

# 從影片序列中建立以人為基礎的影像資料

學生：郭文傑

指導教授：李錫堅教授

國立交通大學資訊工程研究所碩士班

## 摘 要

本論文之研究目的在於建立以人為基礎的影像資料。在一個監控環境中，我們希望建立一個機制，能把各個人在環境中的過程片段串接和組織成爲影像序列，並把對於某人來說，具代表性的影像儲存到影像檔案保管處。若我們想要尋找某人的影像片段時，我們不必從所有的影片中去找，而是從影像檔案保管處去找相關的代表性影像，再利用已串接好的影像序列資訊，去找到那個人在監控環境中的一連串的影像。如此一來，就可以減少不必要的搜尋工作，並且能提升監控系統整體的效能。我們所提出的系統，分成三個部份：人的追蹤、人的樣式比對和具代表性影像的選擇和影像歸檔。

第一個部分，先以一個統計式的方法建立背景模組，這個模組幫助我們取出前景(人)部份。進而開始根據抽取出來的前景來做追蹤。依據目前人所在的位置往外張出一個搜尋範圍，我們提出一個有效且快速的方法，讓我們可以只檢查少數的搜尋點就可以找出人移向的位置。所以我們就可以追蹤人在環境中的一切情況。

第二個部份，人在監控環境中被多相機所拍攝到時，是有許多不同角度，所以同一個人，會有各個角度的影像，這使得我們在做人之間的比對時，常常遇到雖是同一人，但卻被判斷成不同人。我們設計了一個有效的人的比對方法，它利用的是：雖然對同一人來說，被拍到時有許多不同角度，但身體某些部份還是會有相同的樣式。根據這個概念來做人之間的比對。

第三個部份，我們希望從一個人在環境中的片段中，挑選出許多具有代表性的影像。根據人被拍攝到的大小、比例、灰階值、樣式和角度選出重要影像，能表示一個人在環境中的一些重要過程片段。所以對於某一個人來說，取得每個相機拍攝到的重要影像，再利用人的比對來幫助做影像的串接，而每隔一段時間，我們就做影像歸檔的動作，將拍攝到的人的影像串接序列寫到影像檔案保管處。

實驗的部份，我們測試了實際情況中人的追蹤、人在每個相機間的影像串接和在檔案保管處做人的搜尋。實驗結果發現我們所提出來的機制有著不錯的效果，我們可以利用這個機制來協助監控系統。



# Archiving of Human-Based Images from Video Sequences

Student : Wen-Jie Guo

Advisors : Dr. Hsi-Jian Lee

Department of Computer Science and Information Engineering  
National Chiao Tung University

## ABSTRACT

In this paper, the research purpose is to construct archiving of human-based images. A mechanism connecting and organizing video sequences that each person passed through scenes in the surveillance system into a lot of image sequencing and recording representative image frames into archiving for each person is established. If we want to obtain image frames about a specific person, we can obtain corresponding representative frames about him from archiving and exploit information of image sequencing connected to obtain a series of video sequences, in place of searching of all video sequences. In hence, we can reduce much searching workload and improve performance of the surveillance system. The system consists of three stages: human tracking, human matching, and selection of key frames and archiving.

In the first stage, based on a statistical background model, we apply background subtraction to extract components of people. We start to track a person according to the current position of him. Candidate positions of the tracked person are located in a search region centered on the current position. We propose an effective and fast method checking few search points from the search region to obtain the position that the person moves to. So, we can track people thoroughly in the surveillance system.

In the second stage, a person captured by multiple cameras in the surveillance system has many different views. For the same person, having images with multiple views results in many cases that a person is wrongly determined as many different people in human matching. An effective human matching method is designed. It exploits some parts between two compared images of the same person with different views may have similar appearances. Human matching is according to the idea.

In the third stage, we hope to select several representative frames from video sequences that a person passed through scenes. According to the size, ratio, intensity, appearances, and views of a person captured by cameras, several key frames represent significant events for him in the surveillance system are selected. For a specific person, key frames in each camera are obtained, these key frames from all cameras are connected into a lot of image sequencing by human matching, and a lot of image sequencing is recorded into archiving every a time period.

In our experiment, human tracking in real cases, connection of image sequences of each camera, and human searching in archiving are tested. From experiment result, the mechanism proposed with effective performance can be facilitated to the surveillance system.

# Acknowledgements

I heartily express my sincere gratitude to my advisor, Pro. Hsi-Jain Lee, for his proper guidance and patient discussion throughout course of this study.

Many sincere thanks are given to all colleagues in the Document Processing and Character Recognition Laboratory at National Chiao Tung University for their assistance on this thesis. Especially thanks the senior Shan-Lung Zhao for his patient guidance.

Thanks my friends and girlfriend a lot for their help and encouragement in those days. Finally, I would like to express my deep gratitude to my family for their inspiration and support. This thesis is dedicated to all of them.



# TABLE OF CONTENTS

CHINESE ABSTRACT.....	i
ENGLISH ABSTRACT.....	iii
ACKNOELEDGEMENTS.....	v
TABLE OF CONTENTS.....	vi
LIST OF FIGURES.....	ix
LIST OF TALBES.....	x
CHAPTER 1 INTRODUCTION.....	1
1.1 Motivation.....	1
1.2 Problem Definition.....	2
1.2.1 Obtaining of human-based image sequences in multiple cameras.....	2
1.2.2 Obtaining of representative frames to assist matching for a specific person.....	2
1.2.3 Matching of humans in different appearances.....	2
1.3 Survey of Related Research.....	3
1.3.1 Moving Object Detection.....	3
1.3.2 Object Tracking.....	5
1.3.3 Fast Search Algorithm for Movement Estimation.....	7
1.4 Assumptions.....	9
1.5 System Description.....	9
1.5.1 Human Extraction.....	11
1.5.2 Human Tracking in Single Video Sequences....	11
1.5.3 Human Matching in Different Sequences.....	11
1.5.4 Key Frame Selection and Archiving.....	11
1.6 Thesis Organization.....	12

CHAPTER 2 HUMAN TRACKING.....	13
2.1 Human Extraction and Matching.....	13
2.1.1 Human Extraction.....	13
2.1.2 Quarter Search Algorithm.....	16
2.1.3 Feature Matching.....	21
2.2 Construction of the Walking Map.....	23
 CHAPTER 3 APPEARANCE-INDEPENDENT HUMAN MATCHING.....	 26
3.1 Appearance-Independent Matching.....	26
3.2 Application.....	31
 CHAPTER 4 ARCHIVING OF HUMAN-BASED IMAGES..	32
4.1 Selection of Key Frames.....	33
4.2 Functions.....	38
 CHAPTER 5 EXPERIMENTAL RESULT AND DISCUSSION.....	 40
5.1 Experiment Result.....	40
5.1.1 Human Tracking.....	40
5.1.2 Human Matching in Multiple Cameras.....	43
5.1.3 Human Matching in Archiving.....	46
5.2 Analysis of Error Results.....	50
5.2.1 Small Sizes of People.....	50
5.2.2 Occlusion.....	50
5.2.3 Variant Intensity.....	52
5.2.4 Crumbled Entity of Matching.....	52

5.2.5 Different Appearances in Front View and Back View.....	53
5.3 Discussion.....	56
CHPATER 6 CONCLUSION AND FUTURE WORK.....	57
REFERENCE.....	59





# LIST OF FIGURES

Fig. 1.5.1 The system flow diagram.....	10
Fig. 2.1.1.1 Example of foreground segmentation.....	15
Fig. 2.1.2.1 Example of the three-step search algorithm.....	19
Fig. 2.1.2.2 Example of the quarter search algorithm.....	19
Fig. 2.1.3.1 A diagram for feature matching.....	21
Fig. 2.1.1 Example of tracking results.....	22
Fig. 2.2.1 Example representation of a walking path.....	25
Fig. 2.2.2 Example of a directional map.....	25
Fig. 3.1 some parts of two compared people are similar.....	30
Fig. 3.2 Example of appearance-independent matching.....	30
Fig. 4.1.1 Example of key frames with significantly different average intensity.....	35
Fig. 4.1.2 Example of key frames with significantly different appearances.....	35
Fig. 5.1.1.1 Example of tracking results.....	41
Fig. 5.1.1.2 Example of tracking results.....	42
Fig. 5.1.2.1 front view.....	44
Fig. 5.1.2.2 side view.....	45
Fig. 5.1.2.3 back view.....	45
Fig. 5.2.1 Example of occlusion-resistance.....	51
Fig. 5.2.2 Example of without occlusion-resistance.....	51
Fig. 5.2.3 A person has different intensity.....	54
Fig. 5.2.4 Entities of people are crumbled.....	54
Fig. 5.2.5 The person that should be matched is crumbled.....	55
Fig. 5.2.6 Appearances of front and back views may be much different.....	55

# LIST OF TABLES

Table 2.1.2.1 Time complexity of three-step and quarter search algorithms.....	20
Table 2.1.2.2 Comparison of computing time of three-step and quarter search algorithms in real case.....	20
Table 5.1.1.1 Tracking results.....	41
Table 5.1.2.1 Matching rates in multiple cameras.....	44
Table 5.1.3.1 Experiment result of human matching in archiving.....	48
Table 5.1.3.2 Matching rates between views.....	49
Table 5.1.3.3 Experiment result of human matching in archiving after improvement.....	49

