demonstrated over Figure 2.1. [8].

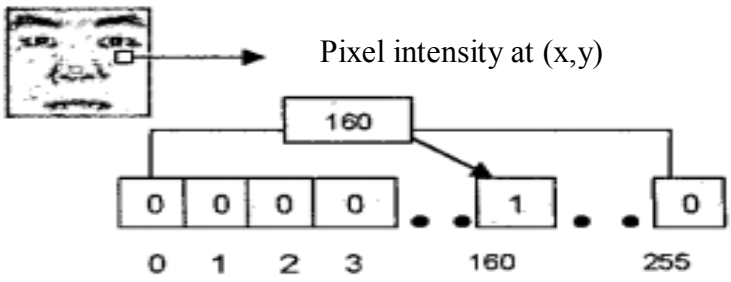

Figure 2.1: Active pixel representation [8].

2.1.3.1 Integrated skin extraction and face detection

Many colour spaces have been utilized in segmenting skin regions in colour images. Literature survey shows that the YcbCr colour space is one of the prosperous colour spaces in segmenting the skin colour accurately, mainly because the chrominance components are virtually independent of luminance component in the space. Although skin colour vary from person to person and they incline to get clustered into a compact region in CbCr space. So the skin colour distribution information in a human face can be utilized as a supplemental feature in relegating the objects into faces and non faces. Chrominance blue component is more prominent around the ocular perceivers in a human face and at the same time the Cr component is less. The chrominance red is more prominent in the mouth area. This information which is typical to human face is withal integrated as a feature in relegation process. This ameliorates the precision of the detection method in colour images at the same time reduces the mendacious positives at marginal increase in computational involution. Figure 2.2 shows the distribution of Cr pixels at the ocular perceiver region and the mouth region in a human face [8].

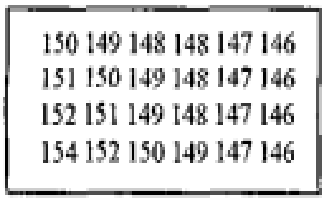

Figure 2.2 (a): Cr values around the eye [8]

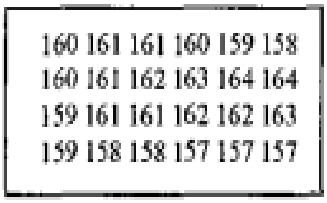

Figure 2.2(b): Cr values around the mouth region [8]

2.1.4 SURF and SIFT for facial detection

2.1.4.1 SURF Detector

SURF, additionally kenned as approximate SIFT, employs integral images and efficient scale space construction to engender keypoints and descriptors very efficiently. SURF uses two stages namely keypoint detection and keypoint description. In the first stage, rather than utilizing DoGs as in SIFT, integral images sanction the expeditious computation of approximate Laplacian of Gaussian images utilizing a box filter. The computational cost of applying the box filter is independent of the size of the filter because of the integral image representation. Determinants of the Hessian matrix are then used to detect the keypoints. So SURF builds its scale space by keeping the image size the same and varies the filter size only [9].

2.1.4.2 SIFT Detector

The SIFT detector has four main stages namely, scale-space extrema detection, keypoint localization, orientation computation and keypoint descriptor extraction. The first stage uses Difference of Gaussians (DoG) to identify the potential keypoints. Several Gaussian blurred images at different scales are engendered from the input image and DoGs are computed from neighbours in scale space. In the second stage, candidate keypoints are located by finding extrema in the DoG images that are locally extremal in space and scale. Spatially unstable keypoints are eliminated by thresholding against the ratio of eigenvalues of the Hessian matrix (unstable edge keypoints have a high ratio, and stable corner keypoints have a low ratio), low contrast keypoints are eliminated and the remaining keypoints are 11ocalized by interpolating across the DoG images. The third stage assigns a principal orientation to each keypoint. The final phase computes a highly distinctive descriptor for each keypoint [9].