ห้องสมุดงานวิจัย สำนักงานคณะกรรมการวิจัยแห่งชาติ
E46964

ROBUST AMBIGUOUS TARGET HANDLING FOR VISUAL OBJECT TRACKING

MR. KAIROEK CHOEYCHUEN

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
(ELECTRICAL AND COMPUTER ENGINEERING)
FACULTY OF ENGINEERING
KING MONGKUT'S UNIVERSITY OF TECHNOLOGY THONBURI
2010

boo246628

ห้องสมุดงานวิจัย สำนักงานคณะกรรมการวิจัยแห่งชาติ
E46964

Robust Ambiguous Target Handling for Visual Object Tracking

Mr. Kairoek Choeychuen M.Eng. (Electrical Engineering)

A Thesis Submitted in Partial Fulfillment
of the Requirements for
the Degree of Doctor of Philosophy (Electrical and Computer Engineering)
Faculty of Engineering
King Mongkut's University of Technology Thonburi

THE NATIONAL RESEARCH COUNCIL OF THAILAND
RESEARCH LIBRARY

Thesis Committee

..... Chairman of Thesis Committee
(Prof. Prabhas Chongstitvattana, Ph.D.)

..... Member and Thesis Advisor
(Assoc. Prof. Kosin Chamnongthai, D.Eng.)

..... Member
(Asst. Prof. Pinit Kumhom, Ph.D.)

..... Member
(Assoc. Prof. Suthep Madarasmi, Ph.D.)

..... Member
(Asst. Prof. Yuttana Kidjaidure, Ph.D.)

..... Member
(Asst. Prof. Tochiaki Kondo, Ph.D.)

..... Member
(Asst. Prof. Werapon Chiracharit, Ph.D.)

Copyright reserved

Thesis Title	Robust Ambiguous Target Handling for Visual Object Tracking
Thesis Credits	36
Candidate	Mr. Kaiorek Choeychuen
Thesis Advisors	Assoc. Prof. Dr. Kosin Chamnongthai
	Asst. Prof. Dr. Pinit Kumhom
Program	Doctor of Philosophy
Field of Study	Electrical and Computer Engineering
Department	Computer Engineering
Faculty	Engineering
B.E.	2553

E46964

Abstract

This thesis presents robust and efficient approaches of visual object tracking with ambiguous targets handling such as partial occlusion, split region or mixed state between the partial occlusion and the split regions. This thesis proposes a hybrid tracking process with ambiguous target handling by combining advantages of the particle filter (PF) and the nearest neighbor data association (NNDA) together. Firstly, we use a compact object models for handling the ambiguous targets that consume low computation. At the first glance, we propose using only a bounding box model and an expected object region as the object models. The experimental results showed that it was sufficient for ambiguous target handling in some circumstances such as failing in color features representation etc. Next, other features were integrated in the preliminary object models to improve the accuracy of tracking through the particle filter-based object tracking called the hierarchical particle filter embedded non-stationary Gaussian object model (HPF-NSGM). Finally, we defined new distance function for the NNDA called as a state classifier for robust determining target state by using a single threshold. The single threshold was defined for classifying target states into three states by considering interaction with the expected object region that consist of fully, partial and no overlap ones. The experimental results showed that a larger threshold for the state classifier increased the number of partial occlusion that can be solved by using the proposed NNDA and HPF-NSGM while the smaller threshold increased the number of faulty new targets which are considered as redundancy. We might lose many existing target histories from the redundancy of the new targets that were not good for event interpretation.

Keywords: Visual Object Tracking / Ambiguous Target Handling / State Classifier / Hierarchical Particle Filter / Non-stationary Gaussian Object Model.

หัวข้อวิทยานิพนธ์	การสร้างความทนทานในการจัดการกับเป้าหมายที่คลุมเครือสำหรับ การติดตามวัตถุจากภาพ
หน่วยกิต	36
ผู้เขียน	นายไกรฤกษ์ เชยชื่น
อาจารย์ที่ปรึกษา	รศ.ดร.โกสินทร์ จำนงไทย
	ผศ.ดร.พินิจ กำหอม
หลักสูตร	ปรัชญาดุษฎีบัณฑิต
สาขาวิชา	วิศวกรรมไฟฟ้าและคอมพิวเตอร์
ภาควิชา	วิศวกรรมคอมพิวเตอร์
คณะ	วิศวกรรมศาสตร์
พ.ศ.	2553

E46964

บทคัดย่อ

วิทยานิพนธ์นี้นำเสนอวิธีการสร้างความทนทานและมีประสิทธิภาพในการติดตามวัตถุภาพกับการจัดการเป้าหมายที่คลุมเครือ เช่น เป้าหมายถูกบัง, บริเวณวัตถุที่ถูกแบ่ง หรือสถานะที่ผสมระหว่างการถูกบังกับบริเวณวัตถุที่ถูกแบ่ง วิทยานิพนธ์นี้เสนอกระบวนการติดตามแบบผสม พร้อมกับการจัดการความคลุมเครือโดยรวมข้อดีของ พาร์ติเคิลฟิลเตอร์ (PF: Particle Filter) และการเชื่อมโยงข้อมูลโดยใช้ข้อมูลข้างเคียง (NNDA: Nearest Neighbor Data Association) เข้าด้วยกัน อันดับแรกเราใช้แบบจำลองที่กระชับ ใช้ทรัพยากรในการคำนวณน้อย สำหรับการจัดการกับเป้าหมายที่คลุมเครือ ในเบื้องต้นเราเสนอการใช้แบบจำลองกล่องล้อมรอบ (Bounding box model) และบริเวณวัตถุเฉลี่ย (Expected object region) ในฐานะแบบจำลองวัตถุ ผลการทดลองแสดงว่ามันเพียงพอสำหรับการจัดการกับการคลุมเครือในบางสถานการณ์ได้ เช่น เมื่อเกิดการผิดพลาดของการแทนวัตถุด้วยสี เป็นต้น ทำให้ไม่สามารถใช้สีในการติดตามวัตถุได้ ลำดับต่อมาลักษณะเด่นอื่นๆ ถูกรวมเข้าไปในแบบจำลองขั้นต้นเพื่อปรับปรุงความถูกต้องของการติดตามผ่าน พาร์ติเคิลฟิลเตอร์ ซึ่งถูกเรียกว่า พาร์ติเคิล ฟิลเตอร์แบบลำดับขั้นซึ่งผนวกกับแบบจำลองเกาส์เซียนที่ไม่คงตัว (HPF-NSGM: Hierarchical-based Particle Filter embedded Non-stationary Gaussian Object Model) ในตอนท้ายเรานิยามฟังกชันระยะใหม่สำหรับ NNDA เรียกว่าตัวคัดแยกสถานะ สำหรับการกำหนดสถานะของเป้าหมายอย่างแข็งแกร่งโดยใช้ค่าตัดสินเพียงค่าเดียว ค่าตัดสินค่าเดียวนี้ถูกกำหนดสำหรับแบ่งสถานะของเป้าหมาย 3 สถานะโดยพิจารณาการเกี่ยวข้องกับ ค่าบริเวณวัตถุเฉลี่ย ซึ่งประกอบด้วย การทับซ้อนกันแบบ เต็ม, บางส่วนและไม่เกี่ยวข้องกัน ผลการทดลองแสดงว่า ค่าตัดสินที่มากเพิ่มจำนวนของการถูกบังบางส่วนซึ่งสามารถถูกแก้ปัญหาได้ง่ายโดยใช้ NNDA และ HPF-NSGM ที่ถูกเสนอ ในขณะที่ค่าตัดสินน้อย เพิ่มจำนวนของเป้าหมายเทียมซึ่งถูกพิจารณาว่าเป็นเป้าหมายฟุ่มเฟือย เราอาจ สูญเสีย ข้อมูลเป้าหมายในอดีตซึ่งไม่ดีสำหรับการตีความเหตุการณ์ที่เกิดขึ้น

E46964

คำสำคัญ : การติดตามวัตถุภาพ / การจัดการเป้าหมายที่คลุมเครือ / ตัวจำแนกสถานะเป้าหมาย / พาร์ติเคิลฟิลเตอร์แบบลำดับขั้น / แบบจำลองวัตถุแบบเกาส์เซียนที่ไม่คงตัว

ACKNOWLEDGEMENTS

I would like to express my profound gratitude to my advisers, Assoc. Prof. Dr. Kosin Chamnongthai and Asst. Prof. Dr. Pinit Kumhom from the Department of Electronic and Telecommunication Engineering, for their valuable guidance and kind suggestions on this research. Special thanks to Prof. Dr. Prabhas Chongstitvattana from the Department of Computer Engineering, Faculty of Engineering, Chulalongkorn University (CU) for kindly serving as the Chairman, and Assoc. Prof. Dr. Suthep Madarasmi from the Department of Computer Engineering, Dr. Werapon Chiracharit from the Department of Electronics and Telecommunications Engineering, Asst. Prof. Dr. Yuttana Kidjaidure from Department of Electronics, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang (KMITL) and Asst. Prof. Dr. Tochiaki Kondo from School of Information, Computer, and Communication Technology, Siridhorn International Institute of Technology (SIIT), Thammasat University for kindly serving as committee members of this thesis and their valuable comments and suggestions.

I would like to acknowledge Prof. Dr. Mubarak Shah and Dr. Arslan Basharat for spending their valuable time to fruitful discussion on various ideas in research during my visit at the Computer Vision Laboratory in University of Central Florida.

I would like to acknowledge the Department of Electronic and Telecommunication Engineering, Rajamangala University of Technology Rattanakosin and the Ministry of Science and Technology, Thailand, for financial support during my doctoral study.

Sincerely thanks to all members of the Computer Vision Research Laboratory (Vision Lab). Finally, I would like to express to my family members for their continuous support, encouragement, understanding and love.

CONTENTS

	PAGE
ENGLISH ABSTRACT	ii
THAI ABSTRACT	iv
ACKNOWLEDGEMENTS	v
CONTENTS	vi
LIST OF TABLES	viii
LIST OF FIGURES	ix
LIST OF SYMBOLS	xi
LIST OF TECHNICAL VOCABULARY AND ABBREVIATIONS	xii
CHAPTER	
1. INTRODUCTION	**1**
1.1 Motivation	1
1.2 The Object Tracking	2
1.3 The Objectives of this Thesis	3
1.4 Research Methodology	3
1.5 Research Contributions	4
1.6 The Organization of this Thesis	4
2. LITERATURE REVIEWS AND BACKGROUND OF PROBELMS OF VISUAL OBJECT TRACKING	**5**
2.1 The Object Detection	5
2.1.1 Recognition-based Object Detection	6
2.1.2 Background subtraction-based Object Detection	7
2.2 The Object Representation	8
2.2.1 Feature-based Object Representation	8
2.2.2 Edge-based Object Representation	10
2.2.3 Appearance-based Object Representation	12
2.3 The Visual Object Tracking with Ambiguous Target Handling	13
2.3.1 Recognition-based Object Tracking	13
2.3.2 Motion-based Object Tracking	18
2.4 The Overview of the Thesis	18
3. FEATURE INTEGRATION FOR VISUAL OBJECT TRACKING	**21**
3.1 Introduction	21
3.2 The Object Models	
3.2.1 Preliminary object models for the motion-based object tracking	23
3.2.2 Extended object models for the recognition-based object tracking	25
3.3 The Color-based Particle Filter	26
3.4 Hierarchical Feature Integration	28
3.5 Linear Combination-based Feature Integration	29
3.6 The Particle Filter Embedded Non-stationary Gaussian Object Model	31
3.6.1 Problem analysis	31
3.6.2 Experimental Results and Discussion	32

4. HYBRID AMBIGUOUS TARGET HANDLING **39**
4.1 Introduction 39
4.2 System Model 41
4.3 The Hybrid Approach 42
4.3.1 The Object Models 42
4.3.2 State Classifier 43
4.3.3 Ambiguous Target Solving 46
4.4 Experimental Results and Discussion 48

5. CONCLUSION **58**

REFERENCES **59**

APPENDIX: Publications **63**

CURRICULUM VITAE **65**

LIST OF TABLES

TABLE		PAGE
4.1	Summary of solvable problems of the NNDA and the Particle filter	40
4.2	Summary of the object models for some tracking processes and problems in this thesis	43
4.3	The 2^{nd} association matrix of five existing targets (track) and five newly detected objects from background-based object detection.	46
4.4	Effect of different weight *(a, b)* in equation 8.	49
4.5	Comparison of the number of object features for tracking the partial target with other approaches	52
4.6	Comparison of time consumption (in seconds) for the proposed approach with adaptive color-based and kernel-based particle filter	53
4.7	The results for the sequence from the CARVIER Data Set	55
4.8	The results for the sequence from the UCF Computer Vision Data Set	55

LIST OF FIGURES

FIGURE		PAGE
1.1	Block diagram of automatic video surveillance (AVS)	1
1.2	Some mistakes of defining the meaningful events (trajectories-based) from some errors of object detection and tracking.	2
1.3	Samples of ambiguous targets: a) The split region, b) The merged region (partial occlusion), and c) The mixed state (split + merged region)	2
1.4	Early object tracking: a) A radar flight tracking system, and b) A sonar tracking system	3
2.1	The example result of background subtraction-based object detection [12]: a) Input image with two persons and swaying trees, b) Image noises due to swaying trees, and c) Final result after post-processing	6
2.2	Example of the temporal texture template proposed by [12]	9
2.3	Example of Gray-Level Co-Occurrence Matrix with distance = 1 and direction = → (right) performed on image I	9
2.4	Selected features for the face classifier [18]	10
2.5	An example of the active contour result of an object at 1500 iterations	11
2.6	An example of the EOH of image I with 4-bin of polar coordinate	11
2.7	An example result of the PSO-based tracking with color models	16
2.8	Thesis's Overview	20
3.1	An example of incorrect detected human obtained from cluttered environment [22]	21
3.2	Condensation algorithm [35]	22
3.3	The object model: a) The detected object, and b) The updated object model	24
3.4	The example of the weight and threshold of the accumulated object regions with w_a=0.6 and Th_a=0.3 respectively	25
3.5	Block diagram of hierarchical-based particle filter (left) applying the idea of the cascade object detector [18] (right)	28
3.6	Hierarchical-based (cascade) particle filter algorithm	29
3.7	Block diagram of linear combination-based particle filter applying the idea of the democratic integration [46]	29
3.8	Linear combination-based (Democratic integration) particle filter algorithm	30
3.9	Estimated triangle waves, shown as red line, with different alphas: a) alpha = 0.1, b) alpha = 0.5, c) alpha = 0.8, and d) alpha = 0.9	32
3.10	Selection of alpha from (3.1) – (3.2) for non-stationary Gaussian model	33
3.11	The qualitative results of the proposed particle filter and the adaptive color-based particle filter [43]	34
3.12	The quantitative results of the proposed particle filter and the adaptive color-based particle filter [43]	36
3.13	Severe change of object shape (non-rigidity) caused of higher tracking errors in the proposed particle filter	36
3.14	Failure in the estimated bounding box and color model updated of the color-based particle filter over time	36
3.15	The comparative results of the root mean square (RMS) of the tracking errors between the proposed particle filter at 500 samples and the color-based particle filter at 500, 800, 1000 and 3000 samples.	37

3.16 Average Euclidean distance, represented as tracking error, of each set of the object models for the particle filter with ground truth 37
4.1 Possible state transitions of a target. 41
4.2 The block diagram of the proposed tracking system. 41
4.3 The components of the distance function D 45
4.4 Possible interaction patterns for the distance function D computation a) No overlapping, b) Partial overlapping, and c) Full overlapping. 45
4.5 Resolving the split state by using the expected object region. 47
4.6 Solving the merged state with the extended bounding box. 48
4.7 The results of the merged state (partial occlusion) handling: a) The opposite path, b) The cross path, c) The Complex Movement Path, d) The Parallel Path of Three Persons 50
4.8 The mom and babies' sequences: a) the trajectory results, b) the binary moving regions of the split regions and the merged regions (the top row), and tracking results of the top row (the bottom row) 51
4.9 The effect of the Larger and Smaller Th_d: a) The sensitivity of faulty new track (dash line) state occurred from the smaller Th_d, and b) The redundancy of the mixed state (dash-dot line) occurred from the lager Th_d 52
4.10 Comparison of tracking results for the Meet-Split-3rdGuy-jpg of the CAVIAR data set: a) Adaptive color-based particle filter [20], b) Kernel-based particle filter [21], and c) The proposed approach 54
4.11 Examples of the UCF sequence the two cars occluded by the trees 55
4.12 A comparison results of solving spurious new track in split region: a) The conventional method [14], and b) The proposed method 56
4.13 A comparison results of solving spurious new track in early confirmable state: a) The conventional method [14], and b) The proposed method 56

LIST OF SYMBOLS

SYMBOL		UNIT
t_k	Frame number k of the image sequence	-
μ_{t+1}	New mean of the recursive width and height	Pixel
σ^2_{t+1}	New variance of the recursive width and height	Pixel
z_{t+1}	Bounding box width and height of the newly detected object	Pixel
α	Forgotten rate of the σ^2_{t+1} and μ_{t+1}	-
p1	The top-left point (x_1, y_1) on the bounding box feature	-
p2	The bottom-left point (x_2, y_2) on the bounding box feature	-
p3	The top-right point (x_3, y_3) on the bounding box feature	-
p4	The bottom-right point (x_4, y_4) on the bounding box feature	-
p5	The center point (x_5, y_5) on the bounding box feature	-
w	Recursive width	Pixel
h	Recursive height	Pixel
σ_x, σ_y	Sample variances of the *p5*	-
σ_w, σ_h	Sample variances of the width and height	-
B_t	Expected object region (EOR)	
F_t	The newly detected object from the background-based object detection at frame t	-
A_t	Accumulated object region	-
Th_a	Threshold for computing the expected object region	-
w_a	Forgotten rate of the B_t	-
s'_t	Potential samples for the particle filter algorithm	-
$s_t^{(n)}$	Sample number n of the particle filter algorithm at time t	-
A	Deterministic component of the dynamic model	-
ω_t	Vector of Gaussian random variable at time t	-
$\pi_t^{(n)}$	Corresponding sample weight of the $s_t^{(n)}$	-
D(r, c)	Distance function of the detected object r from motion-based object detection and detected object c from motion-based object detection	-
Th_d	Gating threshold	-
$Track_{area}(r)$	The estimation of the updated *EOR*	-
$Observation_{area}(c)$	Newly detected object area	-
(X_p, Y_p)	Predicted location of the feature point	-
(X_o, Y_o)	Observation location of the feature point	-
min()	Find the minimum value	-
2^{nd} *AM[...]*	The cell value of the 2nd association matrix	-
MC[...]	Margin column value of a track	-
MR[...]	Margin row value of a observation	-
Track(...)	A track in the 2nd association matrix	-
Obs(...)	A observation in the 2nd association matrix	-

LIST OF TECHNICAL VOCABULARY AND ABBREVIATION

AVS	=	Automatic Video Surveillance
CCTV	=	Closed Circuit Television
HCI	=	Human Computer Interface
CCD	=	Charge Coupled Device
LCD	=	Liquid Crystal Display
FCDT	=	Fast Constrained Delaunay Triangulation
SVM	=	Support Vector Machine
GLCM	=	Gray Level Co-occurrence Matrices
EOH	=	Edge Orientation Histogram
3D	=	Three Dimensional
RGB	=	Red, Green, Blue
SKDA	=	Sequential Kernel Density Approximation
MHT	=	Multiple Hypothesis Tracking
JPDAF	=	Joint Probabilistic Data Association Filter
NNDA	=	Nearest Neighbor Data Association
PF	=	Particle Filter
Bi-models	=	Bi-models of the Gaussian distribution
AM	=	Association Matrix
MR	=	Margin Row
MC	=	Margin Column
2^{nd} AM	=	2^{nd} Association Matrix
NSGM	=	Non-stationary Gaussian Object Model
HPF-NSGM	=	Hierarchical Particle Filter Embedded Non-stationary Gaussian Object Model