

2010

Contextual bayesian inference for visual object tracking and abnormal behavior detection

Philippe L. Bouttefroy
philippe@uow.edu.au

Follow this and additional works at: <https://ro.uow.edu.au/theses>

University of Wollongong

Copyright Warning

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site.

You are reminded of the following: This work is copyright. Apart from any use permitted under the Copyright Act 1968, no part of this work may be reproduced by any process, nor may any other exclusive right be exercised, without the permission of the author. Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material.

Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

Unless otherwise indicated, the views expressed in this thesis are those of the author and do not necessarily represent the views of the University of Wollongong.

Recommended Citation

Bouttefroy, Philippe L., Contextual bayesian inference for visual object tracking and abnormal behavior detection, Doctor of Philosophy thesis, School of Electrical, Computer and Telecommunications Engineering, Faculty of Informatics, University of Wollongong, 2010. <https://ro.uow.edu.au/theses/3117>

NOTE

This online version of the thesis may have different page formatting and pagination from the paper copy held in the University of Wollongong Library.

UNIVERSITY OF WOLLONGONG

COPYRIGHT WARNING

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site. You are reminded of the following:

Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material. Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

Contextual Bayesian Inference for Visual Object Tracking and Abnormal Behavior Detection

A thesis submitted in fulfilment of the
requirements for the award of the degree

Doctor of Philosophy

from

THE UNIVERSITY OF WOLLONGONG

by

Philippe Loïc Marie Bouttefroy
Masters of Engineering Studies
(Telecommunications and Computer Science),
Diplôme d'ingénieur

SCHOOL OF ELECTRICAL, COMPUTER
AND TELECOMMUNICATIONS ENGINEERING
2010

To my hair †

Abstract

Visual object tracking has been extensively investigated in the last two decades for its attractiveness and profitability. It remains an active area of research because of the lack of a satisfactory holistic tracking system that can deal with intrinsic and extrinsic distortions. Illumination variations, occlusions, noise and errors in object matching and classification are only a fraction of the problems currently encountered in visual object tracking. The work developed in this thesis integrates contextual information in a Bayesian framework for object tracking and abnormal behavior detection; more precisely, it focuses on the intrinsic characteristics of video signals in conjunction with object behavior to improve tracking outcomes.

The representation of probability density functions is essential for modeling stochastic variables. In particular, parametric modeling is convenient since it makes possible the efficient storage of the representation and the simulation of the underlying stochastic process. The Gaussian mixture model is employed in this thesis to represent the pixel color distribution for segregation of foreground from background. The model adapts quickly to fast changes in illumination and resolves the problem of “pixel saturation” experienced by some existing background subtraction algorithms. The technique leads to better accuracy in the extraction of the foreground for higher-level tasks such as motion estimation.

The solution of the Bayesian inference problem for Markov chains and, in particular, the well-known Kalman and particle filters is also investigated. The integration of contextual inference is of paramount importance in the aforementioned estimators;

it results in object-specific tracking solutions with improved robustness. The vehicle tracking problem is explored in detail. The projective transformation, imposed by the environment configuration, is integrated into the Kalman and particle filters, which yields the “projective Kalman filter” and the “projective particle filter”. Extensive experimental results are presented, which demonstrate that the projective Kalman and particle filters improve tracking robustness by reducing tracking drift and errors in the estimated trajectory. The constraint on the known nature of the environment is then relaxed to allow general tracking of pedestrians. A mixture of Gaussian Markov random fields is introduced to learn patterns of motion and model contextual information with particle filtering. Such inference results in an increased tracking robustness to occlusions.

The local modeling with the Markov random fields also provides inference on abnormal behavior detection. Since local patterns are unveiled by the Markov random field mixture, detecting abnormal behavior is reduced to the matching of an object feature vector to the underlying local distribution, whereas the global approach, introducing generalization errors, involves complex, cumbersome and inaccurate decisions. Experimental evaluation on synthetic and real data show superior results in abnormal behavior detection for driving under the influence of alcohol and pedestrians crossing highways.

Résumé

Le suivi d'objets visuel a été un domaine de recherche intense durant ces deux dernières décennies pour son attrait scientifique et sa rentabilité. Il reste un sujet de recherche ouvert de par le manque de système de suivi holistique satisfaisant, prenant en compte les distorsions intrinsèques et extrinsèques. Variations d'éclairement, occlusions, bruits et erreurs dans la correspondance et la classification d'objets ne sont qu'une partie des problèmes actuellement rencontrés en suivi d'objets. Le travail développé dans cette thèse intègre l'information contextuelle dans le cadre Bayésien pour le suivi d'objets et la détection de comportements anormaux. Plus précisément, la recherche porte sur les caractéristiques intrinsèques du signal vidéo en conjonction avec le comportement d'objets dans le but d'améliorer les résultats du suivi.

La représentation de fonctions de densité de probabilité est cruciale pour modéliser les variables aléatoires. En particulier, les modèles paramétriques sont pratiques puisqu'ils permettent un stockage compact de la représentation ainsi que la simulation du processus aléatoire sous-jacent. La mixture de Gaussiennes est utilisée dans cette thèse pour représenter la distribution de couleur d'un pixel dans le but de séparer l'avant-plan de l'arrière-plan. Le modèle s'adapte aux changements rapides d'éclairements et résout le problème de "saturation de pixels" rencontré avec certains algorithmes de soustraction d'arrière-plan. Il résulte de cette technique une meilleure précision lors de l'extraction de l'avant-plan pour des tâches de plus haut niveau telles que l'estimation du mouvement.

La solution au problème d'inférence Bayésienne pour les chaînes de Markov, et en particulier, les filtres de Kalman et particulaire, est étudiée. L'intégration d'une inférence contextuelle dans ces estimateurs est primordiale pour améliorer le suivi d'objet. Il en découle des solutions propres à un contexte spécifique. Le problème de suivi de véhicules est également exploré en détails dans cette thèse. La transformation projective, imposée par la configuration de l'environnement, est intégrée dans les filtres de Kalman et particulaire, engendrant le "filtre de Kalman projectif" et le "filtre particulaire projectif". Des résultats expérimentaux exhaustifs sont présentés pour démontrer l'amélioration de la robustesse au suivi par les filtres de Kalman et particulaire projectifs. L'amélioration est caractérisée par la réduction de la dérive du suiveur et la réduction de l'erreur dans l'estimée de la trajectoire. La contrainte sur le caractère connu de l'environnement est ensuite supprimée pour permettre le suivi de piétons. Une mixture de champs aléatoires de Markov Gaussiens est introduite dans l'objectif d'apprendre les motifs de mouvements et de modéliser l'information contextuelle pour le filtrage particulaire. Une augmentation de la robustesse du suivi sous occlusion résulte d'une telle inférence.

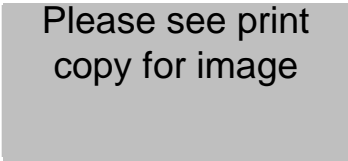
La modélisation locale avec les champs aléatoires de Markov fournit également une inférence pour la détection de comportements anormaux. Puisque les motifs locaux sont révélés par la mixture de champs aléatoires de Markov, la détection de comportements anormaux est réduite à l'étude de la correspondance entre le vecteur de caractéristiques et la distribution locale sous-jacente. L'approche globale, quant à elle, introduit des erreurs de généralisation et implique des décisions complexes, peu élégantes et imprécises. L'évaluation expérimentale de la méthode proposée sur des données synthétiques et réelles présente des résultats supérieurs pour la détection des comportements anormaux de conducteurs en état d'ébriété et de piétons traversant les autoroutes.

Statement of Originality

This is to certify that the work described in this thesis is entirely my own, except where due reference is made in the text.

No work in this thesis has been submitted for a degree to any other university or institution, to the exception of the University Paris 13 (France) with which a cotutelle agreement (Joint Doctorate) has been signed.

Signed



Please see print
copy for image

Philippe Loïc Marie Bouttefroy

21st of January, 2010

Acknowledgments

I would like to express my gratitude to all of those who provided me with the resources to complete my thesis. First, I would like to thank my supervisors, Prof. A. Bouzerdoun and Prof. A. Beghdadi as well as my co-supervisor Dr. Phung, for their insights on my research progresses throughout the thesis. More importantly, I acknowledge their open-mindedness towards postgraduate research that enabled me to explore fields of personal interests rather than being trapped in a predefined path leading to the completion of my thesis. I also owe my genuine appreciation to Prof. A. Bouzerdoun who financially supported me during the first year of this journey, in particular. Second, my deepest expression of gratitude goes to my family. Dad and Mom for their relentless efforts to make me grow as a person, for their support in tough moments and also, for their financial help throughout my education; Aymeric, Marjorie, Séverine and Alice for their brotherhood and for saying the right word at the right moment. The love and care of the family were always shining in the distance to guide and comfort me along the different steps of the thesis. Third, my postgraduate experience wouldn't have been the same without Weerona College. The distractions and the support of the community, in particular Leanne, the SR team and the residents, most definitely helped me to keep my sanity during the thesis. Fourth, I would like to specially thank two people who are very dear to me: Tracey and Rachel. Tracey, thank you for pushing me back to the top of the roller coaster that the thesis writing stage is. Rachel, thank you for the insightful and non-technical conversations we had during our Friday meetings at the North Wollongong pub and the now famous G&T's on the balcony. Fifth, my thoughts go

to Laëtita who has taught me more than I could ever understand and who pushed me to always aim for excellence by her attitude. I would most certainly not be where I am now had our paths not met. Finally, I am indebted to a number of university groups and members, namely the ICT postgraduate students and Roslyn, for their support and joy.

Contents

1	Preliminaries	1
1.1	Introduction	1
1.2	Representation of Video Signals	2
1.2.1	Concepts and Notation	2
1.2.2	Video Acquisition	4
1.2.3	Information Distortion	5
1.2.4	Research Motivation and Assumptions	8
1.3	Contributions of the Thesis	9
1.4	Publications	11
2	Roadmap for the Object Tracking Maze	13
2.1	Introduction	13
2.2	Object Modeling	14
2.2.1	Parametric Representations	15
2.2.2	Non-parametric Representations	18
2.2.3	Object Features	20
2.2.4	Summary of Object Modeling	24
2.3	Object Identification	25

	2.3.1 Object Detection using Supervised Learning	25
	2.3.2 Distribution Representation for Object Detection	28
	2.3.3 Object Segmentation	32
	2.3.4 Summary of Object Identification	36
2.4	Object Tracking	36
	2.4.1 Deterministic Tracking	37
	2.4.2 Probabilistic Tracking	39
	2.4.3 Occlusion Handling	44
	2.4.4 Summary of Object Tracking	46
3	Semi-Constrained Gaussian Mixture Model for Background Sub-	
	traction	49
3.1	Introduction	49
3.2	Density Representation with Gaussian Mixture Model	50
3.3	Background Modeling using the Gaussian Mixture Model	52
	3.3.1 Background/Foreground Classification	56
	3.3.2 State of the Art and Current Shortcomings	57
	3.3.3 Analysis of Background Subtraction with GMM	58
3.4	Semi-Constrained Gaussian Mixture Model	64
	3.4.1 Mean Variable Learning Rate	65
	3.4.2 Standard Deviation Learning Rate	66
	3.4.3 Performance Analysis on Synthetic Data	67
3.5	Experiment Results	70
	3.5.1 Experimental Setup	70
	3.5.2 Controlled Environment	72
	3.5.3 Natural Changes in Illumination	78

3.6	Summary of the GMM for Background Modeling	80
4	Projective Kalman Filter for Vehicle Tracking	83
4.1	Introduction	83
4.2	Constraining the Tracking with the Environment	84
4.2.1	Motivations	85
4.2.2	Linear Fractional Transformation	86
4.3	The Kalman Filter	89
4.3.1	Closed-form Solution to the Bayesian Problem	90
4.3.2	The Extended Kalman Filter	91
4.3.3	The Unscented Kalman Filter	92
4.4	Projective Kalman Filter	93
4.4.1	State and Observation Updates	95
4.4.2	The Mean-shift Procedure	96
4.4.3	Extended versus Unscented Kalman Filter	97
4.5	Vehicle Tracking System	98
4.5.1	Tracker Initialization and Pruning	101
4.5.2	PKF Initialization and Vehicle Detection	101
4.6	Performance Analysis on Vehicle Tracking	103
4.6.1	Experimental Setup and Data	103
4.6.2	Comparison of the PKF and the EKF	105
4.6.3	Effects of the Frame Rate on Tracking	106
4.6.4	Mean-shift Convergence Speed at Low Frame Rates	108
4.7	Summary of the Projective Kalman Filter	111
5	Projective Particle Filter for Vehicle Tracking	113

5.1	Introduction	113
5.2	Sequential Monte Carlo and Particle Filtering	114
5.2.1	A Sub-optimal Bayesian Solution: The Particle Filter	116
5.2.2	Samples Degeneracy and Resampling	118
5.2.3	Particle Filter Summary	119
5.3	Projective Particle Filter	120
5.3.1	Importance Density and Prior	120
5.3.2	Likelihood Estimation	122
5.3.3	System Implementation	123
5.4	Experiments and Results	124
5.4.1	Mean Square Error Performance	125
5.4.2	Importance Sampling Evaluation	128
5.4.3	Tracking Performance and Discussion	128
5.5	Summary of the Projective Particle Filter	130
6	Tracking Through Occlusion with Markov Random Fields	131
6.1	Introduction	131
6.2	Integration of Contextual Information	132
6.2.1	Occlusion Handling	132
6.2.2	Importance of Contextual Information	133
6.2.3	Markov Random Fields	134
6.3	Gaussian Markov Random Field Mixture	137
6.3.1	Learning and Posterior Diffusion for Sparse Random Fields	139
6.3.2	Simulated Annealing	141
6.3.3	MRF Parameters Update	141
6.4	Performance Analysis and Discussion	142

6.4.1	Object Tracking System Implementation	142
6.4.2	Experimental Procedure	143
6.4.3	Mean Square Error Analysis	145
6.4.4	Performance with Total Spatio-temporal Occlusion	146
6.4.5	When Will the Algorithm Fail?	147
6.5	Summary of Tracking Through Occlusion	148
7	Abnormal Behavior Detection with Markov Random Fields	153
7.1	Introduction	153
7.2	Abnormal Behavior Modeling	154
7.3	Related Work	156
7.3.1	Object Descriptor Extraction	156
7.3.2	Activity Modeling	157
7.3.3	Complexity Reduction	158
7.3.4	Behavior Classification	158
7.4	Modeling Behavior with MRFs	159
7.4.1	Feature Vector Dimensionality Reduction	159
7.4.2	Integration of Contextual Information in the MRF	161
7.4.3	Stochastic Clustering Algorithm	162
7.5	Analysis of the Stochastic Learning Algorithm	164
7.5.1	Experimental Setup	164
7.5.2	Distance Measure Selection	166
7.5.3	Performance Analysis	171
7.6	Abnormal Behavior Detection on Highways	174
7.6.1	Experimental Setup	174
7.6.2	Performance Analysis	175

CONTENTS	xv
----------	----

7.6.3 Discussion	178
7.7 Summary of Abnormal Behavior Detection	180
8 Conclusions and Future Research	183
8.1 Thesis Summary	184
8.2 Suggestions for Improvements and Future Research	186
Bibliography	189

List of Figures

1.1	Video formation process	2
1.2	Video structure and representation	3
1.3	Scene projection and distortion	4
1.4	Fixed camera versus moving camera	6
1.5	Displays of an original video and its compressed version	7
1.6	Histogram representations of the spatial and temporal noise	7
2.1	Functional diagram of visual object tracking	14
2.2	Example of rectangular and elliptic shapes	16
2.3	Non-parametric representations of a person	18
2.4	Profile of the 1D and 2D Laplacian of Gaussians.	22
2.5	Maximization of the distance between two hyperplanes	27
2.6	Color histogram representation	29
2.7	Representation of the hidden Markov chain model	40
2.8	Three different types of occlusion	45
3.1	Pixel probability density represented by a mixture model	53
3.2	Original and foreground segmentation with saturated zone	60
3.3	Display of the pixel saturation phenomenon	61

3.4	Percentage of saturated pixels in a video sequence	61
3.5	Background adaptation time for a new mixture component	64
3.6	Performance on synthetic data	68
3.7	Estimated mean to true mean MSE	69
3.8	Number of foreground pixels under illumination changes	73
3.9	Foreground segmentation of the <i>HighwayII</i> video sequence	74
3.10	Foreground segmentation of the <i>People_Walking_1</i> video	75
3.11	Foreground segmentation for office scenes	77
3.12	Foreground segmentation in outdoor environment	79
3.13	Foreground segmentation in indoor environment	81
4.1	Examples of vehicle trajectories	86
4.2	Vehicle projection on the camera plane	87
4.3	Background subtraction on a low definition image	96
4.4	Contribution of the Hessian matrix $\mathcal{H}_{\mathbf{h}}$	99
4.5	Pixel position mean square error for EKF and UKF	100
4.6	Overview of the vehicle tracking algorithm with PKF	101
4.7	Example of tracking in dense vehicle flow	102
4.8	Sequence showing the drift of a tracker	106
4.9	Comparison of the and the proposed tracking algorithm	107
4.10	Effects of the frame rate on the tracking performances	107
4.11	Tracking rate for the PKF and the EKF	109
4.12	Tracking robustness in low frame rate	110
4.13	Mean-shift iterations for PKF and the EKF	111
5.1	Example of vehicle track for PKF and standard filter	125

5.2	Alignment of calculated and extracted trajectories	126
5.3	Position mean square error <i>vs.</i> number of particles	127
5.4	Position mean square error for 5 ground truth labeled vehicles	127
5.5	Position mean square error without resampling step	129
5.6	Drift tracking rate for projective and standard particle filters	129
6.1	Representation of vehicle motion by local mixture of Gaussians	135
6.2	Examples of neighborhoods in a graph.	136
6.3	Examples of cliques for the 8-neighborhood	136
6.4	MRFs update with integration and with diffusion	140
6.5	GMRFMPF and CONDENSATION tracking rates	145
6.6	Tracking with GMRFMPF and CONDENSATION through occlusion	149
6.7	Examples of pedestrian tracking through occlusion	150
6.8	Examples of vehicle tracking through occlusion (case A)	151
6.9	Examples of vehicle tracking through occlusion (case B)	152
7.1	Example of marginal densities of a feature vector	160
7.2	Example of generated vehicle tracks	165
7.3	ROC curves for ABD based on distance	167
7.4	ROC curves for ABD based on local density $p(r \Theta)$	168
7.5	ROC curves for ABD based on Mahalanobis distance measure	170
7.6	ROC curves of stochastic learning algorithm for ABD	171
7.7	ROC curves for the proposed technique and the SOM.	173
7.8	Examples of abnormal behavior on highways.	175
7.9	ROC curve for ABD on highway	176
7.10	Abnormal behavior detection rendering on real data	178

List of Tables

3.1	GMM Parameter Initializing Values	71
4.1	Vehicle Tracking Dataset	104
4.2	Vehicle Tracking System and PKF Parameter Initializing Values . . .	105
5.1	Linear Fractional Transformation Parameters	125
5.2	MSE for the Standard and the Projective Particle Filters	126
6.1	GMRFM Particle Filter Parameter Initializing Values	144
6.2	Comparison of the MSE for GMRFMPF and CONDENSATION . . .	146
6.3	Recovery Rate Under Occlusion	146
7.1	Correct ABD Rate with MRFs	173
7.2	Correct ABD Rate versus Size of SOM	174
7.3	Correct ABD Rate on the Video Dataset	177

List of Algorithms

- 3.1 Generic Gaussian Mixture Algorithm 55
- 4.1 Generic Projective Kalman Filter Algorithm 100
- 5.1 Resampling Algorithm 119
- 5.2 Projective Particle Filter Algorithm 123
- 6.1 GMRFM Particle Filter Algorithm 143

Nomenclature

ABD Abnormal Behavior Detection

ADABOOST Adaptive Boosting

ANN Artificial Neural Network

AVC Advanced Video Coding

BAC Breath Alcohol Content

CCD Charge-Coupled Device

CMOS Complementary Metal Oxide Semiconductor

CONDENSATION Conditional Density Propagation

DCT Discrete Cosine Transform

DUI Driving Under the Influence

DWT Discrete Wavelet Transform

EKF Extended Kalman Filter

EM Expectation-Maximization

EPF Extended Particle Filter

GMM Gaussian Mixture Model

GMPHDF Gaussian Mixture Probability Hypothesis Density Filter

GMRF Gaussian Markov Random Field

GMRFM Gaussian Markov Random Field Mixture

GMRFMPF Gaussian Markov Random Field Mixture Particle Filter

HMM Hidden Markov Model

JPDAF Joint Probability Data Association Filter

LOG Laplacian Of Gaussians

MAP Maximum A Posteriori

MCM Motion Correspondence Matrix

ML Maximum Likelihood

MLP Multi Layer Perceptron

MMSE Minimum Mean Square Error

MPDA Merged Probabilistic Data Association

MPEG Moving Picture Experts Group

MRF Markov Random Field

MSE Mean Square Error

OOP Object-Oriented Programming

PCA Principle Component Analysis

PCNSA Principal Component Null Space Analysis

pdf probability density function

PF Particle Filter

PHD Probability Hypothesis Density

PKF Projective Kalman Filter

PPF Projective Particle Filter

ROC Receiver Operating Characteristic

SIR Sampling Importance Resampling

SIS Sequential Importance Sampling

SOM Self Organizing Map

SSD Sum of Squared Differences

SVD Singular Value Decomposition

SVM Support Vector Machine

UKF Unscented Kalman Filter

UPF Unscented Particle Filter

UT Unscented Transform

WB White Balance