Joint Tracking and Ground Plane Estimation

Junseok Kwon, Ralf Dragon, and Luc Van Gool

***Abstract*—We propose a novel framework that jointly estimates the ground plane and a target's motion trajectory. This results in improvements for both. Estimating their joint posterior is based on Particle Markov Chain Monte Carlo (Particle MCMC). In Particle MCMC, the best target state is inferred by a particle filter and the best ground plane is obtained by MCMC. Compared with conventional sampling methods that iteratively infer the best target states and ground plane parameters, our method infers them jointly. This reduces sampling errors drastically. Experimental results demonstrate that our method outperforms several state-of-the-art tracking methods, while the ground plane accuracy is also improved.**

***Index Terms*—Ground plane estimation, object tracking, particle Markov chain Monte Carlo (Particle MCMC).**

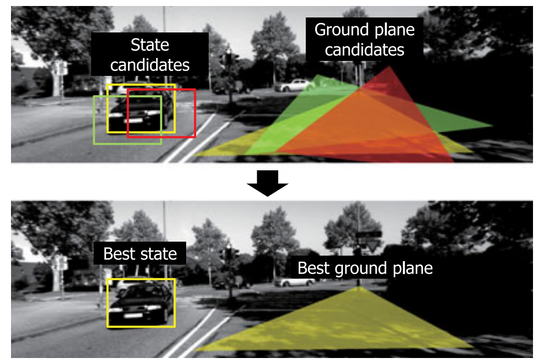

Fig. 1. Basic idea of our method. Our method searches for the best target state and ground plane at the same time, whereas conventional tracking methods find them either separately or iteratively.

I. Introduction

In order to track a target accurately, especially under severe occlusions, several methods [1]–[3] use the ground plane. When considered through orthogonal projection onto that plane, most targets no longer suffer occlusion. This simplifies the tracking. This paper also proposes the ground plane, but adds robustness via a *Feedback Process* and a process of *Joint Inference*. Fig. 1 illustrates this joint optimization.

Feedback Process: The tracking methods in the aforementioned papers [1]–[3] focus on improving the visual tracking with the help of an estimated ground plane. Yet, they do not include a feedback mechanism to enhance the ground plane estimation based on the tracking results. If the ground plane is estimated wrongly, its use may even impair the visual-tracking performance. Instead, our method searches for a better estimate of the ground plane during the tracking process. It assesses the goodness of the ground plane by checking the smoothness of the tracking results when projected onto it.

For their street scene analysis, Ess *et al.* [4] and later Wojek *et al.* [5] propose frameworks where tracking and ground plane estimation are of mutual benefit. Whereas the former work uses depth information (stereo), the latter is purely monocular, as in our case. Both approaches heavily depend on object detection, also for the ground plane estimation (e.g., needing rather precisely estimated feet positions as points on the ground plane). The approach we present here is more versatile and generic.

Manuscript received June 22, 2016; revised August 11, 2016; accepted August 12, 2016. Date of publication August 12, 2016; date of current version September 13, 2016. This work was supported by the European Research Council (ERC) under the project VarCity (#273940) and was supported in part by the Chung-Ang University Research Scholarship Grants in 2016. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Simon Maskell.

J. Kwon is with the School of Computer Science and Engineering, Chung-Ang University, Seoul 156–756, Korea (e-mail: jskwon@cau.ac.kr).

R. Dragon and L. V. Gool are with the Computer Vision Lab, Department of Information Technology and Electrical Engineering, ETH Zurich, Zurich 8092, Switzerland (e-mail: dragon@vision.ee.ethz.ch; vangool@vision.ee.ethz.ch).

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2016.2601085

Joint Inference: For the joint estimation of the target state and the ground plane position, expectation-maximization (EM) [6] and Gibbs sampling [7] come to mind. In such estimation procedures, the target state is considered to be represented by latent variables $\mathbf{X}_t$ and the ground plane position by unknown parameters θ_t, to be determined from the observations. For example, the EM algorithm searches for the best value of the latent variables given the parameters, whereupon it uses the latent variables again to compute a better estimate for the parameters, etc. The Gibbs method samples the latent variables from their distribution conditioned on the parameters. Similarly, it samples the parameters from their distribution conditioned on the latent variables. Both approaches however, have an error propagation problem: if either the latent variables or the unknown parameters are estimated inaccurately, these errors contaminate the other unknowns. This error propagation is due to the iterative estimation of the latent variables and the unknown parameters. To solve this error propagation problem, our method jointly estimates the latent variables and the unknown parameters. To that end, it uses particle MCMC, proposed in [8] and recently used for visual tracking in [9]. Recently, a robust Bayesian algorithm [10] has been proposed to estimate the target state and the ground plane position. However, this algorithm is also an iterative method and thus has aforementioned drawbacks.

The design of the joint posterior that allows to jointly solve for the target state and the ground plane is discussed in Section II. The estimation of the MAP values through Particle MCMC is explained in Section III. Section IV shows experimental results and Section V conclude the paper.

II. Joint Posterior

Our goal of visual tracking is to find the state $\hat{\mathbf{X}}_t$ and the ground plane $\hat{\theta}_t$ given the image observations up to time t, $\mathbf{Y}_{1:t}$, which maximize the joint posterior $p(\mathbf{X}_t, \theta_t|\mathbf{Y}_{1:t})$. We can achieve this goal with the maximum joint a posteriori estimation:

$$\{\hat{\mathbf{X}}_t, \hat{\theta}_t\} = \underset{\{\mathbf{X}_t, \theta_t\}}{\arg} \max p(\mathbf{X}_t, \theta_t|\mathbf{Y}_{1:t}) \qquad (1)$$

where $\mathbf{X}_t$ denotes a target configuration on the image at time t. In (1), θ indicates a parameter that makes a ground plane. The parameter has dynamics and is hence denoted as θ_t at time t. $\mathbf{X}_t = \{X_t^x, X_t^y, X_t^s\}$ includes the x, y center positions and the scale s of the bounding box.

In order to measure smoothness of the target trajectory in the ground plane, we project the target state from the image coordinate to the ground plane using the parameters θ_t. θ_t is estimated from a single camera on the basis of [11]. The method is robust where it only makes the mild assumption that the camera's ego motion $\vec{T}$ and the ground plane normal $\vec{N}$ are orthogonal over longer time spans.

If the conventional MCMC is used for obtaining samples from $p(\mathbf{X}_t, \theta_t|\mathbf{Y}_{1:t})$, then the acceptance ratio is designed in the following fashion:

$$a = \min\left[1, \frac{p(\mathbf{X}_t^*, \theta_t^*|\mathbf{Y}_{1:t})q\{(\mathbf{X}_t, \theta_t)|(\mathbf{X}_t^*, \theta_t^*)\}}{p(\mathbf{X}_t, \theta_t|\mathbf{Y}_{1:t})q\{(\mathbf{X}_t^*, \theta_t^*)|(\mathbf{X}_t, \theta_t)\}}\right] \quad (2)$$

where $q\{(\mathbf{X}_t^*, \theta_t^*)|(\mathbf{X}_t, \theta_t)\}$ is the proposal density that proposes a new state $\mathbf{X}_t^*$ and a new ground plane θ_t^*, which is defined in Section III. However, this acceptance ratio causes two challenges.

1) Because the posterior $p(\mathbf{X}_t, \theta_t|\mathbf{Y}_{1:t})$ is high dimensional, the landscape of the posterior is very rough. In this situation, the acceptance rate obtained by (2) tends to be very low and the sampling method easily gets trapped in local minima.
2) The proposal density $q\{(\mathbf{X}_t^*, \theta_t^*)|(\mathbf{X}_t, \theta_t)\}$ in (2) is hard to design because two variables, $\mathbf{X}_t$ and θ_t, should be proposed simultaneously.

These challenges are efficiently solved by Particle MCMC, as explained in Section III.

III. Particle MCMC

To find the best state and ground plane, our method gets several samples from $p(\mathbf{X}_t, \theta_t|\mathbf{Y}_{1:t})$ using Particle MCMC and chooses the best one. Particle MCMC is a nontrivial combination of MCMC [12] and particle filter [13] methods which takes advantage of the strength of its two components.

A. MCMC

With MCMC, our method generates candidates for the ground plane. MCMC consists of the proposal and the acceptance steps as follows.

Proposal Step: In the proposal step, a new state $\mathbf{X}_t^*$ and a new ground plane θ_t^* are proposed by the proposal density $q\{(\mathbf{X}_t^*, \theta_t^*)|(\mathbf{X}_t, \theta_t)\}$. However, it is unfeasible to directly propose $\mathbf{X}_t^*$ and θ_t^* by using $q\{(\mathbf{X}_t^*, \theta_t^*)|(\mathbf{X}_t, \theta_t)\}$ because of the high dimensionality of $(\mathbf{X}_t, \theta_t)$. Hence, we decompose the original proposal density into two low-dimensional ones:

$$q\{(\mathbf{X}_t^*, \theta_t^*)|(\mathbf{X}_t, \theta_t)\} = q(\theta_t^*|\theta_t) \cdot p(\mathbf{X}_t^*|\theta_t^*, \mathbf{Y}_{1:t}) \quad (3)$$

where $q(\theta_t^*|\theta_t)$ is the proposal density for the ground plane and $p(\mathbf{X}_t^*|\theta_t^*, \mathbf{Y}_{1:t})$ is the conditional density for the target state. Because $\mathbf{X}_t^*$ in $p(\mathbf{X}_t^*|\theta_t^*, \mathbf{Y}_{1:t})$ depends on θ_t^*, our method can only propose a new ground plane θ_t^* in this stage with $q(\theta_t^*|\theta_t)$. We design $q(\theta_t^*|\theta_t)$ as follows:

$$\theta_t^* \text{ with probability } q(\theta_t^*|\theta_t) = e^{-|C(\theta_t^*) - C(\theta_t)|} \quad (4)$$

where the function $C(\theta_t)$ returns the confidence of the ground plane θ_t. To propose a ground plane, our method makes several ground plane candidates by using [11] in advance and probabilistically chooses a ground plane among them based on (4). We further model $p(\mathbf{X}_t|\theta_t, \mathbf{Y}_{1:t})$[1] as follows:

$$p(\mathbf{X}_t|\theta_t, \mathbf{Y}_{1:t}) = e^{-\lambda_1 S(P_{\theta_t}(\mathbf{X}_t))} \quad (5)$$

where λ_1 is a weighting parameter and the function $S(\cdot)$ measures the trajectory smoothness. To measure the trajectory smoothness on the ground plane, our method first projects the states in the trajectory into the ground plane θ_t using the function $P_{\theta_t}(\cdot)$ in [11] and calculates the standard deviation of them. Although our method uses specific functions presented in [11] for $C(\cdot)$ in (4) and $P_{\theta_t}(\cdot)$ in (5), we can use any functions introduced in conventional ground plane estimation methods.

Acceptance Step: In the acceptance step, we determine if the proposed $\mathbf{X}_t^*$ and θ_t^* are accepted or not by the acceptance ratio in (2). With $p(\mathbf{X}_t, \theta_t|\mathbf{Y}_{1:t}) = p(\theta_t|\mathbf{Y}_{1:t})p(\mathbf{X}_t|\theta_t, \mathbf{Y}_{1:t})$ and (3), the acceptance ratio in (2) is changed to

$$\begin{aligned} &\frac{p(\mathbf{X}_t^*, \theta_t^*|\mathbf{Y}_{1:t})q\{(\mathbf{X}_t, \theta_t)|(\mathbf{X}_t^*, \theta_t^*)\}}{p(\mathbf{X}_t, \theta_t|\mathbf{Y}_{1:t})q\{(\mathbf{X}_t^*, \theta_t^*)|(\mathbf{X}_t, \theta_t)\}} \\ &= \frac{p(\mathbf{X}_t^*|\theta_t^*, \mathbf{Y}_{1:t})p(\theta_t^*|\mathbf{Y}_{1:t})q(\theta_t|\theta_t^*)p(\mathbf{X}_t|\theta_t, \mathbf{Y}_{1:t})}{p(\mathbf{X}_t|\theta_t, \mathbf{Y}_{1:t})p(\theta_t|\mathbf{Y}_{1:t})q(\theta_t^*|\theta_t)p(\mathbf{X}_t^*|\theta_t^*, \mathbf{Y}_{1:t})} \\ &= \frac{p(\theta_t^*|\mathbf{Y}_{1:t})q(\theta_t|\theta_t^*)}{p(\theta_t|\mathbf{Y}_{1:t})q(\theta_t^*|\theta_t)} = \frac{p(\mathbf{Y}_{1:t}|\theta_t^*)q(\theta_t|\theta_t^*)}{p(\mathbf{Y}_{1:t}|\theta_t)q(\theta_t^*|\theta_t)} \end{aligned} \quad (6)$$

where $\mathbf{X}_t$ cancels out. In (6), the difficult problem of sampling from the joint posterior $p(\mathbf{X}_t, \theta_t|\mathbf{Y}_{1:t})$ is reduced to that of sampling from the simple posterior $p(\theta_t|\mathbf{Y}_{1:t})$, which is typically defined on a much smaller space. Nevertheless, our samples exactly follow the distribution for the original posterior $p(\mathbf{X}_t, \theta_t|\mathbf{Y}_{1:t})$, when the number of particles tends to infinite, as proven in [8]. On the other hand, conventional sampling methods like Gibbs sampling in [7] transform the original posterior into the simple one. Hence, the samples do not follow the original posterior but represent the approximated posterior $p(\mathbf{X}_t|\theta_t, \mathbf{Y}_{1:t})$ and $p(\theta_t|\mathbf{X}_t, \mathbf{Y}_{1:t})$. Now the remaining task is to design the marginal likelihood $p(\mathbf{Y}_{1:t}|\theta_t)$ in (6).

B. Particle Filter

With particle filter, our method obtains several particles for the target state and calculates $p(\mathbf{Y}_{1:t}|\theta_t)$.

Proposal and Weighting of Particles: Our method obtains the nth sampled particle $\mathbf{X}_t^{(n)}$ for the target state like

$$\mathbf{X}_t^{(n)} \text{ with probability } p(\mathbf{X}_t|\theta_t, \mathbf{Y}_{1:t}) \quad (7)$$

where $p(\mathbf{X}_t|\theta_t, \mathbf{Y}_{1:t})$ is defined in (5). Then, the weight of the particle $w(\mathbf{X}_t^{(n)})$ is calculated by

$$w(\mathbf{X}_t^{(n)}) \equiv p(\mathbf{Y}_t|\mathbf{X}_t^{(n)}) = e^{-\lambda_2 \mathrm{Dist}(\mathbf{Y}_t(\mathbf{X}_t^{(n)}), \mathrm{TM}_t)} \quad (8)$$

where λ_2 is a weighting parameter and $\mathbf{Y}_t(\mathbf{X}_t^{(n)})$ indicates the observation $\mathbf{Y}_t$ inside of the bounding box described by $\mathbf{X}_t^{(n)}$. In (8), the function Dist returns the similarity between $\mathbf{Y}_t(\mathbf{X}_t^{(n)})$ and the reference target model TM_t. The target model TM_t is obtained by the ASLA method in [14].

[1]Only X_t^x and X_t^s are dependent of θ_t in this case.

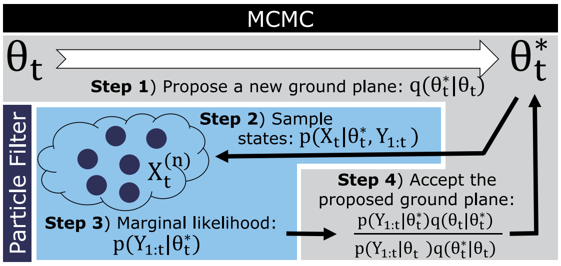

Fig. 2. Whole procedure of our method. Our Particle MCMC method combines the MCMC process (described in Section III-A) with the Particle Filter process (described in Section III-B).

Algorithm 1: Joint Inference of Target State and Ground Plane.

Input: $\hat{\theta}_{t-1}$
Output: $\hat{\mathbf{X}}_t$ and $\hat{\theta}_t$

Black: MCMC, Blue: Particle Filter
for $m = 1$ to M **do**
  Step 1: Propose a new ground plane θ_t^* by (4).
  Step 2: Run the Particle Filter algorithm of which target posterior is $p(\mathbf{X}_t|\theta_t, \mathbf{Y}_{1:t})$.
  for $n = 1$ to N **do**
    Sample the n^{th} particle $\mathbf{X}_t^{(n)}$ by using (7).
    Compute the weight of the sampled particle, $w(\mathbf{X}_t^{(n)})$, by using (8).
  end for
  Step 3: Calculate the marginal likelihood $p(\mathbf{Y}_{1:t}|\theta_t)$ by using N states and (9).
  Step 4: Accept the proposed ground plane θ_t^* with probability a computed by (6).
end for
Step 5: Find the best state $\hat{\mathbf{X}}_t$ among MN states and the best ground plane $\hat{\theta}_t$ among M ground planes using (1).

Design of Marginal Likelihood $p(\mathbf{Y}_{1:t}|\theta_t)$: $p(\mathbf{Y}_{1:t}|\theta_t)$ is decomposed by the chain rule like

$$p(\mathbf{Y}_{1:t}|\theta_t) = p(\mathbf{Y}_1|\theta_t)\prod_{i=2}^{t} p(\mathbf{Y}_i|\mathbf{Y}_{1:i-1}, \theta_t). \tag{9}$$

$p(\mathbf{Y}_i|\mathbf{Y}_{1:i-1}, \theta_t)$ in (9) is then calculated by the average of the weights of all particles:

$$p(\mathbf{Y}_i|\mathbf{Y}_{1:i-1}, \theta_t) = \frac{1}{N}\sum_{n=1}^{N} w(\mathbf{X}_t^{(n)}) \tag{10}$$

where the equality is proven in [8]. In (10), N denotes the total number of particles and $w(\mathbf{X}_t^{(n)})$ is defined in (8). Fig. 2 and Algorithm 1 describe the whole procedure of our method.

IV. Experiment

The proposed method was compared with 10 state-of-the-art tracking methods, IVT [15], SemiT [16], MIL [17], STR'11 [18], ASLA [14], CSK [19], KCF [20], STR'15 [21], ALIEN [22], and TCS [10], which are chosen based on [23]. For IVT, ASLA, and our method, the same number of samples, 500 (M = 10 and N = 50 for our method), was used to track the target. Our method always uses the same parameters throughout all experiments [$\lambda_1 = 0.05$ in (5) $\lambda_2 = 0.05$ in (8)].

A. Performance of Particle MCMC

In Section I, we argued that the Gibbs sampling approach causes the error propagation problem: A noisy ground plane produces an inaccurate target state and the inaccurate target state again makes the ground plane more erroneous. To show the disadvantages of Gibbs sampling and demonstrate the efficiency of our Particle MCMC, we conducted the following experiment. We made a tracker that uses Gibbs sampling to search the best state and ground plane iteratively and compared the tracker with our Particle MCMC based tracker that searches the best state and ground plane jointly.

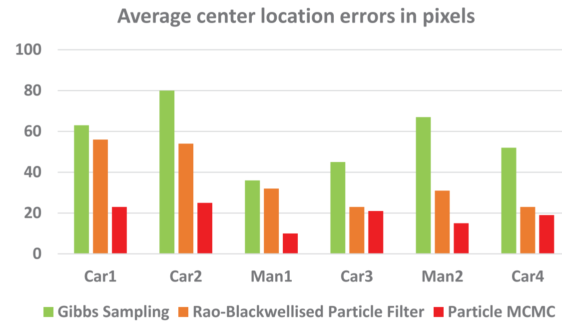

Fig. 3. Tracking performance according to the samplers.

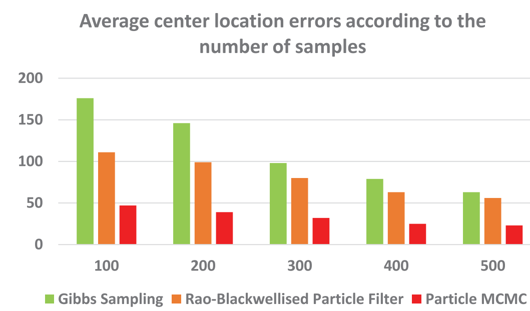

Fig. 4. Tracking performance according to the number of samples.

Fig. 3 shows that our Particle MCMC-based tracker produces more accurate tracking results than the Gibbs sampling-based tracker and the Rao–Blackwellised particle filter [24] based tracker in all sequences. Particle MCMC gave almost 70% improvement of tracking accuracy than the Gibbs sampling-based tracker. Gibbs sampling obtained inaccurate target states due to error propagation. In terms of sampling, the Gibbs sampling approach searches suboptimal states because the approach does not sample from the original posterior but samples from the approximated posterior. Notably, it transforms the original joint posterior into two conditional posteriors and samples from two conditional posteriors iteratively for easy sampling. Instead, our Particle MCMC is a perfect sampling method as number of samples goes infinite, which means samples obtained by Particle MCMC exactly follow the original posterior distribution. As demonstrated in Fig. 3, Particle MCMC also outperforms Rao–Blackwellised particle filter.

A sampler suffers from the high dimensionality of a sample space. Then, the sampler needs more samples to get good tracking accuracy. As shown in Fig. 4, the tracking accuracy significantly decreases as the number of samples decreases in the case of either Gibbs sampling or Rao–Blackwellised particle filter. They suffer from the high dimensionality of a sample space when samples are not sufficient. On the other hand, our tracker based on Particle MCMC produces accurate tracking results even when only 100 samples are used.

Fig. 5 demonstrates that our Particle MCMC produces more accurate ground planes than Gibbs sampling in most sequences. In Gibbs sampling, a good ground plane can be chosen based on the smoothness of an inaccurate trajectory. On the other hand,

Fig. 5. Improvement of ground plane estimation performance. Particle MCMC improved qualitative accuracy of ground plane estimation, where red and green planes denote ground planes obtained by Particle MCMC and Gibbs sampling, respectively.

TABLE I
QUANTITATIVE COMPARISON

	IVT	SemiT	MIL	STR'11	STR'15	ASLA	CSK	KCF	ALIEN	TCS	Ours
Car1	149(3)	120(9)	89(5)	48(44)	29(44)	57(32)	67(44)	64(**55**)	45(40)	**4**(**84**)	**23**(45)
Car2	81(12)	28(36)	56(45)	62(47)	32(55)	56(**60**)	51(28)	39(50)	41(50)	31(55)	**25**(**60**)
Man1	58(4)	26(3)	37(14)	57(18)	32(20)	**10**(**78**)	25(45)	17(47)	15(70)	**6**(**87**)	**10**(77)
Car3	151(7)	**19**(41)	52(42)	35(33)	23(70)	28(**81**)	71(32)	52(45)	55(40)	**17**(**81**)	21(**84**)
Man2	52(1)	18(21)	24(25)	18(25)	15(60)	16(68)	55(50)	48(56)	35(60)	**10**(**84**)	**15**(**74**)
Car4	108(10)	66(6)	100(8)	130(8)	67(19)	71(46)	99(25)	87(33)	50(41)	**20**(**55**)	**19**(**51**)

The numbers indicate average center location errors in pixels. The numbers in () indicate the amount of successfully tracked frames (score >0.5), where the score is defined by the overlap ratio between the predicted bounding box B_p and the ground truth bounding box B_{gt}: $\frac{area(B_p \bigcap B_{gt})}{area(B_p \bigcup B_{gt})}$. Blue is the best result. Red is the second best.

our method chose a good ground plane based on the accuracy of the trajectory, which is optimal in terms of sampling.

B. Comparison With Other Trackers

Tables I demonstrates that our method is competitive with TCS and outperforms CSK, KCF, STR'15, and ALIEN in terms of center location error and success rate. These sequences are very challenging because they contain occlusions, scale changes, and complex camera motions. Our method overcomes these challenges with the ground plane assumption that a target should be on the ground plane. Although our method used additional ground plane information for visual tracking and other trackers did not, this is a fair comparison because our method obtained the ground plane without additional side information from camera calibration. Please note that, compared with TCS, our method has an additional ability of estimating accurate ground planes.

V. CONCLUSION AND DISCUSSION

We propose a novel framework to perform visual tracking and ground plane estimation simultaneously. In the framework, tracking results and ground plane estimates help each other to improve both performances. With Particle MCMC, our method infers the best target state and ground plane jointly and efficiently. In the experiment, our method shows the best tracking performance.

Recently more advanced particle MCMC methods [25], [26] have been proposed. Our future work is to use these methods to improve visual tracking and ground plane estimation performance.

REFERENCES

[1] J. Berclaz, F. Fleuret, and P. Fua, "Robust people tracking with global trajectory optimization," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2006, pp. 744–750.

[2] S. Sternig, T. Mauthner, A. Irschara, P. M. Roth, and H. Bischof, "Multi-camera multi-object tracking by robust hough-based homography projections," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops*, 2011, pp. 1689–1696.

[3] H. Possegger, S. Sternig, T. Mauthner, P. M. Roth, and H. Bischof, "Robust real-time tracking of multiple objects by volumetric mass densities," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2013, pp. 2395–2402.

[4] A. Ess, B. Leibe, K. Schindler, and L. V. Gool, "A mobile vision system for robust multi-person tracking," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 10, pp. 1831–1846, Oct. 2009.

[5] C. Wojek, S. Walk, S. Roth, K. Schindler, and B. Schiele, "Monocular visual scene understanding: Understanding multi-object traffic scenes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 4, pp. 882–897, Apr. 2013.

[6] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Stat. Soc. Ser. B*, vol. 39, no. 1, pp. 1–38, 1977.

[7] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 6, pp. 721–741, Nov. 1984.

[8] C. Andrieu, A. Doucet, and R. Holenstein, "Particle Markov chain Monte Carlo methods," *J. Roy. Stat. Soc. Stat. Methodol. Ser. B*, vol. 72, no. 3, 2010, Art. no. 269–342.

[9] J. Roh, D. W. Park, J. Kwon, and K. M. Lee, "Visual tracking using joint inference of target state and segment-based appearance models," in *Proc. Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf.*, pp. 1–4, Nov. 2013.

[10] J. Kwon, R. Dragon, and L. V. Gool, "Tracking by switching state space models," *Comput. Vis. Image Understanding*, 2016, Accepted for Publication.

[11] R. Dragon and L. V. Gool, "Ground plane estimation using a hidden Markov model," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 4026–4033.

[12] Z. Khan, T. Balch, and F. Dellaert, "MCMC-based particle filtering for tracking a variable number of interacting targets," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 11, pp. 1805–1918, Nov. 2005.

[13] M. Isard and A. Blake, "Condensation: Unifying low-level and high-level tracking in a stochastic framework," in *Proc. 5th Eur. Conf. Comput. Vis.*, 1998, pp. 893–908.

[14] X. Jia, H. Lu, and M.-H. Yang, "Visual tracking via adaptive structural local sparse appearance model," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2012, pp. 1822–1829.

[15] D. A. Ross, J. Lim, R.-S. Lin, and M.-H. Yang, "Incremental learning for robust visual tracking," *Int. J. Comput. Vis.*, vol. 77, no. 1–3, pp. 125–141, 2008.

[16] H. Grabner, C. Leistner, and H. Bischof, "Semi-supervised on-line boosting for robust tracking," in *Proc. 10th Eur. Conf. Comput. Vis.*, 2008, pp. 234–247.

[17] B. Babenko, M.-H. Yang, and S. Belongie, "Visual tracking with online multiple instance learning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2009, pp. 983–990.

[18] S. Hare, A. Saffari, and P. H. S. Torr, "Struck: Structured output tracking with kernels," in *Proc. Int. Conf. Comput. Vis.*, 2011, pp. 263–270.

[19] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista, "Exploiting the circulant structure of tracking-by-detection with kernels," in *Proc. Eur. Conf. Comput. Vis.*, 2012, pp. 702–715.

[20] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista, "High-speed tracking with kernelized correlation filters," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 3, pp. 583–596, Mar. 2015.

[21] S. Hare, S. Golodetz, A. Saffari, V. Vineet, M.-M. Cheng, S. Hicks, and P. Torr, "Struck: Structured output tracking with kernels," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2015, Accepted for Publication.

[22] F. Pernici and A. D. Bimbo, "Object tracking by oversampling local features," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 12, pp. 2538–2551, Dec. 2014.

[23] Y. Wu, J. Lim, and M.-H. Yang, "Online object tracking: A benchmark," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2013, pp. 2411–2418.

[24] A. Doucet, N. de Freitas, K. P. Murphy, and S. J. Russell, "Rao–Blackwellised particle filtering for dynamic Bayesian networks," in *Proc. 16th Conf. Uncertainty Artif. Intell.*, 2010, pp. 176–183.

[25] N. Whiteley, "Comment on "Particle Markov chain Monte Carlo" by C. Andrieu, A. Doucet, and R. Holenstein," *J. Roy. Stat. Soc., Series B*, vol. 72, no. 4, pp. 357–385, 2010.

[26] F. Lindsten, M. I. Jordan, and T. B. Schon, "Particle Gibbs with ancestor sampling," *J. Mach. Learn. Res.*, vol. 15, no. 6, pp. 2145–2184, 2014.