VDI

REIHE 10
INFORMATIK/
KOMMUNIKATION

Fortschritt-Berichte VDI

Dipl.-Math. Roberto D. Henschel,
Hannover

NR. 875

Higher-Order Multiple Object Tracking

BAND
1|1

VOLUME
1|1

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

Higher-Order Multiple Object Tracking

Von der Fakultät für Elektrotechnik und Informatik
der Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des akademischen Grades
Doktor-Ingenieur
(abgekürzt: Dr.-Ing.)
genehmigte Dissertation

von

Dipl.-Math. Roberto D. Henschel
geboren am 10. September 1985 in Berlin, Deutschland

2021

1. Referent:	Prof. Dr.-Ing. Bodo Rosenhahn
2. Referent:	Prof. Dr.-Ing. habil. Christian Heipke
Vorsitzender:	Prof. Dr.-Ing. Jörn Ostermann

Tag der Promotion: 01. Oktober 2021

VDI

REIHE 10
INFORMATIK/
KOMMUNIKATION

Fortschritt-Berichte VDI

Dipl.-Math. Roberto D. Henschel,
Hannover

NR. 875

Higher-Order Multiple Object Tracking

BAND
1|1

VOLUME
1|1

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

VDI verlag

Henschel, Roberto D.

Higher-Order Multiple Object Tracking

Fortschritt-Berichte VDI, Reihe 10, Nr. 875. Düsseldorf: VDI Verlag 2021.

212 Seiten, 67 Bilder, 16 Tabellen.

ISBN 987-3-18-387510-7, ISSN 0178-9627,

76,00 EUR/VDI-Mitgliederpreis: 68,40

Für die Dokumentation: Verfolgung Mehrerer Objekte – Binäres Lineares Programm (BLP) – Binäres Quadratisches Programm (BQP) – Datenassoziationsmodel Höherer Ordnung – Globale Optimierung – Video – Sensorfusion – Inertiale Messeinheit

Keywords: Multiple Object Tracking – Binary Linear Programming (BLP) – Binary Quadratic Programming (BQP) – Higher-Order Data Association Model – Global Optimization – Video – Sensor Fusion – Inertial Sensors

This dissertation deals with camera-based offline multiple object tracking and explores higher-order data association models. Due to their extensive exploitation of the available information, such models are promising approaches in current research. However, they commonly represent NP-hard optimization problems so that their application in practice is challenging.
The first part of this thesis proposes a binary quadratic program that enables to globally fuse signals within a higher-order data association model. This enables to overcome weaknesses of the individual signals. An approximate solver based on the Frank-Wolfe algorithm is presented and analyzed. Its benefit is demonstrated in two setups: fusion of two detectors and combining signals coming from a video and body-worn inertial measurement units. The second part of this thesis proposes an extension of the disjoint path model by higher-order information and connectivity priors, resulting in a binary linear program. Efficient separation algorithms are proposed and integrated into a cutting-plane algorithm, making it possible for the first time to solve higher-order data association globally in practice.

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.

ISBN 987-3-18-387510-7, ISSN 0178-9627

Acknowledgments

This thesis was written in the course of my activity as a scientific research assistant at the *Institut für Informationsverarbeitung* of the Leibniz University Hannover.

First of all, I would like to thank my doctoral advisor Prof. Dr.-Ing. Bodo Rosenhahn for supporting me to become a researcher in the very exciting field of computer vision. I am thankful for the stimulating discussions about work, society, and mathematical topics, the great supervision, and all the support and freedom I received to pursue my ideas.

Also, many thanks to him and Prof. Dr.-Ing. Jörn Ostermann for the excellent working environment. I also thank Prof. Dr.-Ing. habil. Christian Heipke for being the second examiner and Prof. Dr.-Ing. Jörn Ostermann for being the chair of the defense committee.

Special thanks go to Prof. Dr.-Ing. Laura Leal-Taixé for her incredible support and deep discussions, which greatly helped to advance my research and this thesis.

Also, I would like to thank Andrea Hornakova and Dr. Paul Swoboda for the great collaboration, which was very inspiring, successful, and fun.

I would also like to thank Prof. Dr. Konrad Schindler for the opportunity to stay at his research lab, which was really inspiring.

The time at the TNT institute was amazing thanks to all the great colleagues. My special recognition goes to my former office mate Dr.-Ing. Timo von Marcard for making our office such a great place. Having deep discussions about research, mathematics, and non-work topics was great. Also, collaborating and traveling together was outstanding as was all the support I received during the time when I was finishing my thesis. I would like to thank my office mate Timo Kaiser for the excellent time with lots of discussions about anything and for the fun working together on multiple object tracking. Also, I thank Dr.-Ing. Bastian Wandt for many deep discussions and for the joint building of a company. Special thanks go to my colleagues Dr.-Ing. Michele Fenzi, Leonid German, Dr.-Ing. Alina Kutznetsova, Prof. Dr.-Ing. Laura Leal-Taixé, Yasser Samayoa, and Dr.-Ing. Dipl.-Math. techn. Aron Sommer for the great fun at the institute, *e.g.*, table football tournaments, card games or long discussions, and all the fun inside or outside the institute, making the time together unforgettable. Finally, I would like to thank the TNT administrative staff for their technical and administrative support.

I thank my friend Gad Kohls for advising me to study mathematics.

Last but not least, I dedicate a special thanks to my family. Without the unconditional support from my parents Clara and Ramon, as well as my sister Ruth, accomplishing my degrees and this thesis would not have been possible.

Contents

Acronyms

2D	*two-dimensional*
3D	*three-dimensional*
ACF	*Aggregate Channel Features*
Acc	*Accuracy*
BCE	*Binary Cross-Entropy*
BLP	*Binary Linear Program*
BQP	*Binary Quadratic Program*
CE	*Cross Entropy*
CNN	*Convolutional Neural Network*
CS	*Cosine Similarity*
DM	*DeepMatching*
DP	*Disjoint Paths*
DPM	*Deformable Part Model*
FN	*False Negatives*
FP	*False Positives*
fps	*frames per second*
FRCNN	*Faster R-CNN*
FW	*Frank-Wolfe*
FWT	*Frank-Wolfe Tracker*
GAP	*Duality gap*
GPS	*Global Positioning System*

HO-MOT *Higher-Order Multiple Object Tracking*

HOG *Histogram of Oriented Gradients*

ID *Identity*

IDF1 *ID F1*

IDP *ID Precision*

IDR *ID Recall*

IDS *ID Switches*

IMU *Inertial Measurement Unit*

IoU *Intersection over Union*

KLT *Kanade–Lucas–Tomasi feature tracker*

LDP *Lifted Disjoint Paths*

LIDAR *Light Detection and Ranging*

Lif_T *Lifted Disjoint Paths Tracker*

Lif_TsimInt *Lifted Disjoint Paths Tracker using simple linear interpolation*

LP *Linear Program*

mAP *mean Average Precision*

ML *Mostly Lost*

MOT *Multiple Object Tracking*

MOTA *Multiple Object Tracking Accuracy*

MPT *Multiple People Tracking*

MSE *Mean Squared Error*

MT *Mostly Tracked*

PC *Perspective Correction*

Prec *Precision*

PT *Partially Tracked*

QP *Quadratic Program*

RADAR *Radio Detection And Ranging*

ReLU *Rectified Linear Unit*

RPN *Region Proposal Network*

SDP	*Scale-Dependent Pooling*
SVM	*Support Vector Machine*
TPR	*True Positive Rate*
TNR	*True Negative Rate*
VHN	*Visual Heading Network*
VIMPT	*Video Inertial Multiple People Tracking*
VIT	*Video Inertial Tracker*

Notation

Numbers and Arrays

a	A scalar
A	A set
$\mathbf{a}$	A vector
$\mathbf{A}$	A matrix
$\mathbf{A}^\intercal$	Transpose of matrix $\mathbf{A}$
$\mathbf{A}^{-1}$	Inverse of square matrix $\mathbf{A}$
$\langle \mathbf{a}, \mathbf{b} \rangle$	Scalar product of $\mathbf{a}$ and $\mathbf{b}$
$\mathbf{a} \star \mathbf{b}$	Convolution of $\mathbf{a}$ and $\mathbf{b}$
$\frac{df}{dx}$	Derivative of f with respect to x
$\frac{d^2 f}{dx^2}$	Second derivative of f with respect to x
∇f	Gradient of f
$\lfloor x \rfloor$	Integer part of x
$\|\|\mathbf{a}\|\|$	L^2-norm of $\mathbf{a}$
$\det(\mathbf{A})$	Determinant of $\mathbf{A}$
$\mathbf{I}$	Identity matrix
$\mathbf{1}$	Matrix of ones
$\mathbf{0}$	Zero matrix
$\mathbb{E}[X]$	Expectation of random variable X
$[n]$	Set of natural numbers from 1 to n
$[n]_0$	Set of natural numbers from 0 to n
$[n_1 : n_2]$	Set of natural numbers from n_1 to n_2

Symbols

$\mathfrak{f}$	Frame index
n_R	Number of frames in a recording
R	Set of frames indices of a recording

$\mathbf{d} \in D$	A detection $\mathbf{d}$ in a set of detections D
$D_\mathfrak{f}$	Set of all detections in frame $\mathfrak{f}$
$\gamma \in \Gamma$	A trajectory γ in a set of trajectories Γ
$\text{supp}(\gamma)$	Set of frames for which trajectory γ contains detections
ϕ	Unary feature
ψ	Pairwise feature
$\mathcal{P}$	Deterministic polynomial time complexity class
$\mathcal{NP}$	Nondeterministic polynomial time complexity class
L	A loss
$\wedge$	Logical AND
$\vee$	Logical OR
$\mathcal{O}(n)$	Big O notation
P	Probability measure

Graphs

$\mathcal{G}$	A graph
$v \in \mathcal{V}$	A vertex v in a vertex set $\mathcal{V}$
$e \in \mathcal{E}$	An edge e in an edge set $\mathcal{E}$
$\mathcal{G}_\mathcal{V}$	The vertex set of graph $\mathcal{G}$
$\mathcal{G}_\mathcal{E}$	The edge set of graph $\mathcal{G}$
n_{nod}	Number of nodes
P	A path
$vw\text{-paths}(\mathcal{G})$	The set of paths in $\mathcal{G}$ starting at v and ending in w
$\mathcal{G}[\widetilde{\mathcal{V}}]$	The subgraph of $\mathcal{G}$ induced by the vertex set $\widetilde{\mathcal{V}}$
$d_G(v)$	Neighborhood of node v within the graph $\mathcal{G}$
$\mathcal{V}_\mathfrak{f}$	All nodes at frame $\mathfrak{f}$
$\mathbf{c}$	Vertex weights
$\mathbf{q}$	Edge weights
$\breve{\mathcal{G}} = (\breve{\mathcal{V}}, \breve{\mathcal{E}})$	Lifted graph $\breve{\mathcal{G}}$ with edge set $\breve{\mathcal{E}}$ and vertex set $\breve{\mathcal{V}}$
$\breve{q}_{\breve{e}}$	Weight of lifted edge $\breve{e}$
$\mathcal{R}$	Reachability relation
n_{obj}	Number of labels

Optimization

P	Polyhedron
$\text{P}_\mathcal{B}$	Polyhedron P with additional binary constraints

$\mathrm{P}_{\mathcal{B}}^{\circ}$	Continuous relaxation of $\mathrm{P}_{\mathcal{B}}$
$\mathrm{P}(\mathbf{A}, \mathbf{b})$	A polyhedron in canonical $\mathcal{H}$-representation
$\mathrm{conv}(A)$	Convex hull of a set A
$H^{\leq}$	Closed half-space
$\mathbf{x}, \mathbf{y}$	Binary indicator variables
$[v \rightharpoonup k]$	Linear index to indicator variable for the assignment of node v to label k
$\mathrm{WGL}(\mathcal{G})$	Weighted graph labeling problem defined on graph $\mathcal{G}$
$\mathrm{WGL}_{\mathrm{NMS}}(\mathcal{G})$	Problem $\mathrm{WGL}(\mathcal{G})$ with additional non-maxima suppression
$\mathrm{RWGL}(\mathcal{G})$	Continuous relaxation of $\mathrm{WGL}(\mathcal{G})$
$\mathrm{P}_{\mathcal{B}}(\mathcal{G})$	Underlying polyhedron of problem $\mathrm{WGL}(\mathcal{G})$
$\mathrm{P}_{\mathcal{B}}^{\mathrm{NMS}}(\mathcal{G})$	Underlying polyhedron of problem $\mathrm{WGL}_{\mathrm{NMS}}(\mathcal{G})$

Abstract

This dissertation deals with methods for camera-based multiple object tracking (MOT). More precisely, the task is to compute the association between objects of a specified class and corresponding image contents of a video recording. To tackle this extremely difficult problem, the so-called tracking-by-detection paradigm is usually employed: First, object detections are generated for the entire recording. Then, an association between detections and objects is computed. Finding the correct assignment is called the data association problem. A solution to the problem provides the trajectories for the desired objects.

Since the tracking-by-detection paradigm is computed sequentially, errors of the detector as well as wrongly assessed temporal consistencies between detections can lead to propagation of errors. Therefore, the employed data association model substantially determines the accuracy of the computed trajectories. In order to achieve highly accurate results, this thesis focuses on building robust data association models. At the same time, standard detectors are used to meaningfully compare the corresponding results with previous approaches.

Many established methods use data association models that exploit temporal consistency only between detections that directly follow each other in a trajectory. However, such simple models are highly susceptible to the errors mentioned above. This thesis presents more robust tracking methods by using higher-order data association models (higher-order multiple object tracking). Here, all pairs of detections associated with a trajectory, and not only the consecutive ones, contribute to the evaluation of the consistency of a trajectory. To comprehensively exploit the entire information of a video recording, this thesis formulates two data association models, each as a global optimization problem. Accordingly, the recordings are processed offline, using all information available. However, the underlying optimization problems are $\mathcal{NP}$-hard, which makes it difficult to compute good solutions. A suitable optimization method is presented for each proposed data association model. The corresponding optimizer yields in practice near-optimal or even (to the best of our knowledge for the first time) provable global optimal solutions, depending on the model used.

In the first part of this thesis, a method for improved utilization of the signals available at a point in time is presented. While the standard tracking approach uses only object detections, the proposed method allows multiple input signals to be fused globally for the tracking task. A higher-order data association model is proposed that evaluates consistency within a signal as well as between different signals. By using complementary signals, weaknesses of individual signals can be compensated and advantages can be combined. The proposed data association model is based on an $\mathcal{NP}$-hard weighted graph labeling problem. Due to the complexity of the problem, computing an optimal solution is difficult. A suitable approximate optimization method for the graph labeling problem is presented. Evaluations show that near-optimal solutions are generated with this method. The benefits of the fusion are analyzed using two applications of the graph-labeling formulation. (i) To exploit more image information, person detections are combined with head detections. It is shown that the fusion achieves much better results than when only using person detections. In particular, the fusion helps to detect and remove false-positive person detections, as these often do not have matching head

detections. In addition, the fusion approach results in persons being tracked for longer periods of time, since in the case of missing person detections, head detections can be used to locate and track persons. (ii) A video is fused with inertial measurement units (IMU). For this purpose, it is assumed that each person to be tracked wears an IMU on his or her back. Acceleration and orientation measurements from the IMUs are linked with corresponding values estimated from the video recording. The corresponding graph labeling problem generates trajectories that are temporally consistent with respect to the video recording and the IMU signals. The fusion leads to significantly better trajectory results compared to purely video-based MOT methods, especially when the visual information is impaired (*e.g.*, due to motion blur or similarly dressed persons). Missing detections can be reconstructed very robustly by the fusion so that the method has a lower dependence on the quality of the detections compared to purely video-based methods. In addition, the proposed fusion allows people to be identified in the image since each trajectory is associated with an IMU. Overall, the methods from the first part of this thesis demonstrate that the proposed fusion formulation enables to exploit provided data more extensively by being able to process more image information and integrate more signals. This substantially improves tracking accuracy.

Nonetheless, object detections provide valuable information not fully exploited by existing methods. Higher-order data association models are either not used at all due to their complexity or are based on heuristic optimization methods. Both cases can lead to false associations.

In contrast, in the second part of this thesis, an optimization method is presented that allows for the first time to solve a suitable higher-order data association model by means of global optimization despite being $\mathcal{NP}$-hard. This enables to exploit long-term temporal information and long-range temporal interactions. To this end, a novel data association model is proposed and described as a binary linear program. Efficient separation algorithms are presented to solve the optimization problem within a cutting-plane method. The global optimization enabled the method to outperform the state of the art on all tested datasets by a large margin. In addition, conducted experiments show that the method benefits significantly from the use of long-term information. On the datasets used, the presented method leads to nearly optimal assignment accuracies for given detections. Future work can therefore focus on other areas, such as a more accurate extraction of detections. Overall, the second part of this thesis shows that improved exploitation of the information provided over time leads to substantial improvements in trajectory results. For the first time, a global optimization method has been successfully used to solve higher-order data association models.

Keywords: Multiple Object Tracking, Video, Higher-Order Data Association Models, Sensor Fusion, Binary Linear Program, Binary Quadratic Program, Global Optimal Solution

Kurzfassung

Diese Dissertation befasst sich mit Verfahren zur kamerabasierten Verfolgung mehrerer Objekte (Multiple Object Tracking, abgekürzt MOT). Genauer besteht die Aufgabe in der Zuordnungsberechnung zwischen Objekten einer gewählten Objektklasse und zugehörigen Bildinhalten einer Videoaufnahme. Um dieses äußerst schwere Problem anzugehen wird für gewöhnlich das sogenannte Tracking-durch-Detektionen Paradigma verwendet: Als Erstes werden für die gesamte Aufnahme Objektdetektionen erzeugt. Im zweiten Schritt werden die Zuordnungen zwischen Detektionen und Objekten berechnet. Das Finden der korrekten Zuordnungen wird als Datenassoziationsproblem bezeichnet. Aus der Lösung ergeben sich die Trajektorien der Objekte.

Da das Tracking-by-Detection-Paradigma sequentiell berechnet wird, können sowohl Fehler des Detektors, als auch falsch bewertete zeitliche Konsistenzen zwischen Detektionen zu Fehlerfortpflanzungen führen. Entsprechend bestimmt das verwendete Datenassoziationsmodell die Genauigkeit der Trajektorien wesentlich. Um möglichst genaue Resultate zu erreichen, fokussiert sich diese Arbeit auf die Erforschung robuster Datenassoziationsmodelle. Gleichzeitig werden Standard-Detektoren verwendet, um die entsprechenden Ergebnisse aussagekräftig mit vorherigen Ansätzen vergleichen zu können.

Viele der etablierten Verfahren nutzen für die Datenassoziation lediglich die zeitliche Konsistenz zwischen Detektionen aus, welche in einer Trajektorie direkt aufeinanderfolgen. Solch einfache Modelle sind jedoch stark anfällig gegenüber den erwähnten Fehlern. In dieser Arbeit werden robustere Tracking-Verfahren durch die Verwendung von Datenassoziationsmodellen höherer Ordnung (Higher-Order Multiple Object Tracking) präsentiert. Dabei tragen nicht nur aufeinanderfolgende, sondern alle Paarungen von Detektionen, welche einer Trajektorie zugeordnet werden, zur Bewertung der Konsistenz einer Trajektorie bei. Um die Gesamtinformationen einer Videoaufnahme umfassend auszunutzen, formuliert diese Arbeit die Datenassoziationsmodelle als globale Optimierungsprobleme. Entsprechend werden die Aufnahmen offline, unter Verwendung aller verfügbaren Informationen verarbeitet. Die zugehörigen Optimierungsprobleme sind jedoch $\mathcal{NP}$-schwer. Entsprechend ist es schwierig, gute Lösungen zu berechnen. Zu jedem vorgeschlagenen Datenassoziationsmodel wird ein passendes Lösungsverfahren präsentiert, welches in der Praxis je nach verwendetem Model nahezu optimale oder sogar (nach unserem Wissen erstmalig) beweisbar global optimale Ergebnisse liefert.

Im ersten Teil dieser Arbeit wird ein Verfahren zur verbesserten Ausnutzung der zu einem Zeitpunkt bereitstehenden Signale präsentiert. Während der Standard-Tracking-Ansatz nur Objektdetektionen verwendet, erlaubt die vorgeschlagene Methode, mehrere Eingangssignale global für die Tracking-Aufgabe zu fusionieren. Ein Datenassoziationsmodel höherer Ordnung wird vorgeschlagen, bei dem die Konsistenz sowohl innerhalb eines Signals, als auch zwischen verschiedenen Signalen bewertet wird. Durch die Verwendung komplementärer Signale können Schwächen einzelner Signale kompensiert und Vorteile kombiniert werden. Das vorgeschlagene Datenassoziationsmodell basiert auf einem $\mathcal{NP}$-schweren gewichteten Graph-Labeling Problem. Auf Grund der Komplexität des Problems ist die Berechnung einer Optimallösung schwierig. Es wird ein dafür passendes, approximatives Optimierungsverfahren für das Graph-Labeling Problem vorgestellt. Die Evaluierungen zeigen, dass damit nahezu optimale Lösungen

erzeugt werden. Der Nutzen der Fusionierung wird anhand von zwei Anwendungen der Graph-Labeling Formulierung analysiert. (i) Um mehr Bildinformationen auszunutzen, werden Personendetektionen mit Kopfdetektionen kombiniert. Es zeigt sich, dass durch die Fusion wesentlich bessere Ergebnisse als bei ausschließlicher Verwendung von Personendetektionen erreicht werden. Insbesondere hilft es falsch-positive Personendetektionen zu erkennen und entfernen, da diese oft keine passenden Kopfdetektionen haben. Ferner führt der Fusionierungssansatz zu einer längeren Verfolgung von Personen, da im Falle von fehlenden Personendetektionen die Kopfdetektionen zur Lokalisierung und Verfolgung der Personen verwendet werden. (ii) Es werden Inertialmesseinheiten (IMUs) mit einer Videoaufnahme fusioniert. Dazu wird angenommen, dass jede zu verfolgende Person eine IMU am Rücken trägt. Es werden die Beschleunigungs- und Orientierungsmessungen der IMUs mit entsprechend aus der Videoaufnahme geschätzten Werten gekoppelt. Das entsprechende Graph-Labeling Problem erzeugt Trajektorien, welche sowohl zeitlich konsistent bezüglich der Videoaufnahme, als auch zu den zugehörigen IMU-Signalen sind. Die Fusionierung führt insbesondere dann zu erheblich besseren Trajektorienergebnissen gegenüber rein videobasierten MOT-Verfahren, wenn die visuellen Informationen beeinträchtigt sind (etwa durch Bewegungsunschärfe sowie bei ähnlich gekleideten Personen). Fehlende Detektionen lassen sich durch die Fusionierung sehr robust rekonstruieren, sodass die Methode eine geringere Abhängigkeit von der Qualität der Detektionen, verglichen mit rein videobasierten Verfahren, aufweist. Darüber hinaus ermöglicht die Fusionierung, Personen im Bild zu identifizieren, da jede Trajektorie einer IMU zugeordnet wird. Insgesamt zeigen die Methoden aus dem ersten Teil dieser Arbeit, dass die vorgeschlagene Fusionsformulierung es ermöglicht, bereitgestellte Daten umfassender zu nutzen, da mehr der vorhandenen Bildinformationen und Signale integriert werden können. Dadurch verbessert sich die Tracking-Genauigkeit erheblich.

Nichtsdestotrotz stellen Objektdetektionen wertvolle Informationen bereit, welche durch bestehende Verfahren nicht komplett ausgenutzt werden. Datenassoziationsmodelle höherer Ordnung werden auf Grund der Komplexität entweder gar nicht verwendet oder basieren auf heuristischen Optimierungsverfahren, wodurch falsche Zuordnungen erzeugt werden können.

Im Gegensatz dazu wird im zweiten Teil dieser Arbeit ein Optimierungsverfahren vorgestellt, welches es erstmalig ermöglicht, ein dafür passendes Datenassoziationsmodell höherer Ordnung mittels globaler Optimierung zu lösen, wodurch sich insbesondere Langzeitinformationen und -interaktionen ausnutzen lassen, obwohl das Problem $\mathcal{NP}$-schwer ist. Dazu wird ein neues Datenassoziationsmodell vorgeschlagen und als ein binäres lineares Programm beschrieben. Es werden effiziente Separierungsalgorithmen vorgestellt, um das Optimierungsproblem mittels eines Schnittebenenverfahrens zu lösen. Durch die globale Optimierung konnte das Verfahren auf allen getesteten Datensätzen den Stand der Technik erheblich verbessern. Außerdem zeigen durchgeführte Experimente, dass die Methode wesentlich von der Verwendung von Langzeitinformationen profitiert. Die vorgestellte Methode führt auf den verwendeten Datensätzen zu einer nahezu optimalen Zuordnungsgenauigkeit bei gegebenen Detektionen. Zukünftige Arbeiten können sich daher auf andere Bereiche, wie der genaueren Extraktion von Detektionen konzentrieren. Insgesamt zeigt der zweite Teil dieser Arbeit, dass die verbesserte Ausnutzung der über die Zeit bereitgestellten Informationen zu einer erheblichen Verbesserung der Trajektorienergebnisse führt. Erstmalig wurde ein globales

Optimierungsverfahren zur Lösung von Datenassoziationsmodellen höherer Ordnung erfolgreich eingesetzt.

Schlagwörter: Verfolgung mehrerer Personen, Video, Zuordnungsmodelle höherer Ordnung, Sensorfusion, Binäres lineares Programm, Binäres quadratisches Programm, Global optimale Lösung