

VEHICLE TRACKING IN WIDE AREA MOTION IMAGERY : A FACILITY LOCATION MOTIVATED COMBINATORIAL APPROACH

Ahmed Elliethy, Gaurav Sharma

Dept. of Electrical and Computer Engineering, University of Rochester, Rochester, NY 14627
{ahmed.s.elliethy, gaurav.sharma}@rochester.edu

ABSTRACT

We propose a practical combinatorial approach for addressing the challenging task of vehicle tracking in Wide area motion imagery (WAMI) by leveraging a pixel accurate co-registered vector road-map. Specifically, guided by the co-registered road network, we obtain a sparse trellis graph linking each vehicle detection (VD) in a WAMI frame with a road reachable VD in the next frame, which then allows us to enumerate all possible hypotheses tracks. The globally optimal selection of tracks over a K frame window is then formulated as a minimum cost K capacitated facility location problem, where each hypothesis track is allocated VDs in each frame and assigned a cost that models desirable properties of vehicle track. Computation of optimized combinations of jointly feasible tracks that minimize the total cost for the tracks becomes feasible in our formulation for moderate values of K by utilizing available solvers for the facility location problem. The approach automatically selects the optimal number of tracks and provides flexibility in defining costs for tracks globally across the K frames. Vehicle tracking results obtained over test WAMI datasets indicate that our proposed method provides significant better performance than two other state of the art alternatives.

Index Terms— Wide area motion imagery, vehicle tracking, K capacitated facility location problem, vector road network

1. INTRODUCTION

New aerial imaging platforms offer motion imagery with rich spatio-temporal information that enables a host of new applications. We focus particularly on urban area wide area motion imagery (WAMI) that offers a high resolution picture sequences covering a “city-scale” area within each frame at temporal rates of 1-2 frames per second [1–3]. In this setting, we consider the problem of tracking the many vehicles present in the field of view. Specifically, given vehicle detections (VDs) in each frame, the goal is to associate detections that correspond to the same vehicle over the entire set of the WAMI frames into a vehicle track. Generally speaking, the vehicle tracking in WAMI is a challenging problem due to many factors such as the large number of the vehicles encountered that do not necessarily have strong discriminative features, the low temporal sampling rate, and frequent vehicles occlusions.

In this paper, we propose a novel combinatorial approach to solve the vehicle tracking in WAMI globally across K WAMI frames that are pixel accurately co-registered with a vector road network (RN). From the VDs detected in the entire set of the K WAMI frames, our approach starts with building a trellis graph (TG) [4] whose vertices are the VDs and the edges represent possible association from one VD to another. The pixel accurate co-registered vector (RN) allows us to sparsify the TG by limiting the association possibilities (number of edges) from each VD to only reach-

able VDs via the RN. Then, we parse the TG to generate all possible hypotheses tracks where each hypothesis track is assigned a cost that penalizes deviations from common behaviors of a vehicle track. Finally, we model the tracking problem as K capacitated facility location problem (KCFLP) [5], where track \equiv facility and VD \equiv client. Analogous to the KCFLP, our goal is to select a subset tracks from the hypotheses tracks such that the sum of the selected track costs is minimum where the selected subset tracks must satisfy two constraints: (a) each VD is assigned to only one track and (b) the selected tracks must have capacity K , i.e., an assigned VD in each frame. Our idea is presented on an illustrative example in Fig. 1. The proposed approach is scaled naturally to address large number of WAMI frames. Specifically, we partition these frames into multiple K frame temporal windows, solving the KCFLP for each window, and combining the estimated tracks within the temporal windows using the proposed approach. The TG is formed in this case by linking the estimated tracks within each K temporal window instead of linking VDs.

Prior work on vehicle tracking also exploits the context provided by a co-registered RN as well as other geographic information system (GIS) sources to enhance vehicle tracking performance [6, 7]. In [8], co-registered RN information is utilized to constrain the successive frame to frame VDs associations. Similarly, in [9], road orientation is estimated based on the movement of the current estimated vehicle tracks, which is used to constrain the assignment in future frames. However, these techniques are prone to ID switches especially when vehicles are moving in two-way road. Also wrong associations encountered across frame pairs are persistent, i.e., can not be corrected in future frames.

Another category of tracking techniques approaches the assignment problem globally over set of frames. Although the assignment problem is NP-hard when estimated for more than two frames, there are polynomial time complexity algorithms that have globally optimum solution in specific cases such as the network flow formulation [10–12]. One limitation of the network flow based techniques is that the cost of a hypothesis track must be decomposable as a product of pairwise cost terms and thus preclude the ability to incorporate higher order motion models (such as constant velocity/acceleration motion model) into the track cost. Other techniques approximately solve the assignment problem globally for more than two frames with utilization of higher order motion models [13, 14]. These techniques use a heuristic approach that iteratively solves the assignment problem between two frames while keeping the assignment between the other frames fixed, and therefore require good initial assignment solution to converge to good local optimum, i.e., the quality of the final assignment solution is sensitive to good initialization.

Compared with previous tracking techniques, our approach also utilizes a co-registered RN information but in a fundamentally different manner. Specifically, the co-registered RN not only sparsifies the

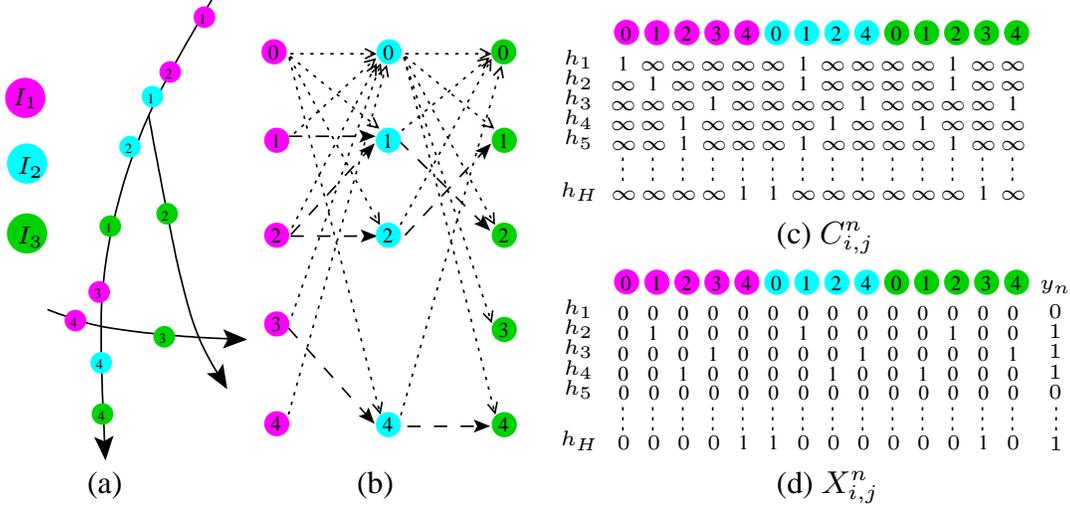


Fig. 1: Illustrative example shows the main steps of our proposed approach. In (a), we sketch a set of VDs contained in 3 frames (I_1 , I_2 , and I_3) represented in (magenta, cyan, and green) colors respectively, and overlaid with the co-registered RN. (b) shows the TG formed from the VDs shown in (a) augmented with a dummy node (0) in each frame to handle spurious and missed-detections. The RN help us to reduce the number of association possibilities and render the TG sparse. The assignment cost of the j^{th} VD within the i^{th} frame to a hypothesis track h_n generated from the TG is represented by $C_{i,j}^n$ and we show $C_{i,j}^n$ for all i, j, n in (c). In our example, h_2 is composed of z_1^1 , z_1^2 , and z_3^3 , and thus the assignment cost for these VDs is 1 and for the other VDs is ∞ . Solving the KCFLP yields hypotheses tracks h_2, h_3, h_4 , and h_H as an optimal solution. The solution comprises the hypothesis selection binary indicators y_n , and additionally, the VDs assignments to the selected hypotheses tracks represented by the assignment binary indicators $X_{i,j}^n$ for all i, j, n as shown in (d).

TG and consequently render our combinatorial approach tractable, but also plays an important role in assessing hypotheses tracks that are generated from TG as we discuss shortly. Moreover, our novel formulation of the tracking problem as KCFLP : (a) allows cost to be defined flexibly to assess the hypotheses tracks globally across the K frames, (b) estimates the optimal number of tracks (\equiv opened facilities) automatically, and (c) does not require any initialization since all hypotheses tracks are generated in advance.

This paper is organized as follows. Section 2 explains our formulation for vehicle tracking problem. Results and a comparison against alternative methods are presented in Section 3. We conclude the paper in Section 4.

2. OUR COMBINATORIAL APPROACH FOR VEHICLE TRACKING

Our goal in this paper is to estimate vehicle tracks in WAMI within a temporal window consisting of a set of K WAMI frames $\mathcal{I} = \{I_i\}_{i=1}^K$, where frame I_i contains N_i VDs and the WAMI frames in \mathcal{I} are co-registered with a vector road network (for example using [15, 16]). In the following we discuss the three main parts of our proposed algorithm. First, we describe how we form a trellis graph from the VDs. Then, we illustrate the hypotheses tracks generation from the trellis graph. Finally, we outline the optimal tracks selection formulation.

2.1. Trellis graph formation

Let $\mathcal{Z} = \{z_j^i | i \in \{1, \dots, K\}, j \in \{0, \dots, N_i\}\}$ be the set of all VDs in \mathcal{I} , where z_j^i is the location of the j^{th} VD in the i^{th} frame and z_0^i is a dummy VD that we add for frame I_i to account for spurious and missed-detections. Figure 1 (a) shows an example of a set of VDs contained in 3 frames and overlaid on the co-registered RN.

We construct the trellis graph $G = (\mathcal{Z}, \mathcal{E})$ whose nodes are the

set \mathcal{Z} and its edges $\mathcal{E} = \{(z_j^i, z_m^n) | z_m^n \in \mathcal{N}(z_j^i)\}$. The ordered pair (z_j^i, z_m^n) defines a directed connection link from z_j^i to z_m^n where $\mathcal{N}(z_j^i)$ is the neighbourhood of z_j^i that contains all VDs that can be associated with z_j^i . We state that $z_m^n \in \mathcal{N}(z_j^i)$ if two conditions are satisfied: (a) $n = i + 1$ and (b) $\vartheta(z_j^i, z_m^n) \leq \tau$, where $\vartheta(z_j^i, z_m^n)$ is the minimum distance of travel on the road network from z_j^i to z_m^n , and τ is a threshold that is determined based on the maximum distance a vehicle can travel between successive WAMI frames. The minimum distance of travel $\vartheta(a, b)$ is determined by the shortest route on the road network between the two locations a and b , which is estimated in our current implementation by Dijkstra's shortest path algorithm [17].

Figure 1 (b) shows the TG for the VDs shown in (a). The RN help us to reduce the number of association possibilities and render the TG sparse as shown in the figure. For example, z_4^2 (VD # 4 in I_2) can not be reached via the RN from z_4^1 (VD # 4 in I_1), i.e., $\vartheta(z_4^1, z_4^2) = \infty$ in this example, because there is no route between them, and therefore there is no link between them in the TG although they are spatially close.

2.2. Hypotheses tracks generation

From the constructed TG, we obtain the set of all possible hypotheses tracks $\mathcal{H} = \{h_1, \dots, h_H\}$, where H is the number of all possible hypotheses tracks generated from the TG. A hypothesis track h_n consists of a sequence of nodes that define a path on TG which starts with a node (VD) in I_1 and ends with a node in I_K , i.e., $h_n = (z_{n_1}^1, z_{n_2}^2, \dots, z_{n_K}^K)$.

Note that, multiple hypotheses tracks can claim the assignment of the same VD. In other words, the hypotheses tracks can conflict with each other, because in reality, a VD can be claimed only by one track. Our goal is to optimally select a non-conflicting subset from the hypotheses tracks such that no VD is assigned to more than one

selected track. Similarly, the goal of the KCFLP is to select a subset from a set of potential facilities such that each client is assigned to only one facility. Thus, we formulate our tracking problem as KCFLP and benefit from the efficient and effective algorithms proposed for the long studied KCFLP [5].

To formulate our tracking problem as KCFLP, we need to generate all hypotheses tracks (\equiv facilities) from the TG, defining the cost of assigning a VD to a hypothesis track (\equiv client to facility assignment cost), and defining a cost for selecting a hypothesis track (\equiv facility opening cost). We define the cost of assigning a VD to a hypothesis track as

$$C_{i,j}^n = \begin{cases} 1 & \text{if } z_j^i \in h_n, \\ \infty & \text{otherwise.} \end{cases} \quad (1)$$

Example of $C_{i,j}^n$ for all i, j, n shown in Fig. 1 (c). Since h_2 is composed of z_1^1, z_1^2 , and z_2^3 in this example, the assignment cost for these VDs is 1 and for the other VDs is ∞ .

Each hypothesis track h_n is assigned a selection cost f_n that penalizes deviations from common behavior of a vehicle track. Specifically, we define f_n to be proportional to (1) track motion smoothness, (2) directional chamfer distance [18] between track and the RN, (3) track length excluded dummy VDs. Formally,

$$f_n = p_m(h_n)p_r(h_n)p_l(h_n), \quad (2)$$

where

$$p_m(h_n) \propto \frac{1}{K-1} \sum_{i=2}^{K-1} \|z_{n_{i+1}}^{i+1} - 2z_{n_i}^i + z_{n_{i-1}}^{i-1}\| \quad (3)$$

is defined as in [13] and penalizes the motion irregularities¹ of vehicle track,

$$p_r(h_n) \propto \frac{1}{K} \sum_{i=1}^K \min_m d(z_{n_i}^i, \chi^m) + \lambda |\theta(z_{n_i}^i) - \theta(\chi^m)| \quad (4)$$

measures the directional chamfer distance between h_n and the RN, which penalizes deviations of vehicle track from the RN in both distance and orientation mismatches; where χ^m is the nearest point to $z_{n_i}^i$ in the RN, $\theta(z_{n_i}^i)$ is the orientation angle of h_n at $z_{n_i}^i$, $\theta(\chi^m)$ is the orientation angle of the road contains the point χ^m , and λ is a constant. Finally,

$$p_l(h_n) \propto \exp(-l(h_n)) \quad (5)$$

penalizes short vehicle tracks or tracks with many missed-detections, where $l(h_n) = |\{z_j^i | z_j^i \in h_n, j \neq 0\}|$ counts the number of VDs in h_n excluding dummy VDs.

2.3. Optimal tracks selection

Given : (a) the set of detections \mathcal{Z} , (b) the set of hypotheses tracks \mathcal{H} , (c) the cost $C_{i,j}^n$ of assigning z_j^i to h_n for all i, j, n , and (d) the cost f_n of the hypothesis track h_n for all n , our goal is to select a subset from the hypotheses tracks \mathcal{H} and the assignment of VDs to the selected subset of the hypotheses tracks that minimize the sum of the costs for the selected subset of the hypotheses tracks and for the VDs assignment to the selected subset of the hypotheses tracks. To accomplish this, we define a binary indicator variable y_n that takes the value 1 if h_n is selected and 0 otherwise. Similarly, we define a binary indicator variable $X_{i,j}^n$ that takes the value 1 if z_j^i is assigned

to h_n and 0 otherwise. Our tracking problem goal now is to find y_n and $X_{i,j}^n$ for all i, j, n , that minimize

$$\sum_{n=1}^H \sum_{i=1}^K \sum_{j=0}^{N_i} C_{i,j}^n X_{i,j}^n + \sum_{n=1}^H f_n y_n, \quad (6)$$

subject to the constraints

$$\sum_{j=0}^{N_i} X_{i,j}^n = y_n, \quad \forall n, \forall i, \quad (7)$$

$$\sum_{n=1}^H X_{i,j}^n = 1, \quad \forall i, \forall j \neq 0, \quad (8)$$

$$X_{i,j}^n \in \{0, 1\}, \quad \forall n, \forall i, \forall j, \quad (9)$$

$$y_n \in \{0, 1\}, \quad \forall n. \quad (10)$$

The constraint (7) ensures that if h_n is selected (i.e., $y_n = 1$), it must have a single assigned VD in each frame. Otherwise, (if not selected, i.e., $y_n = 0$), there is no any assigned VD to it in any frame. In other words, VDs are assigned only to a selected h_n . The constraint (8) ensures that a VD assigned only to one selected track. The constraints (9) and (10) force $X_{i,j}^n$ and y_n to be binary, respectively. Note that the constraint (7) is equivalent to

$$\sum_{i=1}^K \sum_{j=0}^{N_i} X_{i,j}^n = K y_n, \quad \forall n,$$

which, in the KCFLP terminology, states that each facility must have exactly capacity K , i.e., exactly K clients must be assigned to facility. This constraint establishes the mapping between our tracking problem and the KCFLP. Figure 1 (d) shows the estimated y_n and $X_{i,j}^n$ for all i, j, n in the illustrative example, where, the hypotheses tracks h_2, h_3, h_4 , and h_H are being selected.

3. EXPERIMENTAL RESULTS

We evaluated our approach on two WAMI data sets: (1) CorvusEye dataset that is recorded using the CorvusEye 1500 Wide-Area Airborne System [19] for the Rochester, NY region, and (2) Wright-Patterson Air Force Base (WPAFB) 2009 dataset [20], which was recorded over the WPAFB, OH region. For the vector road map, we use OpenStreetMap (OSM) [21], which provides each road in the road network in a vector format along with properties of each road such as its type (highway, residential, etc), one or two-way road, number of lanes, etc. We use the method in [16] for co-registering the WAMI frames to the vector road map. We obtain the VDs by background subtraction, where the background is estimated using median filter as in [9]. We implement our approach using C++ and adopt the solver in [22] for solving (6). Our unoptimized tracker implementation runs at 0.16 frame/second for² $K = 5$.

In order to evaluate our tracking methodology in different scenarios, we carefully create three test sequences such that each test sequence has its own characteristics. The test sequence ‘‘Seq1’’ is formed by cropping a region from the CorvusEye dataset that contains a forked one-way roads with different directions and also a lot of occluders (bridges, trees, etc.). The sequences ‘‘Seq2’’

¹The proposed approach is also applicable with more sophisticated motion models than the simple constant velocity model we use in this paper.

²The value of K poses a trade-off between the quality of the estimated tracks and the computational cost. Experimentally, we found $K = 5$ provides estimated tracks with good quality at reasonable computational cost.



Fig. 2: Test sequences used in our evaluation. Each test sequence contains 30 frames and we show the first frame from “Seq1”, “Seq2”, and “Seq3”, in (a), (b), and (c) respectively. All roads in “Seq1” and “Seq3” are one-way roads, while roads in “Seq2” are two-way roads.

and “Seq3” are formed by cropping two different regions from the WPAFB dataset. Both “Seq2” and “Seq3” cover regions that are clearly observed (no occluders) but all roads in “Seq2” are two-way roads, while roads in “Seq3” are one-way roads. All test sequences are ground truth labeled and composed of 30 frames. We show the first frame from each sequence in Fig. 2.

We compare our method with the methods in [8] and [9] that also exploit the RN information for estimating vehicle tracks in WAMI. We quantify the performance of the tracking methods by the measures defined in [23] which are (1) the total number of ID switches for the tracked vehicles compared to their ground truth label ($IDS\downarrow$), (2) the number of mostly tracked vehicles ($MT\uparrow$), and (3) the number of mostly lost vehicles ($ML\downarrow$). In addition to these measures, we report the Multiple Object Tracking Accuracy ($MOTA\uparrow$) defined in [24] (with $c_s = 1$).

Seq	GT	Method	$MT\uparrow$	$ML\downarrow$	PT	$IDS\downarrow$	$MOTA\uparrow$
1	77	[8]	38	4	35	85	0.963
		[9]	32	4	41	105	0.954
		Prop.	40	6	31	62	0.973
2	54	[8]	43	11	2	91	0.887
		[9]	44	6	4	95	0.882
		Prop.	49	4	1	4	0.995
3	42	[8]	38	0	4	9	0.985
		[9]	38	0	4	16	0.974
		Prop.	39	0	3	3	0.994

Table 1: Tracking performance comparison between our proposed approach (prop.) and the methods in [8] and [9] evaluated on our three test sequences. The tracking performance measures are the total number of ID switches ($IDS\downarrow$), the number of mostly tracked vehicles ($MT\uparrow$), the number of mostly lost vehicles ($ML\downarrow$), the number of partially tracked vehicles (PT), and Multiple Object Tracking Accuracy ($MOTA\uparrow$)

Table 1 shows for each test sequence the number of ground truth tracks (GT) contained in this sequence, in addition to the tracking performance measures mentioned above for each tracking method. From the table, we can draw three important conclusions. First, the performance of our proposed approach is significantly better than the methods in [8] and [9] for all test sequences. This is because our

approach generates all hypotheses tracks first, and then optimally selects the best subset from them, disregarding low quality tracks that have high selection cost f_n . Second, in the case of two-way roads as in “Seq2”, the methods in [8] and [9] perform poorly. Because the road direction is non-informative in this case, the methods in [8] and [9] face a challenging situation when an ambiguous VD can be associated with one of two conflicting tracks that are corresponding to vehicles approaching each other from the opposite directions of the road. A wrong association in this delicate scenario may result in either early termination of the track as it will not find good association in the next frame (i.e., lower MT) or propagation of wrong associations in future frames (i.e., higher IDS). On the other hand, our proposed approach assigns the ambiguous VD to both conflicting tracks that correspond to the opposite directions approaching vehicles in this scenario, and optimally selects the best among them considering a global cost defined in (2). This optimal selection strategy of our approach is reflected in the results reported for “Seq2” in Table 1, which show significant improvement of the tracking performance for our proposed approach compared with [8] and [9]. Finally, in relatively easy sequences such as “Seq3” which contains only one-way roads without occluders, all methods behave reasonably well.

4. CONCLUSION

In this paper, we propose a novel combinatorial approach to solve the vehicle tracking in WAMI globally across K frames. Our approach is motivated by pixel accurate co-registered vector road network with the WAMI frames that allows us to limit the association possibilities for each VD and consequently limits the number of considered hypotheses tracks and renders our approach tractable. Our approach formulates the tracking problem as a K capacitated facility location problem that provides an elegant way to optimally select subset tracks from the hypotheses tracks that minimize the sum of costs for the selected tracks. Results obtained over three test sequences that represent different tracking scenarios show a significant performance improvement for proposed approach compared with two state of the art alternative methods.

5. ACKNOWLEDGMENT

We thank Bernard Brower of Harris Corporation for making available the CorvusEye [19] WAMI datasets used in this research.

6. REFERENCES

- [1] K. Palaniappan, R. M. Rao, and G. Seetharaman, "Wide-area persistent airborne video: Architecture and challenges," in *Distributed Video Sensor Networks*. Springer, 2011, pp. 349–371.
- [2] E. Blasch, G. Seetharaman, S. Suddarth, K. Palaniappan, G. Chen, H. Ling, and A. Basharat, "Summary of methods in wide-area motion imagery (WAMI)," in *Proc. SPIE*, vol. 9089, 2014, pp. 90 890C–90 890C–10.
- [3] R. Porter, A. Fraser, and D. Hush, "Wide-area motion imagery," *IEEE Sig. Proc. Mag.*, vol. 27, no. 5, pp. 56–65, Sept 2010.
- [4] C. Schlegel and L. Perez, *Trellis and Turbo Coding: Iterative and Graph-Based Error Control Coding*, ser. IEEE Series on Digital & Mobile Communication. Wiley, 2015.
- [5] B. Korte and J. Vygen, *Combinatorial Optimization: Theory and Algorithms*, 5th ed. Springer Publishing Company, Incorporated, 2012.
- [6] B. Bradford, E. M. Dixon, J. Sisskind, and W. D. Reynolds, Jr., "Target tracking with GIS data using a fusion-based approach," in *Proc. SPIE*, vol. 8053, 2011, pp. 80 530F–80 530F–11.
- [7] J. Prokaj and G. Medioni, "Using 3D scene structure to improve tracking," in *IEEE Intl. Conf. Comp. Vision, and Pattern Recog.*, June 2011, pp. 1337–1344.
- [8] J. Xiao, H. Cheng, H. Sawhney, and F. Han, "Vehicle detection and tracking in wide field-of-view aerial video," in *IEEE Intl. Conf. Comp. Vision, and Pattern Recog.*, June 2010, pp. 679–684.
- [9] V. Reilly, H. Idrees, and M. Shah, "Detection and tracking of large number of targets in wide area surveillance," in *Proc. European Conf. Computer Vision*, ser. Lecture Notes in Computer Science, 2010, vol. 6313, pp. 186–199.
- [10] L. Zhang, Y. Li, and R. Nevatia, "Global data association for multi-object tracking using network flows," in *IEEE Intl. Conf. Comp. Vision, and Pattern Recog.*, June 2008, pp. 1–8.
- [11] J. Berclaz, F. Fleuret, E. Turetken, and P. Fua, "Multiple object tracking using k-shortest paths optimization," *IEEE Trans. Pattern Anal. Mach. Intel.*, vol. 33, no. 9, pp. 1806–1819, 2011.
- [12] H. Pirsiavash, D. Ramanan, and C. C. Fowlkes, "Globally-optimal greedy algorithms for tracking a variable number of objects," in *IEEE Intl. Conf. Comp. Vision, and Pattern Recog.*, 2011, pp. 1201–1208.
- [13] R. T. Collins, "Multitarget data association with higher-order motion models," in *IEEE Intl. Conf. Comp. Vision, and Pattern Recog.*, 2012, pp. 1744–1751.
- [14] X. Shi, H. Ling, J. Xing, and W. Hu, "Multi-target tracking by rank-1 tensor approximation," in *IEEE Intl. Conf. Comp. Vision, and Pattern Recog.*, June 2013, pp. 2387–2394.
- [15] A. Elliethy and G. Sharma, "Vector road map registration to oblique wide area motion imagery by exploiting vehicles movements," in *IS&T Electronic Imaging: Video Surveillance and Transportation Imaging Applications*, San Francisco, California, 2016, pp. VSTIA–520.1–8. [Online]. Available: <http://ist.publisher.ingentaconnect.com/contentone/ist/ei/2016/00002016/00000003/art00008>
- [16] ———, "Automatic registration of vector road maps with wide area motion imagery by exploiting vehicle detections," *IEEE Trans. Image Proc.*, vol. 25, no. 11, pp. 5304 – 5315, Nov. 2016.
- [17] D. Medhi and K. Ramasamy, *Network Routing: Algorithms, Protocols, and Architectures*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2007.
- [18] M.-Y. Liu, O. Tuzel, A. Veeraraghavan, and R. Chellappa, "Fast directional chamfer matching," in *IEEE Intl. Conf. Comp. Vision, and Pattern Recog.*, June 2010, pp. 1696–1703.
- [19] "CorvusEye™ 1500 Data Sheet," http://www.exelisinc.com/solutions/corvuseye1500/Documents/CorvusEye500_DataSheet_AUG14.pdf.
- [20] "AFRL WPAFB 2009 data set," <https://www.sdms.afrl.af.mil>.
- [21] "OpenStreetMap," <http://www.openstreetmap.org>, accessed Mar. 2016.
- [22] "Gurobi optimizer reference manual," 2015. [Online]. Available: <http://www.gurobi.com>
- [23] B. Wu and R. Nevatia, "Tracking of multiple, partially occluded humans based on static body part detection," in *IEEE Intl. Conf. Comp. Vision, and Pattern Recog.*, vol. 1, 2006, pp. 951–958.
- [24] R. Kasturi, D. Goldgof, P. Soundararajan, V. Manohar, J. Garofolo, R. Bowers, M. Boonstra, V. Korzhova, and J. Zhang, "Framework for performance evaluation of face, text, and vehicle detection and tracking in video: Data, metrics, and protocol," *IEEE Trans. Pattern Anal. Mach. Intel.*, vol. 31, no. 2, pp. 319–336, Feb 2009.