UNIVERSITI TEKNIKAL MALAYSIA MELAKA

STUDY ON HAAR CLASSIFIER FOR MACHINE LEARNING SYSTEM THAT CAN DETECT SPECIFIED OBJECT IN VISION SYSTEM

This report submitted in accordance with requirement of the Universiti Teknikal Malaysia Melaka (UTeM) for the Bachelor Degree of Manufacturing Engineering (Robotic & Automation) with Honours.

By

CHAN HORNG TAT

FACULTY OF MANUFACTURING ENGINEERING

2009 / 2010

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

BORANG PENGESAHAN STATUS LAPORAN PROJEK SARJANA MUDA

TAJUK: **Study on Haar Classifier for Machine Learning System That Can Detect Specified Object in Vision System**

SESI PENGAJIAN: 2009/10 Semester 2

Saya **CHAN HORNG TAT**

mengaku membenarkan Laporan PSM ini disimpan di Perpustakaan Universiti Teknikal Malaysia Melaka (UTeM) dengan syarat-syarat kegunaan seperti berikut:

1. Laporan PSM adalah hak milik Universiti Teknikal Malaysia Melaka dan penulis.
2. Perpustakaan Universiti Teknikal Malaysia Melaka dibenarkan membuat salinan untuk tujuan pengajian sahaja dengan izin penulis.
3. Perpustakaan dibenarkan membuat salinan laporan PSM ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. **Sila tandakan (√)

☐	SULIT	(Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia 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:

__________________________ __________________________

Alamat Tetap:
45, JALAN HAJI ABDUL KARIM, 31
OFF JALAN SUNGAI JATI,
41200 KLANG,
SELANGOR

Cop Rasmi:

Tarikh: ____________________ Tarikh: ____________________

** Jika Laporan PSM ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa/organisasi berkenaan dengan menyatakan sekali sebab dan tempoh laporan PSM ini perlu dikelaskan sebagai SULIT atau TERHAD.

DECLARATION

I hereby, declared this entitled „Object detection system‟ is the results of my own research except as cited in references.

Signature	:	___________________
Author‟s name	:	CHAN HORNG TAT
Date	:	5 May 2010

APPROVAL

This report is submitted to the Faculty of Manufacturing Engineering of UTeM as a partial fulfillment of the requirements for the Degree of Bachelor of Manufacturing Engineering (Robotic and Automation) with Honors. The member of the supervisory committee is as follow.

………………………………

Supervisor	:	Mr Shariman Bin Abdullah
Date	:	5 May 2010
Stamp	:	

ABSTRACT

Nowadays, visual system is widely applied in the manufacturing process. However, implementation of visual system is highly cost due to expensive camera and software is needed. Therefore, a low cost object detection system with webcam is shown. The main purpose on this project is to create a Haar classifier that able to detect specified object. Haar classifier is one type of detection system application that apply boosted rejection cascade. It has applied decision tree with boosting theory in order to have better performance and high percentage of detection. 2 main categories that related to this topic are machine learning and machine vision. Haar classifier can be consider as a software that able to learn and differentiate the difference in a still image and it also can say as have vision to "see" and differentiate the images. The Haar classifier that is used in this project is based on Viola and Jones theory. Objects used for detection system are "T" shape object and transparent mouse to differential the effect of complexity object in the performance of system.

ABSTRAK

Baru-baru ini, kilang – kilang banyak menggunakan sistem pengesanan dalam proses kilang. Tetapi penggunaan sistem pengesanan melibatkan alat-alat yang berharga tinggi. Dalam projek ini, satu sistem pengesanan yang berharga rendah ditujukan. Tujuan Projek ini adalah untuk menciptakan satu sistem pengesanan untuk mencari benda yang dikehendaki daripada satu gambar. “Haar Classifier” adalah sejenis sistem pengesanan yang bergantung pada penolakan bahan yang tidak diperlukan. Dalam projek ini, 2 aspek yang perlu diketahui adalah “machine vision” dan “machine learning”. “Haar classifier” dapat dianggap sebagai satu perisian komputer yang berkebolehan untuk belajar dari gambar-gambar dan mempunyai “mata” untuk melihat dan membezakan benda yang berlainan. Teori bagi projek ini adalah berasal dari “Viola & Jones” yang menggunakan “AdaBoost” untuk menciptakan satu “classifier” yang mempunyai pengesanan yang tinggi tetapi penolakan yang rendah. Benda digunakan untuk sistem pengesan adalah benda berbentuk “T” dan “Mouse” yang berlutsinar.

DEDICATION

This report is dedicated to my parents of blessed memory, who raise me to be a responsible and careful person. Other than that, I would like to thank them for supporting me in this project.

ACKNOWLEDGEMENT

I would like to take this opportunity to express my highest gratitude and appreciation to my supervisor, Mr. Shariman for continuing supports, supervises encouragement, patience and constructive opinions throughout this project.

Other than that, I am grateful for the cooperation and guidance which are given by my friends. I would like to thank them for their willingly support and answer all my question with detail explanation. It does help me a lot in my project.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

BBN	–	Bayesian Belief Networks
BL	–	Bayesian Learning
BMP	–	Bitmap
DT	–	Decision Trees
GA & GP	–	Genetic Algorithms and Genetic Programming
MB	–	Megabyte
ML	–	Machine Learning
MLL	–	Machine Learning Library
NN	–	Neural Networks
OCR	–	Optical Character Recognition
OPENCV	–	Open Source Computer Vision Library
PC	–	Personal Computer
PLC	–	Programmable Logic Controller
PSM	–	Final Year Project
SVM	–	Support Vector Machines

CHAPTER 1

INTRODUCTION

Nowadays, artificial intelligence is getting popular among us. The machine is always being wonder whether it is possible to obtain intelligence like human. A lot of research is running in order for machine to obtain the intelligence of human. One of human intelligence ability is detect and differentiate objects. Therefore, in this project, an object detection based on Haar classifier is discussed. Haar classifier is one of the machines learning approach that can be used for object detection and it is based on the boosted rejection cascade to eliminate the unnecessary. This method has been widely using in various applications such as face detection. Thus, in this project, the method to apply Haar classifier for object detection will be clearly discussed and this method is involves machine learning techniques which require training the computer on the specified object sample image and then test the machine learning with another set of test set to verified the object.

1.1 Objective

- Study on Haar classifier for machine learning system that can detect specified object in vision system
- Study on the method to use OpenCV in creates sample and Haar training.
- Find image database that contain positive and negative image of specified object.
- Create a own object detection system
- Increase the percentage of detect specified object

1.2 Problem Statement

For detection system, face detection is easily and commonly found but object detection for specified object is not easily obtained from anywhere. Due to the detection system is specified on certain object, the system has to be specially made. Other than that, for an object detection system, it usually has a low percentage in detect specified object with small amount of sample images. It requires many sample images in order to increase the accuracy of the system. Other than that, as in production line in industry, detection system is always needed to check for correct orientation of object and condition of object.

1.3 Scope

- Applying not more than 500 sample positive image for training Haar classifier
- Object detection system is used to detect specified object in a still image.
- Images generated by CreateSample utility on OpenCV is not more than 5000.

1.4 Result Expected

The object detection system is manages to differentiate specified object in a still image. The system is able to detect more than 1 object if there are 2 specified objects in a still image. Other than that, the system will either highlight or circle the specified object that wanted.

1.5 Gantt Chart

	PSM I (Weeks)														PSM II (Weeks)													
Tasks	1	2	3	4	5	6	7	8	9	10	11	12	13	14	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Project determination	■	■																										
Review topic		■	■	■																								
Literature review					■	■	■	■	■	■																		
Methodology									■	■	■	■																
Report Submission												■	■															
Preparation & Presentation													■	■														
Project Planning															■	■												
Data collection																■	■	■										
Training Classifier																	■	■	■	■	■	■	■	■	■			
Testing																					■	■	■	■	■	■		
Modification																							■	■	■	■		
Report Submission																										■	■	
Preparation & Presentation																											■	■

Figure 1.1: Gantt chart PSM I and PSM II

CHAPTER 2

LITERATURE REVIEW

The main focus on this project is about the Study on Haar classifier for machine learning system that can detect specified object in vision system. Machine vision is the application of computer vision to industry and manufacturing, whereas machine learning is a scientific discipline that is concerned with the machine or computer able to learn based on data, such as from sensor data or databases. In this section, it included the detail description of some aspects that is related to the main focus such machine vision, machine learning, object detection, object recognition, and Haar classifier.

2.1 Machine Vision

To define machine vision, human vision must be known first. Human vision is one of the five human senses, and it is the one that carries the richest information content. It also refers as a response of the eye and brain to light. Anything that responds to and processes light can be thought of as having vision. Other than that, humans rely on vision for almost all aspects of living, including particularly locomotion, searching for sustenance, and manufacturing myriad edifices, devices, and food products [Roy Davies, 2003].

Machine vision which is known as computer vision is the scientific discipline whereby meaningful and explicit descriptions of physical objects from the world around are constructed by images. In other word, it is the artificial response of a device to spectral radiation, such as using the simplest devices, photocells. Other than that, it can produces abstractions or measurements from geometrical properties and comprises techniques for estimating features in images, relating feature measurements to the geometry of objects in space, and interpreting this geometric information. This overall task is generally called as image understanding [David. Kosiba, 2006].

The objective of a machine vision system is to create a model of the real world from images or sequences of images. Since images are two-dimensional projections of the three-dimensional world, the information is not directly available and must be recovered. This recovery requires the inversion of a many-to-one mapping therefore to reclaim this information, the knowledge about the objects in the scene and projection geometry is required. Besides that, at every stage in a machine vision system, knowledge of the application is requiring to make the decisions. Therefore, the important point in machine vision systems is on maximizing the automatic operation at each stage by using knowledge. The knowledge used by the system can includes models of features, image formation, objects, and relationships among objects. Without use of knowledge, machine vision systems can only for limited applications. Therefore, to provide more flexibility and robustness, knowledge is represented explicitly and is directly used by the system. Knowledge is also used by the designers of machine vision systems in many implicit as well as explicit forms. In fact, the efficacy and efficiency of a system is usually governed by the quality of the knowledge used by the system. Difficult problems are often solvable only by identifying the proper source of knowledge and appropriate mechanisms to use it in the system [David. Kosiba, 2006].

2.1.1 Fundamentals of Machine Vision

In order to design and develop successful machine vision systems, a clear and full understanding of all aspect of the system must be established. The processes of machine vision from initial image formation to final scene interpretation must be clearly understood. In this section, process involved in machine vision will be briefly explained. [David. Kosiba, 2006]

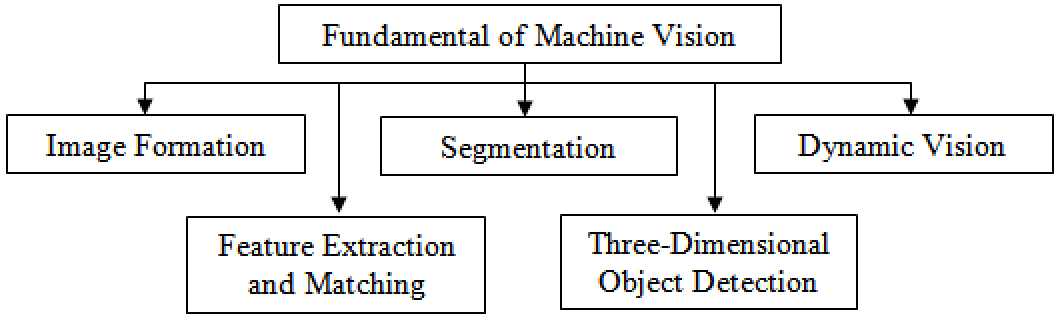

Figure 2.1: Fundamentals of Vision

2.1.1.1 Image Formation

When a sensor records received radiation as a two-dimensional function, an image is formed. The brightness or intensity values in an image can be representing different physical entities. As an example, in a typical image obtained by a video camera, the intensity values represent the reflectance of light from various object surfaces in the scene, whereas, in a thermal image, they represent the temperature of corresponding regions in the scene. Other than that, multiple images of the same scene are often captured using different types of sensors to facilitate more robust and reliable interpretation of the scene. Therefore selecting an appropriate image formation system plays a key role in the design of practical machine vision systems. [David. Kosiba, 2006]

2.2.1.2 Segmentation

Before abstract representations and descriptions can be generated, an image must be analyzed and its relevant features must be extracted. The success of higher level scene interpretation algorithms is strongly depending on the selection of these so-called low-level operations. One of the first operations that a machine vision system must be performed is the separation of objects from the background. This operation is commonly called as segmentation. It can be approached two ways, such as

a) Edge-based method – the boundaries of objects are used to partition an image and it emphasis o the detection the dissimilarities in the neighborhoods of point.
b) Region-based method – it is based on grouping of pixels according to certain similarities

2.1.1.3 Feature Extraction and Matching

Segmented images are often represented in a compact form to facilitate further abstraction. Methods used for object recognition and description is always match the representation schemes that been chosen. Therefore, the task of object recognition is requires matching an object description in an image to models of known objects. Since object recognition requires matching process, it is shown that matching plays an important role in other aspects of information recovery from images and there are few type of matching, such as pattern point matching and template matching. [David. Kosiba, 2006]

2.1.1.4 Three-Dimensional Object Recognition

In real world, everything is composed of three-dimensional solid objects. When an object is viewed for the first time, human typically gather information about that object from many different viewpoints. This process of gathering detailed object information and storing that information by human is referred to as model formation. Once human is familiar with many objects, the objects can be identified from an arbitrary viewpoint without further investigation. Furthermore, human now is able to identify, locate, and qualitatively describe the orientation of objects in black-and-white photographs. This basic capability is significant to machine vision because it involves the spatial variation of a single parameter within a framed rectangular region that corresponds to a fixed, single view of the world. [David. Kosiba, 2006]

2.1.1.5 Dynamic Vision

At the beginning, machine vision systems were only concerned with static scenes but the world is dynamic. Therefore there is a need for designing machine vision systems that capable of analyzing dynamic scenes. For a machine vision system engaged in the performance of nontrivial real-world operations and tasks, the ability to cope with moving and changing objects and viewpoints is critical. The input to a dynamic-scene analysis system is a sequence of image frames and the camera used to acquire the image sequence may itself be in motion. Therefore, the changes in a scene may be due to camera motion, object motion, illumination changes, or changes in object structure, size, or shape. However, changes in a scene are usually assumed to be due to camera and/or object motion. [David. Kosiba, 2006]

A scene usually contains several objects. An image of the scene at a given time represents a projection of part of the scene; the part of the scene depends on the position of the camera. Four cases represent the possibilities for the dynamic-camera/world setup:

a) Stationary camera, stationary objects (SCSO)
b) Stationary camera, moving objects (SCMO)
c) Moving camera, stationary objects (MCSO)
d) Moving camera, moving objects (MCMO)