UNIVERSITI PUTRA MALAYSIA

MOTION TRACKING WITH WEB CAMERA IMAGES BASED ON SPATIAL PROPERTIES

ANDREW LEE CHOON GUAN

FK 2003 21

MOTION TRACKING WITH WEB CAMERA IMAGES BASED ON SPATIAL PROPERTIES

By

ANDREW LEE CHOON GUAN

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirement for the Degree of Master of Science

May 2003

UPM

Dedicated to my family,

Hwee Chin
John Lee
Jane Lee

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in partial fulfilment of the requirements for the degree of Master of Science

MOTION TRACKING WITH WEB CAMERA IMAGES BASED ON SPATIAL PROPERTIES

By

ANDREW LEE CHOON GUAN

May 2002

Chairman: **Abdul Rahman Ramli, Ph.D.**

Faculty: **Engineering**

Machine Vision provides a cheap and flexible mean of tracking objects in motion when implemented by a web camera. The low resolution digital images, capturing the different instances of the scene of object in motion yields information which can be used to lay a historical track of the object.

The implementation of such a system involved the separation of the objects from the background using threshold segmentation techniques. Although it accepted the variation of natural lighting, it assumed that the background was lighter than the objects. By that method, the objects which have the potential to move, were separated from the stationary background.

The segmentation scheme implemented was a robust automated scheme, and form the preprocessing stage of the whole system.

A difference picture technique that detects changes between two consecutive images was implemented to separate the objects which moved from the stationary objects. The technique's inherent problem is the generation of too much errors due to signal problems and changing illumination. In order to track motion a noise elimination process was required to filter the difference image.

The historical positions of the moving object, was marked relative to the image pixel frame work. It involved pointing the object position to an image pixel within the object image. This involved the comparison between the difference picture and the primitive and a reduction scheme to reduce the object pixels to one which is within the object.

When multiple objects were in motion, an object labeling scheme was done to distinguish each object uniquely so that they can be correlated correctly in between images. The *RGB* properties was used as the object's unique property to label the moving objects. If the objects are similar in color shape and size, and are in close proximity to each other, the tracking might yield errors. This problem is similar to any tracking systems, include the human eyes.

The tracking record of the objects in motion was by linking the historical positions of the moving objects. The linking track was done by extrapolating the positions in different frames of image by a straight line. The track was superimposed upon the last frame of image captured, forming a compact record of object tracking.

A tracking system was designed and implemented using the web camera. It was able to track more than one object moving at the same time. The tracks of each of the different colored objects were correctly correlated.

The web camera is capable to run machine vision application as seen here in the task of object tracking. If the limitation of its resolution and the capture speed is of no consequence to the output, the web camera can be used in other machine vision applications.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia untuk memenuhi sebahagian keperluan untuk ijazah Master Sains

PENGESANAN PENJEJAKAN MENGGUNAKAN IMEJ KAMERA WEB MELALUI CIRI TEMPATAN

Oleh

ANDREW LEE CHOON GUAN

May 2003

Pengerusi: **Hj. Abdul Rahman Ramli, Ph.D**

Fakulti : **Kejuruteraan**

Penglihatan mesin membekalkan suatu cara yang murah dan fleksibel untuk menjejak objek yang sedang bergerak apabila digunakan dengan kamera web. Resolusi rendah imej digital, menangkap imej objek dalam gerakan pada masa yang berbeza dapat menghasilkan maklumat yang boleh digunakan untuk membina suatu ejakan masa untuk objek tersebut.

Pelaksanaan sistem sebegini melibatkan pemisahan objek daripada latar belakang dengan menggunakan teknik bahagian bendul. Walaupun ia

belakang lebih terang daripada objek. Dengan kaedah itu, objek yang ada potensi untuk bergerak dipisahkan latar belakang yang tidak bergerak.

Skim pembahagian yang dilaksanakan itu adalah suatu skim automatik yang tegap, dan membentuk peringkat pra-pemprosesan untuk seluruh sistem.

Teknik imej yang berbeza digunakan untuk mengesan perubahan telah dilaksanakan untuk memisahkan objek yang bergerak daripada objek yang tidak bergerak. Teknik ini menghasilkan banyak kesalahan kerana masalah signal dan keterangan yang berubah. Untuk mengesan pergerakan suatu proses penyingkiran kebisingan telah dilaksanakan.

Kedudukan objek bergerak ditandakan berbanding dengan kerangka piksel imej. Ia melibatkan petunjuk kedudukan objek ke suatu piksel imej dalam imej objek. Ini melibatkan perbandingan antara imej berbeza dengan imej pertama. Skema penurunan untuk mengurangkan piksel objek kepada yang dalam lingkungan objek tersebut dilaksanakan.

Ketika pelbagai objek bergerak, suatu skim penglabelan objek dibuat untuk membezakan setiap objek secara unik supaya objet-objet ini dapat dihubungkan dengan betul antara imej. Maklumat RGB digunakan sebagai maklumat unik objek untuk menglabelkan objek bergerak. Apabila ciri-ciri objek adalah sama, dan objek-objek berdekatan antara satu sama lain,

masalah mungkin timbul dalam penjejakan. Ini adalah sama dengan sistem lain termasuk penglihatan manusia.

Rekod penjejakan objek dalam gerakan dibuat dengan menghubungkan kedudukan mengikut perbezaan masa objek bergerak. Hubungan penjejakan dibuat dengan mengekstrapolasi kedudukan bingkai gambaran yang berbeza dengan suatu barisan lurus. Penjejakan tersebut ditindih atas bingkai imej yang terakhir yang akan membentuk suatu rekod padat penjejakan objek.

Suatu sistem penjejakan direkabentuk dan dilaksanakan menggunakan kamera web. Ia dapat menjejaki lebih daripada sebuah objek yang bergerak pada masa yang sama. Korelasi untuk setiap jejak objek berwarna yang berbeza itu dilakukan dengan betul.

Kamera web berupaya untuk menjalankan aplikasi penglihatan mesin sebagaimana yang boleh dilihat dalam aktiviti penjejakan objek. Jika had resolusinva dan kelajuan tangkapan tidak mempengaruhi keluaran, kamera web itu dapat digunakan di dalam aplikasi penglihatan mesin yang lain.

ACKNOWLEDGEMENTS

I would like to express my deep gratitude to my supervisor, Dr. Abdul Rahman Ramli, for his guidance. His motivation, ideas and encouraging concern had been very important to me juggling a full-time work and doing the rather intensive course.

I am very thankful to my supervisory committee, Puan Roslizah binti Ali and En. Syed Abdul Rahman for giving me their valuable time.

I appreciate with deep gratitude all my fellow researchers in the Department of Computer , Communication and System Engineering, as well as all the helpful staff.

UPM

TABLE OF CONTENTS

		Page
DEDICATION		ii
ABSTRACT		iii
ABSTRAK		vi
ACKNOWLEDGEMENTS		ix
APPROVAL		x
DECLARATION		xi
LIST OF FIGURES		xvi
LIST OF TABLES		xvii
LIST OF SYMBOLS AND ABBREVIATIONS		xviii
CHAPTER		
1	INTRODUCTION	1
	1.1 Objectives of the thesis	2
	1.2 Thesis organization	4
2	LITERATURE REVIEW	6
	2.1 Digital Imaging	7
	2.1.1 The Digital Hardware	7
	2.1.2 The Image	8
	2.1.3 The Single Frame Image	10
	2.1.4 Successive Frames of Dynamic Image	11
	2.2 Machine Vision	11
	2.2.1 Motion Tracking	12
	2.2.2 The Machine Vision Tracking Solutions	14
	2.2.2.1 Detection and Long Term Tracking of Moving Objects in Aerial Video	14
	2.2.2.2 Robust Multiple Car Tracking with Occlusion Reasoning	15
	2.2.2.3 Optical Flow Based Realtime Object Tracking	15
	2.2.3 Camera verses Target Dynamic Interplay	15
	2.3 Modular Design of Target Acquisition and Tracking	16
	2.3.1 Change Detection	16
	2.3.2 Time-Varying Edge Detection	19
	2.3.3 Difference Picture.	19
	2.3.4 Correlation for Multiple Objects Tracking	22
	2.4 Conclusion	24

3. BACKGROUND THEORY
3.0 Introduction 26
3.1 Segmentation of Objects to Background 26
3.1.1 The P-tile method. 28
3.1.2 The Mode method 28
3.1.3 Iterative Thresholding 30
3.1.4 Adaptive Thresholding 31
3.1.5 Variable Thresholding 31
3.1.6 Double Thresholding 32
3.1.7 Thresholding Selection Using Gradient Operator 33
3.2 Separation of Moving Objects from Stationary Objects 34
3.2.1 Detection of Change between Two Images 35
3.3 Morphological Operator 36
3.3.1 Noise Subtraction 36
3.4 Object Tagging for Single Object in Motion 41
3.4.1 Intermediate Image Union 43
3.5 Motion Path Tracing 44
3.6 Separating Moving Object from Other Moving Objects 45
3.7 Conclusion 46

4. METHODOLOGY 47
4.1 Overview 47
4.2 Different Automated Segmentation Techniques 51
4.2.1 The P-Tile Technique 51
4.2.2 The Mode Technique 52
4.2.3 Iterative Thresholding Technique 53
4.2.4 Adaptive Thresholding Technique 54
4.2.5 Double Thresholding Technique 54
4.3 Image Subtraction Technique 55
4.4 Technique to Remove Noises 56
4.5 Motion Detection 56
4.5.1 Object Marking for Single Object in Motion 57
4.5.2 Intermediate Image Union 57
4.5.3 Object Location 59
4.5.4 Labeling of Multiple Objects 59
5.5.4.1 Normalized *RGB* Properties 60
5.5.4.2 Labeling and Correlating the Objects 61
4.6 Motion Path Marking 62

5. RESULTS AND DISCUSSION 69
5.1 Result of Segmentation 64
5.1.1 Simple Global Thresholding 64
5.1.2 Automatic Thresholding – P-tile method 67
5.1.3 Automatic Thresholding –The Mode Method 68
5.1.4 Iterative Thresholding 69
5.1.5 Adaptive Thresholding 70
5.1.6 Double Thresholding 71
5.1.7 Discussion about the Segmentation Techniques 72
5.2 Image Subtraction 72
5.3 Noise Suppression 74
5.4 Object Marking for Single Object in Motion 75
5.5 Locating of Object Position 77
5.5.1 Reducing Location to a Single Pixel Coordinate 78
5.6 Connecting $I_1(x,y)$ with $I_2(x,y)$, to form the Track Path 78
5.7 Multiple Objects Labeling and Association using *RGB* properties 80
5.8 Multiple Objects Location and Morphing of Multiple Objects Co-ordinates 82

7. CONCLUSION 83

REFERENCES 86

VITA 88

LIST OF FIGURES

Figure		Page
3.1	Histogram of pixels and gray intensity value	29
3.2	Image of an object (1) in a background	37
3.3	Dilation process	38
3.4	Erosion process	38
3.5	Filling of holes by 4-neighbourhood condition	39
3.6	Erosion using the 4-neighbourhood condition	40
3.7	Registering the object position	41
3.8	Images of the same scene in two different times	42
3.9	Difference image	42
3.10	Intermediate Image Union	43
3.11	Results after the image union	44
3.12	Track being drawn to show the historical positions	45
4.1	The experimental setup	47
4.2	Structure of the processes	49
4.3	Flow Chart of the preliminary processes	50
4.4	Image subtraction with noise removing	56
4.5	The 'AND' operator	58
5.1	Bitmap(144x176) Image Im_1	64
5.2	Image Im_2	65
5.3	Histogram of gray level of image Im_1	65
5.4	Histogram of gray level of image Im_2	66
5.5	Binary images showing results of segmentation	67
5.6	P-tile Binary images	68
5.7	Binary image produced by the Mode method	69
5.8	Histogram showing the Mode Technique	70
5.9	Binary image produced by the Iterative method	71
5.10	Image produced by the Adaptive method	71
5.11	Binary image produced by the Double thresholding	72
5.21	Segmented I_1	74
5.22	Segmented I_2	74
5.23	Difference image	75
5.24	Difference image	76
5.25	Result of the 'AND' operator	77
5.26	Object in image	78
5.27	Reduction of object to Single Pixel	78
5.28	The track of the object in a single displacement	79
5.29	The track of the moving object through four positions.	80
5.31	Two objects moving through three different positions	82

LIST OF TABLES

Table		Page
5.1	Comparison of results of the P-tile method with different % of object pixels assumed	68
5.2	Comparison of the various segmentation techniques	73
5.3	This shows the error between the marked positions and the actual center of images I_1 and I_2	79
5.4	Show the comparison of *RGB* properties and the normalized *RGB*	81

LIST OF SYMBOLS AND ABBREVIATIONS

AADP	:	Absolute Accumulative Difference Picture
ADP	:	Accumulative Difference Picture
AND	:	Logic Operator
CD	:	Compact Discs
CCD	:	Charged Couple Devices
DP	:	Difference Picture
$D(x,y)$	:	Binary Difference Image
GPS	:	Global Positioning System
$I(x,y)$	:	Binary Segmented Image
$Im(x,y)$	:	*RGB* Image Captured by Web Camera
imshow	:	Matlab@tm Function to show image file
imread	:	Matlab@tm Function to read image file
imhist	:	Matlab@tm Function to plot histogram of image
MCSO	:	Moving Camera, Stationary Object
MCMO	:	Moving Camera, Moving Object
NDP	:	Negative Difference Picture
PADP	:	Positive Accumulative Difference Picture
RGB	:	Red, Blue, and Green Values of Image
PDP	:	Positive Difference Picture
SCSO	:	Stationary Camera, Stationary Object
SCMO	:	Stationary Camera, Moving Object

CHAPTER 1

INTRODUCTION

Moving objects tracked using aerial video (Walter *et al* 1999) can be applied to both military uses in target acquisitions systems. Translating this to the local scene, video stream from satellite can be used to manage ships passing through the Straits of Malacca. Although such long distance tracking requires expensive state-of-the-art high resolution optical hardware, the solutions done by Walter using low end hardware are similar.

The robust multiple car tracking system done in Berkley (Wiklund *et al* 2001) can be used for the traffic management system as well as in the underground car parking management. The traffic light system could response to the situation, allowing for maximum flow of vehicles.

The tracking of people in real time (Ryozo *et al* 1995) can be used as security surveillance for a secured area. Applying to industrial management, the work flow of a factory can be studied and improved upon by tracking the movements of the workers. Similarly the buying habits of the customer in a supermarket can be tracked using the video information.

This thesis records an implementation of a low cost web camera in an object tracking solution by applying progressive separating algorithms. The attempt to use a short range and low resolution web camera in such a machine vision application, was because it is convenient and a cheap way to generate digitized data of a scene.

It will be shown that the algorithms applied and tested in the web camera images can be applicable to the high-end equipments (high resolutions) because the fundamental image format captured are the same.

Motion Detection involves the ability to detect the objects that moved and to tag them. Motion Tracking includes the marking of the path of temporal positions the object moved. This historical trace can be recorded in various forms, for example the coordinate readout. It can also be a graphical track which is drawn connecting the various positions of the object.

The implementation was done using the web camera to capture the various positions of round colored buttons on a white board. The track of motion is the graphic line linking the various historical positions.

1.1 Objectives of the thesis.

The simulation of motion tracking by the web camera must;

- Be automated without the need to tweak the values of the perimeters
- Lay a historical track of the object in motion
- Lay separate tracks when multiple objects are in motion

The images captured at two different times represent the time element of the situation. Object was deemed to have moved when the position of the object are different

between the two images captured. The comparison of two images quantitatively is called the different image technique.

Difference image technique applied alone is inadequate to track motion. The causes of the changes between two frames of capture could be due to small insignificant motion, the change in illumination and even the digital capture technology itself. A simple difference image technique will pick up all these noises making the tracking of a single moving object (big and moving significantly) not possible.

The research aim to;

- solve the problem of the unwanted changes (or noises) using dilation and erosion techniques
- implementation of automated separation schemes progressively
    - narrowing the artifacts in a digitized image of the scene to objects and background using segmentation schemes
    - differentiation of moving objects from non-moving objects using spatial properties of the image
- includes the correlation of each of the moving objects with itself through various time of image capture
- draw multiple separable tracks for each of the different moving objects

- link all the solutions to one automated scheme of tracking using web camera images

Allegorically it can be thought of as a funneling of many possibilities to one, or an inverted pyramidal hierarchy of reduction of choices.

1.2 Thesis Organization.

This thesis proposes the use of web camera for the capture of images and processed by personal computer using MATLAB@tm codes as platform, to track object in motion.

Research background to the hardware, computer imaging and algorithms which forms the whole process of tracking of objects in motion is from chapter 2 to chapter 4.

Chapter 2 deals with broad aspect of work done in target acquisition and tracking by other researchers, highlighting specific differences that may be used to catalogue other works in the same general heading. It includes the principle of digital imaging, covering hardware and software, single image capture and dynamic capture of many frames. It distills to a modular approach of target acquisition and tracking.

Chapters 3 deals in detail the algorithms suggested in the modular approach, highlighting problems and possible solutions.

Chapter 4 describe the methodology to be implemented for the web camera for the object tracking.

Chapter 5, are results of experiments done, and discussion on the applied solutions and results to the algorithms dealt in chapters 3 and 4. The results published were from tests done both with images residing in the MATLAB$_{@tm}$ library, as well as from fresh capture of images through the web camera.

Chapter 6, summarizes the success and difficulties of the research, and suggests some application to the model as well as future areas of possible research.

CHAPTER 2

LITERATURE REVIEW

Machine vision, the extension of digital imaging, is the inclusion of intelligent interpretation to image processing. It is the application of selected programmable mathematical algorithms on digitized images. It has been incorporated into many uses, for example, object recognition, pattern recognition, intrusion detection and motion tracking. Some research on machine vision were published as early as 1970's, but there were an escalation of research done during the 1980's. (Jain *et al* 1995) It was by no accident that it was also during that time that computers were made smaller, faster, cheaper and more portable. Computing power is the backbone of machine vision, without which machine vision is not possible.

In the 1990's, another significant development was the flooding of digital cameras into the market. Almost at every quarter, digital cameras were made and sold, cheaper, and with better resolutions with the increased number of pixel per frame of image. One of the handy low-end cameras that can be easily attached to the PCs through the USB port is the Web Camera. Its' low-resolution images, using small amount of memory is ideal for transmission through the Internet for tele-conferencing purpose.

Softwares that were hardwired into the camera, as well as that provided through CD's and disks, include mainly image processing functions. Enhancing of images can be