Fortschritt-Berichte VDI

VDI

Reihe 10

Informatik/
Kommunikation

Dipl.-Math. Oliver Müller,
Hannover

Nr. 860

Graphical Model MAP Inference with Continuous Label Space in Computer Vision

tnt

Institut für Informationsverarbeitung
www.tnt.uni-hannover.de

Graphical Model MAP Inference with Continuous Label Space in Computer Vision

Der Fakultät für Elektrotechnik und Informatik

der Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des akademischen Grades

Doktor-Ingenieur

genehmigte

Dissertation

von

Dipl. Math. Oliver Müller

geboren am 16. Juni 1986 in Berlin.

2018

Hauptreferent:	Prof. Dr.-Ing. Bodo Rosenhahn
Korreferent:	Prof. Dr. Bastian Leibe
Vorsitzender:	Prof. Dr.-Ing. Markus Fidler

Tag der Promotion: 20. September 2017

Fortschritt-Berichte VDI

Reihe 10

Informatik/
Kommunikation

Dipl.-Math. Oliver Müller,
Hannover

Nr. 860

Graphical Model MAP Inference with Continuous Label Space in Computer Vision

Institut für Informationsverarbeitung
www.tnt.uni-hannover.de

Müller, Oliver
Graphical Model MAP Inference with Continuous Label Space in Computer Vision
Fortschr.-Ber. VDI Reihe 10 Nr. 860. Düsseldorf: VDI Verlag 2018.
156 Seiten, 54 Bilder, 3 Tabellen.
ISBN 978-3-18-386010-4, ISSN 0178-9627,
€ 57,00/VDI-Mitgliederpreis € 51,30.

Für die Dokumentation: Maschinelles Sehen – Probabilistisch graphische Modelle – MAPInferenz – Markov-chain Monte-Carlo – Slice-Sampling – Produkt-Slice-Sampling – Artikulierte Objektverfolgung – Visuelle Objektverfolgung – Poseschätzung

This thesis deals with monocular object tracking from video sequences. The goal is to improve tracking of previously unseen non-rigid objects under severe articulations without relying on prior information such as detailed 3D models and without expensive offline training with manual annotations. The proposed framework tracks highly articulated objects by decomposing the target object into small parts and apply online tracking. Drift, which is a fundamental problem of online trackers, is reduced by incorporating image segmentation cues and by using a novel global consistency prior. Joint tracking and segmentation is formulated as a high-order probabilistic graphical model over continuous state variables. A novel inference method is proposed, called S-PBP, combining slice sampling and particle belief propagation. It is shown that slice sampling leads to fast convergence and does not rely on hyper-parameter tuning as opposed to competing approaches based on Metropolis-Hastings or heuristic samplers.

Bibliographische Information der Deutschen Bibliothek
Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliographie; detaillierte bibliographische Daten sind im Internet unter www.dnb.de abrufbar.

Bibliographic information published by the Deutsche Bibliothek
(German National Library)
The Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliographie (German National Bibliography); detailed bibliographic data is available via Internet at www.dnb.de.

Als Manuskript gedruckt. Printed in Germany.

ISSN 0178-9627

ISBN 978-3-18-386010-4

Acknowledgement

This thesis was written in the course of my activity as a scientific research assistant at the Institut für Informationsverarbeitung (TNT), Leibniz Universität Hannover.

First of all, I would like to thank my doctoral advisor Prof. Dr.-Ing. Bodo Rosenhahn for giving me the opportunity to work in his group, for many inspiring discussions, and for his continuing support and guidance. Many thanks are due to Prof. Dr. Bastian Leibe for being my second reviewer, and Prof. Dr.-Ing. Markus Fidler for serving as chair of my thesis defense committee.

I would also thank Prof. Dr.-Ing. Jörn Ostermann and Prof. Dr.-Ing. Bodo Rosenhahn for providing an outstanding working environment at the TNT.

Special thanks go to all of my colleagues at the TNT, who make this lab such a fun place to work. I especially would like to thank Dr.-Ing. Kai Cordes for his guidance during my study and diploma thesis, which motivated me to start working towards a doctoral degree. I also thank Prof. Dr.-Ing. Michael Ying Yang for his support and for sharing his deep knowledge on probabilistic graphical models with me. I further thank my officemate Karsten Vogt for many inspiring, and constructive discussions.

Special thanks to Matthias Schuh, Dr.-Ing. Martin Pahl, Dr.-Ing. Marco Munderloh, and Thomas Wehberg for their outstanding technical and administrative support, as well as to the secretaries, Hilke Brodersen, Doris Jaspers-Göring, Melanie Huch, and Pia Bank for their invaluable administrative work. This institute would not have been such an outstanding place to work without their continuing efforts.

Last but not least, special thanks go to my family and my friends. Without the unconditional support of my parents Rosel and Michael Müller, this work would not have been possible.

Contents

Abbreviations

2D	two-dimensional
3D	three-dimensional
ADMM	alternating direction method of multipliers
BP	belief propagation
DAG	directed acyclic graph
DD	dual decomposition
DPM	deformable parts model
DPMP	diverse particle max-product
FMP	flexible mixtures-of-parts
HOG	histogram of oriented gradients
KL	Kullback-Leibler divergence
LP	linear program
MAP	maximum a posteriori
MATLAB®	Matrix Laboratory, a proprietary programming language and IDE developed by MathWorks
MCMC	Markov chain Monte-Carlo
MH	Metropolis-Hastings
MH-PBP	Metropolis-Hastings particle belief propagation
MP-BP	max-product belief propagation
MRF	Markov random field
MuPAD®	a computer algebra system bundled with MATLAB®
OCT	optical coherence tomography
OTB	online tracking benchmark
PBP	particle belief propagation
PCP	percentage of correct parts
PGM	probabilistic graphical model
RGB	red, green, and blue

RMSD	root-mean-square deviation
S-PBP	slice-sampling particle belief propagation
SVM	support vector machine
TRBP	tree-reweighted belief propagation
VDPM	visibility-aware deformable parts model
VDPM-e	visibility-aware deformable parts model without edge
VOT	visual object tracking

Symbols and Notation

Symbols for Probability Theory

$\mathcal{N}(x; \mu, \sigma)$	normal distribution with mean μ and standard deviation σ
Ω	sample space
P	probability distribution
$p(x)$	probability density function
$\mathbf{X}$	random variable vectors
X, Y, Z	random variables
$\mathcal{X}, \mathcal{Y}, \mathcal{Z}$	state space
x	random variable values
$\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$	random variable value vectors

Symbols for Probabilistic Graphical Models

$\mathcal{C}$	set of cliques
c	clique
$\mathbf{d}$	data
$\mathcal{E}$	edges
$E(x)$	energy function
$\mathcal{F}$	factor vertices
$\mathcal{G}$	graphical model
$g(\lambda)$	dual function
k, l	state indices

$\mathcal{L}$	local polytope
L_s	number of states for discrete state space $\mathcal{X}_s$
$\mathcal{L}(\{\mathbf{y}_\tau\}_\tau,\mathbf{y},\lambda)$	(augmented) Lagrangian function
$\lambda \in \Lambda$	Lagrange multipliers
$\mathcal{M}$	marginal polytope
$m_{t\to s}(x_s)$	belief propagation message from vertex t to s
$\overline{m}_{t\to s}(x_t)$	pre-message from vertex t to s
$\hat{m}_{t\to s}(x_s)$	approximate belief propagation message from vertex t to s
$\hat{\mu}_s(x_s)$	approximate max-marginal function over variable x_s
$\mu_s(x_s)$	max-marginal function over variable x_s
$\mathcal{N}$	neighborhood system
$n = 1, ..., N$	BP/DD/ADMM iteration
$\mathrm{Pa}(s)$	parents of vertex s
$\mathcal{P}_s$	particle set for vertex s
$\phi_s(x_s), \phi_{s,t}(x_s, x_t), \phi_c(x_c)$	unary potential, pairwise potential, and clique potential
$\psi_s(x_s), \psi_{s,t}(x_s, x_t), \psi_c(x_c)$	unary energy, pairwise energy, and clique energy
r, s, t	vertex indices
$\mathcal{S}$	visiting schedule
$\tau \in \mathcal{T}$	subproblems
θ	parameter vector
$\mathcal{V}$	random variable vertices
Z	partition function

Symbols for MCMC

A	slice

$i = 1, ..., p$	particle index
$m = 1, ..., M$	Markov chain index
$\mu(x)$	MCMC target density function
$\nu_0(x)$	MCMC initial density function
$\nu(x)$	marginal probability
$x_s^{(i)\langle m\rangle}$	particle at the m-th MCMC iteration
$\bar{x}_s^{(i)\langle m\rangle}$	candidate particle
$q(x \mid y)$	MCMC proposal density function
$q_\sigma(x \mid y)$	Gaussian proposal density with mean y and standard deviation σ
$T(x, y)$	MCMC transition kernel
$\mathcal{U}$	auxiliary state space
$\mathfrak{U}(A)$	uniform distribution over region A
u	auxiliary variable
Z	normalization constant
$\{u^{\langle m\rangle}\}_m$	Markov chain auxiliary variables
$\{x_s^{(i)}\}_i$	particle
$\{x^{\langle m\rangle}\}_m$	Markov chain variables

Symbols for Part-based Object Tracking

$E_{\text{con}}(\mathbf{p}, \mathbf{y})$	global consistency energy
$E_{\text{DPM}}(\mathbf{p})$	deformable parts model energy
$E_{\text{seg}}(\mathbf{y})$	segmentation energy
$E_{\text{VDPM}}(\mathbf{p}, \mathbf{v})$	visibility-aware deformable parts model energy
$\mathbb{I}[\,\cdot\,]$	indicator function
$\mu(\mathbf{p})$	patch mask function

$\boldsymbol{\mu}$	patch mask matrix
$\mathbf{p}$	vector of object poses
$\mathbf{v}$	visibility v_i for each DPM patch i
$\mathbf{x}$	indicator vector over object poses
$\mathcal{X}^{\text{pose}}$	pose state space
$\mathcal{X}^{\text{seg}}$	segmentation state space
$\mathbf{y} \in \{0,1\}^N$	foreground/background segmentation of an image with N pixels

Abstract

This thesis deals with the development and improvement of maximum a posteriori (MAP) inference approaches in probabilistic graphical models (PGMs) and their application on challenging computer vision problems.

Many challenging computer vision tasks are modeled as MAP inference problems in PGMs. MAP inference is the problem of finding the most probable configuration of random variables for a given target problem in the exponentially large space of possible outcomes. PGMs are a family of powerful modeling languages which unify two fundamental concepts: uncertainty and graphical models. Many real-world phenomena can be modeled in form of probability distributions over continuous-valued random variables. A PGM is a language to model these distributions which typically involve a very large number of random variables. Conditional independence assumptions of the random variables play a key role in retrieving tractable models.

In the first part of this thesis, a general purpose framework for MAP inference in PGMs over continuous-valued random variables based on stochastic inference methods is developed. A novel approach, the slice-sampling particle belief propagation (S-PBP) algorithm, is developed which achieves more accurate and faster MAP estimates than heuristic sampling or Metropolis-Hastings sampling approaches. The proposed approach generates sample proposals from the max-marginal distributions using the slice sampling algorithm. By exploiting the message-passing nature of the applied MAP inference approach, the dependence on hyper-parameters is reduced and a significant speedup is achieved.

The second part of this thesis is dedicated to the application of the developed inference approaches to computer vision applications. Hereby, the main focus is in online tracking of articulated objects. The visual tracking of previously unseen objects in videos or video streams is a fundamental task in computer vision. A novel framework is proposed for part-based object tracking. The problem of automatic model initialization and the reduction of tracker drift by incorporating higher-order constraints and image segmentation cues to the tracker is addressed. A global consistency prior is proposed which enables inference of both part-based tracking and image segmentation in a joint probabilistic model. Experiments show that the joint formulation leads to improved image segmentation results as well as reduced drift in online object tracking.

Keywords: computer vision, probabilistic graphical models, MAP inference, Markov-chain Monte-Carlo, slice sampling, product slice sampling, articulated online tracking, visual object tracking, pose estimation

Kurzfassung

Das Ziel dieser Arbeit ist die Erstellung und Verbesserung von Optimierungsverfahren zur Inferenz in probabilistischen graphischen Modellen (PGMs) und deren Anwendung auf Probleme im Bereich Computer-Vision.

Viele Computer-Vision Probleme werden heutzutage als MAP-Inferenz Probleme in PGMs behandelt. MAP-Inferenz beschreibt hierbei das Finden der wahrscheinlichsten Kombination von Werten im exponentiell wachsenden Raum möglicher Lösungen eines Problems. Probabilistische graphische Modelle sind eine Familie von Modellierungssprachen zur Beschreibung von Verbundwahrscheinlichkei-ten über einer Menge von Zufallsvariablen. Hierbei werden zwei grundlegende Prinzipien miteinander vereint: Die Modellierung von Unsicherheiten und die Modellierung (bedingter) Unabhängigkeit von Zufallsvariablen mittels Knoten und Kanten in einem Graph.

Der erste Teil dieser Arbeit behandelt das Problem der MAP-Inferenz bei reellwertigen Zufallsvariablen mit Hilfe stochastischer Inferenzverfahren. Slice-sampling particle belief propagation (S-PBP) ist ein neu entwickelter Ansatz, welcher eine genauere MAP-Schätzung in kürzerer Zeit erlaubt als andere stochastische Verfahren. Eine Kernkomponente stochastischer Suchverfahren ist die Erzeugung von Stichproben im Lösungsraum. Bisherige Verfahren sind entweder heuristisch motiviert oder verwenden Vorschlagsverteilungen deren Parameter Problem-abhängig eingestellt werden müssen. Der in dieser Arbeit vorgestellte Ansatz erzeugt Stichproben direkt aus den Max-Marginal-Verteilungen des graphischen Modells mit Hilfe des Slice-Sampling Verfahrens. Durch Ausnutzung des Message-Passing Mechanismus der verwendeten Optimierungsverfahren wird die Parameter-Abhängigkeit verringert und eine Beschleunigung des Verfahrens erreicht.

Im zweiten Teil der Arbeit werden die zuvor entwickelten Inferenzverfahren auf Probleme im Bereich Computer-Vision angewandt. Das Hauptaugenmerk liegt hierbei auf der Online-Verfolgung artikulierter Objekte in Videosequenzen. Die visuelle Verfolgung zuvor unbekannter Objekte in Videos ist ein fundamentales Problem des maschinellen Sehens. Es wird ein neuer Ansatz zur Teile-basierten Objektverfolgung entwickelt. Das Problem der automatischen Modellinitialisierung und der Reduktion des Driftens in Teile-basierten Modellen wird durch die Integration von Bildsegmentierung und zusätzlicher Bedingungen höherer Ordnung behandelt. Es wird ein globaler Konsistenzterm vorgeschlagen, der die Teile-basierte Poseschätzung und die Bildsegmentierung in einem gemeinsamen probabilistischen Modell vereint. Experimente zeigen das die gemeinsame Schätzung von Pose und Segmentierung zum einen die Bildsegmentierung verbessert, als auch das Driften der geschätzten Pose effektiv verringert.

Stichworte: maschinelles Sehen, Probabilistische graphische Modelle, MAP-Inferenz, Markov-chain Monte-Carlo, Slice-Sampling, Produkt-Slice-Sampling, Artikulierte Objektverfolgung, Visuelle Objektverfolgung, Poseschätzung