Online multi-model particle filter-based tracking to study bedload transport

Hugo Lafaye de Micheaux, Christophe Ducottet, Philippe Frey

ONLINE MULTI-MODEL PARTICLE FILTER-BASED TRACKING TO STUDY BEDLOAD TRANSPORT

Hugo Lafaye de Micheaux ⋆† *Christophe Ducottet* ⋆ *Philippe Frey* †

⋆ Université de Lyon, UJM-Saint-Étienne, CNRS, UMR5516, Laboratoire Hubert Curien
F-42023 Saint-Étienne, France
† Université Grenoble Alpes, Irstea, UR ETGR, F-38402 St-Martin-d'Hères, France

ABSTRACT

Multi-object tracking is a difficult problem underlying many computer vision applications. In this work, we focus on sediment transport experiments in a flow were sediments are represented by spherical calibrated beads. The aim is to track all beads over long time sequences to obtain sediment velocities and concentration. Classical algorithms used in fluid mechanics fail to track the beads over long sequences with a high precision because they incorrectly handle both miss-detections and detector imprecision. Our contribution is to propose a particle filter-based algorithm including an adapted multiple motion model. Additionally, this algorithm integrates several improvements to account for the lack of precision of the detector. The evaluation was made using a test sequence with a dedicated ground-truth. The results show that the method outperforms state-of-the-art concurrent algorithms.

***Index Terms*—** Particle filtering, multi-model tracking, multi-object tracking, detector confidence, bedload transport.

1. INTRODUCTION

Visual target tracking is a recurrent problem in computer vision, it has been used in many applications, *e.g.* video surveillance, sports analysis, or traffic safety. In particular, object tracking is used in fluid mechanics to track particles with techniques as PIV [1] and PTV [2]. We focus here on bedload sediment transport experiments. The aim is to track all spherical beads (see Fig. 1) over a long time to obtain particle velocities and concentrations [3], for studying bedload granular rheology [4], size segregation [5] and associated morphology [6].

In this context, most of the proposed algorithms are deterministic and rely on a standard two step approach: (1) object detection, (2) object tracking by solving a data association problem [7]. However, these are not robust enough to track objects over very long sequences with a good precision for object positions. This is mainly because they incorrectly handle miss-detections and do not take into account the lack of precision of the detector.

This research is funded by Irstea, labex OSUG@2020 and by the Rhône-Alpes region as part of its Environment Academic Research Community.

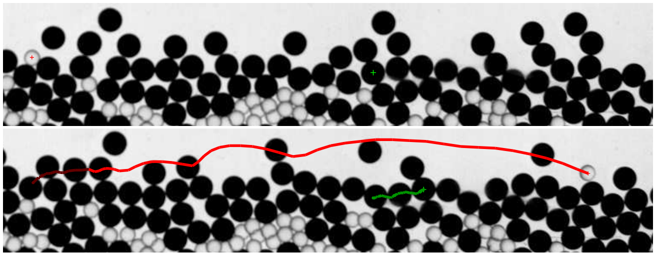

Fig. 1. Example of trajectories for two tracked beads (cropped images). Green color is for a coarse black bead and red is for a smaller transparent bead.

In recent years, particle filter-based approaches have been successfully used in many applications [8, 9, 10]. The main benefit of these methods is to enable the estimation of the distribution of the target's system state incorporating a non linear motion model. In this context, Breitenstein *et al.* [11] proposed a very efficient online multi-target tracking approach accounting for the detector confidence to handle miss-detections. However, this method is not directly adapted to our problem mainly because it uses a single motion model unable to describe all occurring object interactions.

In this paper, using the same framework, we propose to introduce a multiple motion model dedicated to our simulation of bedload transport. Each motion model corresponds to a specific state of motion of our objects: resting, rolling or bouncing. We also introduce improvements to correctly handle the lack of precision of the detector providing tracking over long sequences with a high precision for objects position.

The paper is structured as follows. After discussing the related work in next section, Sec. 3 describes our algorithm and design choices. Sec. 4 presents the evaluation results of tracking performances and a comparison with concurrent algorithms. Finally, Sec. 5 concludes the paper.

2. RELATED WORK

Various visual tracking approaches have been reviewed [12, 13] and they can be separated in two categories: detection-based methods and Bayesian filtering ones. In fluid mechanics, researchers usually employ the first one, *i.e.* taking the object detector output as a straight truth and then solving the association problem on pairwise frames [7, 14]. This deterministic approach lacks precision since it strongly relies on detectors that can be unreliable and imprecise.

Sequential monte carlo methods or particle filters were introduced to provide better estimations and predictions [15, 16, 17]. Basically, they consist of a dynamic model for prediction and an observation model to evaluate the likelihood of a predicted state. The main steps are: random generation of a sample of potential states called particles, propagation of these particles through the motion model and resampling of the distribution according to the observation model. To handle difficult tracking situations in a multi-object context, Breitenstein *et al.* [11] proposed an online and automatic particle filter-based algorithm that exploits the intermediate output of the object detector as a graded observation model. As most particle filter-based algorithms, it relies on a single motion model which can be problematic in an environment with complex dynamics.

To face this limitation, multiple dynamics models such as switching dynamical model approaches [18], have been recently developed under particle filters for robotics and navigation [19] or video surveillance [20]. These methods use the fact that object motions can be classified in different motion states according to the situation. Each motion state is dealt with specific state evolution model for prediction. However, it requires to know all motion states beforehand with a mathematical description for each. In contrast, our method exploits the mechanistic knowledge we have in our application.

3. MULTI-OBJECT TRACKING ALGORITHM

Our fully automatic multi-object tracking algorithm uses the same principle as [11]. Each trajectory is estimated by an individual particle filter (tracker), automatically initialized when a new object is detected and terminated when the object is leaving the field of view. At each time step we perform: (1) object detection, (2) association of each detection to a tracker, (3) update of the state of each particle filter according to the motion model and the new detection.

3.1. Object detector

Black beads are detected by thresholding the image and taking the center of objects having an area close to the mean bead area. As transparent beads appear as faint dark rings of different shapes because of their neighboring beads, we use a specific chain of morphological operations to detect them. We apply a *hconvex* operator [21] on the image, then a normalized cross-correlation with a ring-shape model and finally we keep the relevant maxima using an adjusted threshold. The object detector returns a set of detections D.

To handle missing detections on transparent beads, we use the detector confidence as a graded observation model [11]. Based on the raw output of the detector, the detector confidence must give an estimation of the likelihood of a detection at each candidate pixel. In our case, for a given detection $d \in D$, we propose to define the confidence value as a decreasing function of the distance to the detection position $\boldsymbol{x}_d$ with a maximum equal to the cross-correlation output $\mathrm{xcorr}(d)$ of the detector, *i.e.* the correlation value before thresholding. Formally, for each position $\boldsymbol{x}$ of the image, the detector confidence density $d_c(\boldsymbol{x})$ is defined as:

$$d_c(\boldsymbol{x}) = \frac{1}{k_D} \sum_{d \in D} \mathrm{xcorr}(d) \exp(-\lambda \|\boldsymbol{x} - \boldsymbol{x}_d\|) \qquad (1)$$

where λ denotes the exponential decay constant (see Sec. 3.4), and k_D is the normalization coefficient.

3.2. Data association

To assign at most one detection to at most one tracker, a data association process is needed. Due to the high number of detections and the long time series, a greedy algorithm was found to be an effective solution. Given the matching cost $c(tr, d)$ between all tracker-detection pairs, the greedy algorithm iteratively selects the best candidate and removes the corresponding concurrent associations. To limit the number of possible tracker-detection pairs, the set of possible detections associated to a given tracker is limited to the detections located inside a circular region centered at the predicted position $\boldsymbol{x}_{\widehat{tr}}$ given by this tracker. This predicted position is made assuming a constant velocity model.

Given a tracker tr and a detection d, the matching cost is:

$$c(tr, d) = \underbrace{\alpha \frac{\|\boldsymbol{x}_d - \boldsymbol{x}_{\widehat{tr}}\|}{r_s}}_{\text{distance term}} + \underbrace{\beta \frac{|\boldsymbol{v}_d|}{|\boldsymbol{v}_{max}|}}_{\text{velocity term}} \qquad (2)$$

where $\boldsymbol{v}_d$ is the measured velocity of detection d[1], the parameters α and β represent the proportion of each term (see Sec. 3.4), and r_s is the radius of the circular searching region calculated from the fluid velocity. The distance term promotes the detections closest to the prediction. With the velocity term, we want to promote low velocities as it allows to have better association in case of beads collisions.

3.3. Particle filtering

Each tracker is described by a fixed number N of particles, having a state $\{\boldsymbol{x}_p, \boldsymbol{v}_p, s_p\}$ where $\boldsymbol{x}_p = (x, y)$ denotes the position, $\boldsymbol{v}_p = (u, v)$ the velocity and s_p the motion state. We employ the Sequential Importance Resampling filter, also

[1]The measured velocity $\boldsymbol{v}_d$ corresponds to the displacement between the detection position $\boldsymbol{x}_d$ and the estimated position at previous time step.

known as the bootstrap filter [16], to approximate the probability function. For one target, it works as follows: (1) Estimation of particle states, positions and velocities, (2) Particle importance weighting, (3) Normalization of weights, (4) Resampling, (5) Estimation of final state, position and velocity of the target. The main difference with [11], is the use of the velocity estimation of a particle as an observation in the importance weighting.

State model. In motion-based stochastic tracking [22], explicit motion measurements are used to guide predictions. In our application, beads have different behaviors according to their location, velocity and neighborhood. We can distinguish three distinct motion states: resting (not moving), rolling (rolling/sliding on others) and saltating (bouncing on others). To propagate the particles, we first update their state and then we apply a motion model to update the position and the velocity. To update the state, we draw the new state according to a conditional probability table (see Sec. 3.4). The motion models are:

$$\text{resting: } (x,y)_t = (x,y)_{t-1} + \varepsilon^{rest}_{(x,y)} \tag{3a}$$

$$\text{rolling: } (x,y)_t = (x,y)_{t-1} + (u)_{t-1}\Delta t + \varepsilon^{roll}_{(x,y)} \tag{3b}$$

$$(u)_t = (u)_{t-1} + \varepsilon^{roll}_{(u)} \tag{3c}$$

$$\text{saltating: } (x,y)_t = (x,y)_{t-1} + (u,v)_{t-1}\Delta t + \varepsilon^{salt}_{(x,y)} \tag{3d}$$

$$(u,v)_t = (u,v)_{t-1} + \varepsilon^{salt}_{(u,v)} \tag{3e}$$

where $\varepsilon^{rest}_{(x,y)}$, $\varepsilon^{roll}_{(x,y)}$, $\varepsilon^{salt}_{(x,y)}$ are the process noises on the position and $\varepsilon^{roll}_{(u)}$, $\varepsilon^{salt}_{(u,v)}$ are the process noises on the velocity, they are all independently drawn from zero-mean normal distributions ; Δt depends on the framerate of the sequence.

Particle weighting. The importance weight $w_{tr,p}$ for a particle p of tracker tr is given by the conditional likelihood of the new observation given the propagated particle. Given detection d^* associated to the tracker, we have:

$$w_{tr,p} = \mathcal{I}(tr)\big(\underbrace{\delta p_{\mathcal{N}}(\boldsymbol{x}_p - \boldsymbol{x}_{d^*})}_{\text{particle position}} + \underbrace{\gamma p_{\mathcal{N}}(\boldsymbol{v}_p - \boldsymbol{v}_{d^*})}_{\text{particle velocity}}\big) + \underbrace{\eta d_c(\boldsymbol{x}_p)}_{\text{det. conf}} \tag{4}$$

where $\mathcal{I}(tr)$ is an indicator function that returns 1 if a detection is associated to the tracker and 0 otherwise. δ, γ and η are set experimentally (see Sec. 3.4).

The particle position (resp. velocity) term calculates the distance between particle position (resp. velocity) and the measured detection position (resp. velocity), and is evaluated under a normal distribution $p_{\mathcal{N}}$. The detector confidence term evaluates the detector confidence density at the particle position.

Resampling and estimating. After the weighting, the resampling is used to solve particle degeneracy, namely removing the particles of small weight and reproducing those of large weights [23]. Finally, the position and the velocity are estimated by averaging the resampled particles.

3.4. Implementation details

The number of particles N was fixed to $N = 100$, being a good compromise between computational cost and tracking reliability. The parameters in equations (1), (2) and (4) have been set experimentally and remained the same for the different experiments used in Sec. 4.1. λ was set to 1 in eq. (1). In eq. (2), α and β were set respectively to 0.75 and 0.25 to give more importance to the distance term. The proportion parameters (δ, γ, η) in eq. (4) were fixed at $(20, 1, 1)$. Therefore, when a detection is associated, the particle weight is mainly influenced by its position and also a little by its velocity, to reduce the imprecision of the detection position.

Each motion state in eqs. (3) has its own variance for the process noise on position and velocity. The initial particle states were set as rolling, and their initial positions and velocities were drawn from a Normal distribution with positions centered at the detection center and velocities centered to null velocity. To handle the difficulty of dealing with new trackers, we increased the variances to make the motion model more flexible during the first 3 frames. A tracker survives only 10 frames without associated detection and is then automatically terminated.

For a given particle, the transition from one state to another is controlled with conditional probabilities. We have estimated these probabilities on a representative training sequence. Concerning the particle velocity at the beginning of a track, the estimation was set to the measured particle velocity (*i.e.* the displacement between the particle position and the estimated position at previous time step).

4. RESULTS

In this section, we present the results of the proposed tracking algorithm with three motion models (3MM) (see Sec. 3.3) against two simplified versions using only one motion model: one with a constant velocity model (1CVM) and the other with a null velocity model (1NVM). These simplified versions can be considered as direct adaptations of Breitenstein *et al.* [11] algorithm. The results are compared to a ground truth to have information about the false positives and false negatives. We also use the CLEAR MOT metrics [24] to evaluate the tracking performance.

4.1. Experiments and ground truth

We worked on a 1000-frames sequence recorded at 130 fps with approximately 400 beads per frame (about 300 coarse and 100 small beads). The sequence is split in two parts, one for optimizing the parameters, and the other part for evaluation. We created a ground truth on the sequence by selecting manually the detections on each frame. To get the ground truth of trajectories, we applied a simple data association algorithm on the ground truth of detections just created.

Algo. Models	% Correct Tracks	MOTP	FN	FP	Id Sw.	MOTA
3MM	**98.35%**	**0.64px**	**0.07%**	**0.16%**	**4**	**99.77%**
1CVM [11]	95.05%	0.77px	0.09%	0.17%	23	99.72%
1NVM [11]	89.90%	0.89px	0.14%	0.54%	84	99.28%

Table 1. Tracking evaluation results for the three algorithm configurations. It shows the percentage of correct tracks (*i.e.* more than 95% of the track is correct) and the CLEAR MOT metrics [24] such as precision (MOTP), false negative rate (FN), false positive rate (FP), number of mismatches (Id Sw.) and accuracy (MOTA).

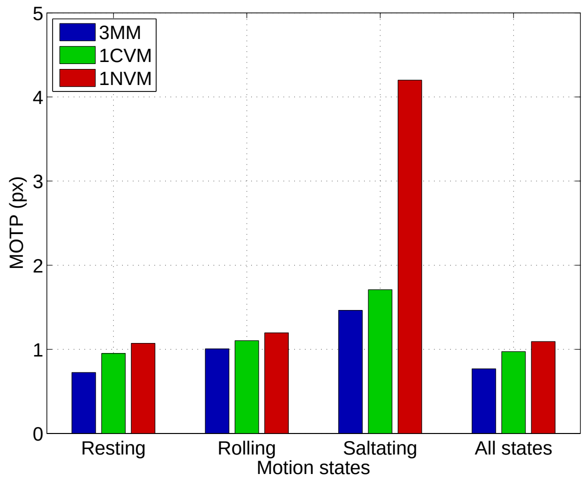

Fig. 2. Precision score MOTP on transparent beads tracking according to motion state (resting, rolling, saltating) for each algorithm model (3MM, 1CVM, 1NVM).

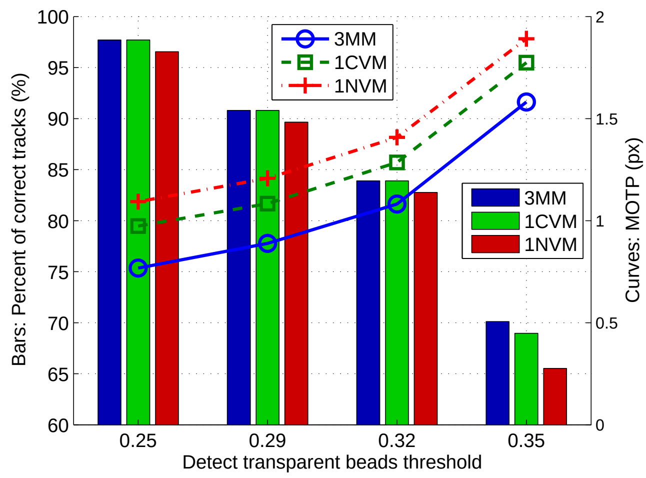

Fig. 3. Evaluation of tracking performance on transparent beads for different detection thresholds. The percentage of correct tracks (bars, left scale) and the precision score MOTP (curves, right scale) are plotted.

4.2. Evaluation of the tracking algorithm

In Tab. 1, we present the evaluation of the three tracking algorithms 3MM, 1CVM and 1NVM. The results show that 3MM is always better than the two others. Especially for the percentage of correct tracks (a track is considered as correct on the total length if more than 95% of it has no false negatives or mismatches), the 3MM reaches 98.35%. We also computed the CLEAR MOT metrics to evaluate the precision score MOTP (average error in estimated position) and the accuracy score MOTA (accounts for all object configuration errors made by the tracker, false negatives, false positives and mismatches). Here again, the 3MM appears to be the most precise and accurate.

In Fig. 2, we plotted the precision MOTP of the tracking of transparent beads for each motion state described in Sec. 3.3. The precision depends on the motion states, the higher the velocity (saltating $>$ rolling $>$ resting), the worse the precision. However, our 3MM performs better on each motion state especially on saltation.

In Fig. 3, we varied the threshold of the transparent beads detector in order to see the tracking performance with different levels of detection. It illustrates the influence of this threshold on the percentage of correct tracks and the precision MOTP on transparent beads. Again, 3MM performs better even when the miss detection rate increases due to the increasing threshold.

5. CONCLUSION

We have presented a new online particle filter algorithm based on multiple dynamic models for automatic multi-object tracking over long sequences. Having *a priori* information about the object mechanical dynamics, we were able to approach the real trajectories. This allows us to study bedload transport with higher confidence.

Our multiple motion model based algorithm provides a high tracking precision and accuracy when applied on detectors of different quality. Moreover, the approach has been shown to outperform state-of-the-art algorithms presenting a single motion model such as constant or null velocity model.

A possible extension would be to use the state of neighboring objects as an information to help choosing between motion states and bring more estimation precision. Finally, the tracking algorithm could be applied on very long sequences to observe lower frequency phenomena.

6. REFERENCES

[1] J. Westerweel, G. E. Elsinga, and R. J. Adrian, "Particle image velocimetry for complex and turbulent flows," *Annual Review of Fluid Mechanics*, vol. 45, pp. 409–436, 2013.

[2] K. Ohmi and H.-Y. Li, "Particle-tracking velocimetry with new algorithms," *Measurement Science and Technology*, vol. 11, no. 6, pp. 603, 2000.

[3] P. Frey, "Particle velocity and concentration profiles in bedload experiments on a steep slope," *Earth Surface Processes and Landforms*, vol. 39, no. 5, pp. 646–655, 2014.

[4] R. Maurin, J. Chauchat, B. Chareyre, and P. Frey, "A minimal coupled fluid-discrete element model for bedload transport," *Physics of Fluids (1994-present)*, vol. 27, no. 11, pp. 113302, 2015.

[5] V. Bacchi, A. Recking, N. Eckert, P. Frey, G. Piton, and M. Naaim, "The effects of kinetic sorting on sediment mobility on steep slopes," *Earth Surface Processes and Landforms*, vol. 39, no. 8, pp. 1075–1086, 2014.

[6] A. Recking, V. Bacchi, M. Naaim, and P. Frey, "Antidunes on steep slopes," *Journal of Geophysical Research: Earth Surface*, vol. 114, no. F4025, 2009.

[7] V. Hergault, P. Frey, F. Métivier, C. Barat, C. Ducottet, T. Böhm, and C. Ancey, "Image processing for the study of bedload transport of two-size spherical particles in a supercritical flow," *Experiments in Fluids*, vol. 49, no. 5, pp. 1095–1107, nov 2010.

[8] H. Choset and K. Nagatani, "Topological simultaneous localization and mapping (SLAM): toward exact localization without explicit localization," *Robotics and Automation, IEEE Transactions on*, vol. 17, no. 2, pp. 125–137, 2001.

[9] Y. Rathi, N. Vaswani, A. Tannenbaum, and A. Yezzi, "Tracking deforming objects using particle filtering for geometric active contours," *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, vol. 29, no. 8, pp. 1470–1475, 2007.

[10] W.-L. Lu, K. Okuma, and J. J. Little, "Tracking and recognizing actions of multiple hockey players using the boosted particle filter," *Image and Vision Computing*, vol. 27, no. 1, pp. 189–205, 2009.

[11] M. D. Breitenstein, F. Reichlin, B. Leibe, E. Koller-Meier, and L. Van Gool, "Online multiperson tracking-by-detection from a single, uncalibrated camera," *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, vol. 33, no. 9, pp. 1820–1833, sep 2011.

[12] A. W. M. Smeulders, D. M. Chu, R. Cucchiara, S. Calderara, A. Dehghan, and M. Shah, "Visual tracking: An experimental survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, vol. 36, no. 7, pp. 1442–1468, jul 2014.

[13] Y. Wu, J. Lim, and M.-H. Yang, "Online object tracking: A benchmark," in *IEEE Conference on Computer vision and pattern recognition (CVPR)*. jun 2013, pp. 2411–2418, IEEE.

[14] N. T. Ouellette, H. Xu, and E. Bodenschatz, "A quantitative study of three-dimensional lagrangian particle tracking algorithms," *Experiments in Fluids*, vol. 40, no. 2, pp. 301–313, nov 2006.

[15] A. Doucet, A. de Freitas, and N. Gordon, *Sequential Monte Carlo methods in practice*, Springer Science & Business Media, 2001.

[16] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to nonlinear/non-gaussian bayesian state estimation," in *IEE Proceedings F (Radar and Signal Processing)*. IET, 1993, vol. 140, pp. 107–113.

[17] M. Isard and A. Blake, "Condensation–conditional density propagation for visual tracking," *International journal of computer vision*, vol. 29, no. 1, pp. 5–28, 1998.

[18] M. Isard and A. Blake, "A mixed-state condensation tracker with automatic model-switching," in *IEEE 6th International Conference on Computer Vision (ICCV)*. jan 1998, pp. 107–112, IEEE.

[19] Y. Boers and J. N. Driessen, "Interacting multiple model particle filter," *IEE Proceedings-Radar, Sonar and Navigation*, vol. 150, no. 5, pp. 344–349, 2003.

[20] Y. Zhai, M. B. Yeary, S. Cheng, and N. Kehtarnavaz, "An object-tracking algorithm based on multiple-model particle filtering with state partitioning," *Instrumentation and Measurement, IEEE Transactions on*, vol. 58, no. 5, pp. 1797–1809, 2009.

[21] P. Soille, *Morphological image analysis: Principles and Applications*, Springer-Verlag Berlin Heidelberg, 1999.

[22] J.-M. Odobez, D. Gatica-Perez, and S. O. Ba, "Embedding motion in model-based stochastic tracking," *Image Processing, IEEE Transactions on*, vol. 15, no. 11, pp. 3514–3530, nov 2006.

[23] A. Milstein, J. N. Sánchez, and E. T. Williamson, "Robust global localization using clustered particle filtering," in *AAAI/IAAI*, 2002, pp. 581–586.

[24] K. Bernardin and R. Stiefelhagen, "Evaluating multiple object tracking performance: the CLEAR MOT metrics," *Journal on Image and Video Processing*, , no. 246309, 2008.