Automated 3D trajectory measuring of large numbers of moving particles

Hai Shan Wu,[1,3] **Qi Zhao,**[2] **Danping Zou,**[1] **and Yan Qiu Chen**[1,4]

[1]*School of Computer Science, Fudan University, Shanghai, China*
[2]*School of Information Science and Engineering, Fudan University, Shanghai, China*
[3]*hswu@fudan.edu.cn*
[4]*chenyq@fudan.edu.cn*

Abstract: Complex dynamics of natural particle systems, such as insect swarms, bird flocks, fish schools, has attracted great attention of scientists for years. Measuring 3D trajectory of each individual in a group is vital for quantitative study of their dynamic properties, yet such empirical data is rare mainly due to the challenges of maintaining the identities of large numbers of individuals with similar visual features and frequent occlusions. We here present an automatic and efficient algorithm to track 3D motion trajectories of large numbers of moving particles using two video cameras. Our method solves this problem by formulating it as three linear assignment problems (LAP). For each video sequence, the first LAP obtains 2D tracks of moving targets and is able to maintain target identities in the presence of occlusions; the second one matches the visually similar targets across two views via a novel technique named maximum epipolar co-motion length (MECL), which is not only able to effectively reduce matching ambiguity but also further diminish the influence of frequent occlusions; the last one links 3D track segments into complete trajectories via computing a globally optimal assignment based on temporal and kinematic cues. Experiment results on simulated particle swarms with various particle densities validated the accuracy and robustness of the proposed method. As real-world case, our method successfully acquired 3D flight paths of fruit fly (*Drosophila melanogaster*) group comprising hundreds of freely flying individuals.

OCIS codes: (150.6910) Three-dimensional sensing; (100.4999) Pattern recognition, target tracking; (110.6880) Three-dimensional image acquisition; (120.0120) Instrumentation, measurement, and metrology.

References and links

1. T. Vicsek and A. Zafiris, "Collective motion," Arxiv preprint arXiv:1010.5017 (2010).
2. C. Reynolds, "Flocks, herds and schools: A distributed behavioral model," Comput. Graph. **21**, 25–34 (1987).
3. T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, "Novel type of phase transition in a system of self-driven particles," Phys. Rev. Lett. **75**, 1226–1229 (1995).
4. I. Couzin, "Collective cognition in animal groups," Trends Cogn. Sci. **13**, 36–43 (2009).
5. M. Nagy, Z. Ákos, D. Biro, and T. Vicsek, "Hierarchical group dynamics in pigeon flocks," Nature **464**, 890–893 (2010).
6. H. Hirschmuller and D. Scharstein, "Evaluation of Cost Functions for Stereo Matching," in IEEE Computer Society Conference on Computer Vision and Pattern Recognition, (2007), pp. 1–8.
7. C. Rasmussen and G. Hager, "Probabilistic data association methods for tracking complex visual objects," IEEE Trans. Pattern Anal. Mach. Intell. **23**, 560–576 (2001).
8. I. Cox and S. Hingorani, "An efficient implementation of Reid's multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking," IEEE Trans. Pattern Anal. Mach. Intell. **18**, 138–150 (2002).

#138377 - $15.00 USD Received 22 Nov 2010; revised 16 Jan 2011; accepted 23 Jan 2011; published 6 Apr 2011

9. Z. Khan, T. Balch, and F. Dellaert, "MCMC data association and sparse factorization updating for real time multitarget tracking with merged and multiple measurements," IEEE Trans. Pattern Anal. Mach. Intell. **28**, 1960–1972 (2006).
10. A. Cavagna, A. Cimarelli, I. Giardina, G. Parisi, R. Santagati, F. Stefanini, and M. Viale, "Scale-free correlations in starling flocks," Proc. Natl. Acad. Sci. U.S.A. **107**, 11865 (2010).
11. Z. Wu, N. I. Hristov, T. L. Hedrick, T. H. Kunz, and M. Betke, "Tracking a Large Number of Objects from Multiple Views," in IEEE 11th International Conference on Computer Vision, (2009), vol. 1.
12. Anonymous, "No fruit fly an island?" Nat. Methods **6**, 395 (2009).
13. K. Branson, A. Robie, J. Bender, P. Perona, and M. Dickinson, "High-throughput ethomics in large groups of Drosophila," Nat. Methods **6**, 451–457 (2009).
14. H. Dankert, L. Wang, E. Hoopfer, D. Anderson, and P. Perona, "Automated monitoring and analysis of social behavior in Drosophila," Nat. Methods **6**, 297–303 (2009).
15. S. Fry, N. Rohrseitz, A. Straw, and M. Dickinson, "TrackFly: Virtual reality for a behavioral system analysis in free-flying fruit flies," J. Neurosci. Methods **171**, 110–117 (2008).
16. G. Maimon, A. Straw, and M. Dickinson, "A simple vision-based algorithm for decision making in flying Drosophila," Curr. Biol. **18**, 464–470 (2008).
17. D. Grover, J. Tower, and S. Tavaré, "O fly, where art thou ?" J. R. Soc. Interface **5**, 1181–1191 (2008).
18. A. Straw, K. Branson, T. Neumann, and M. Dickinson, "Multi-camera real-time three-dimensional tracking of multiple flying animals," J. R. Soc. Interface (2010).
19. C. Wang, C. Liang, and C. Lee, "Three-dimensional nanoparticle tracking and simultaneously membrane profiling during endocytosis of living cells," Appl. Phys. Lett. **95**, 203702 (2009).
20. F. Cheong, B. Krishnatreya, and D. Grier, "Strategies for three-dimensional particle tracking with holographic video microscopy," Opt. Express **18**, 13563–13573 (2010).
21. M. Piccardi, "Background subtraction techniques: a review," IEEE Trans. Syst. Man Cybern. **4**, 3099–3104 (2004).
22. B. Anderson, J. Moore, and J. Barratt, *Optimal filtering*, (Prentice-Hall, 1979).
23. S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*, (Artech House, 1999).
24. Y. Bar-Shalom, *Tracking and Data Association*, (Academic Press Professional, 1987).
25. R. Jonker and A. Volgenant, "A shortest augmenting path algorithm for dense and sparse linear assignment problems," Computing **38**, 325–340 (1987).
26. H. Kuhn, P. Haas, I. Ilyas, G. Lohman, and V. Markl, "The Hungarian method for the assignment problem," Masthead **23**, 151–210 (1993).
27. H. Du, D. Zou, and Y. Chen, "Relative Epipolar Motion of Tracked Features for Correspondence in Binocular Stereo," in IEEE 11th International Conference on Computer Vision, (2007), pp. 1–8.
28. R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, (Cambridge University Press, 2003).
29. D. Forsyth and J. Ponce, *Computer Vision: a Modern Approach*, (Prentice-Hall, 2002).
30. A. Perera, C. Srinivas, A. Hoogs, and G. Brooksby, "Multi-Object Tracking Through Simultaneous Long Occlusions and Split-Merge Conditions," in IEEE Computer Society Conference on Computer Vision and Pattern Recognition, (2006), vol. 1.
31. Q. Zhao and Y. Chen, "High Precision Synchronization Of Video Cameras Using A Single Binary Light Source," J. Electron. Imaging **18**, 040501 (2009).
32. R. Tsai, "A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-shelf TV Cameras and Lenses," IEEE J. Robot. Autom. **3**, 323–344 (1987).

1. Introduction

Complex motion pattern displayed by natural particle systems moving in three-dimensions (3D), such as insect swarms, bird flocks and fish schools, is a fascinating natural phenomenon and has long attracted a great deal of attention from scientists of a diversified range of disciplines [1]. Although many numerical models [2, 3] have been proposed to simulate such phenomena, theoretical assumptions behind them without the support of empirical trajectory data are insufficient to reveal the nature of such phenomena [4]. However, such spatial-temporal position data is rare mainly due to technological difficulties. Most recently, [5] discovered the hierarchical property of group dynamics via tracking pigeon flock in 3D through GPS devices. Although useful, it is infeasible to attach such advanced sensor devices to a group consisting of hundred individuals. Moreover, when handling groups with tiny and lightweight individuals, such as fruit flies or bees, this approach would affect their flight behaviors. By contrast, a more feasible non-contact and non-intrusive solution is to use multiple video cameras to capture the

#138377 - $15.00 USD Received 22 Nov 2010; revised 16 Jan 2011; accepted 23 Jan 2011; published 6 Apr 2011

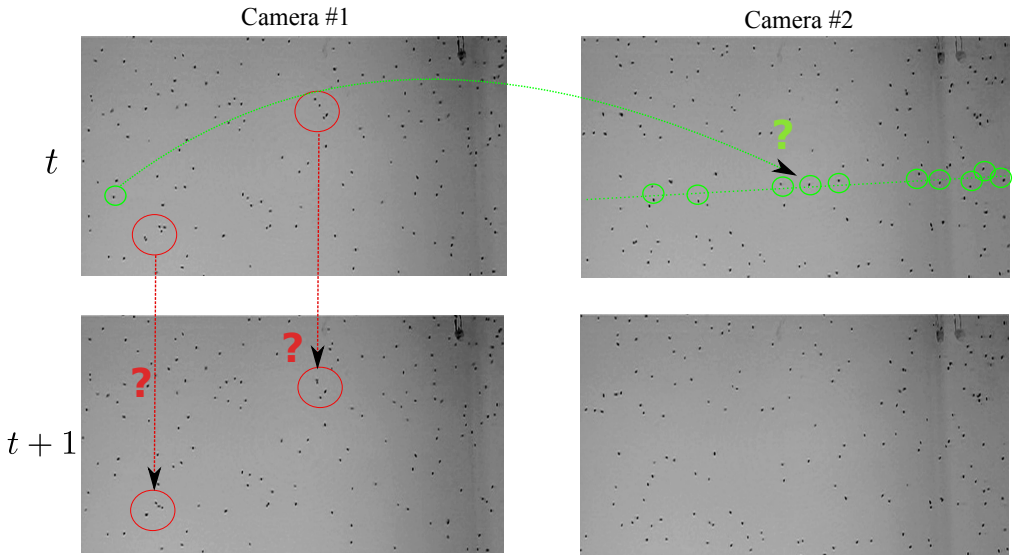

Fig. 1. Measuring 3D individual trajectory of large numbers of moving particles is challenging. As red color illustrates, tracking moving groups is difficult in the presence of occlusions, newly entering targets and clutters etc. Meanwhile, when individuals resemble each other, epipolar constraint (green line denotes the epipolar line) alone is insufficient to reduce matching ambiguity. Four images here were chosen from our *Drosophila* tracking experiments below.

dynamic scene and then to retrieve the 3D individual trajectories from video sequences.

The key challenge to this scheme, as shown in Fig. 1, is twofold: matching visually similar targets across various camera views to compute the 3D positions and maintaining their identities in the presence of occlusions throughout the whole sequences to obtain complete tracks. Therefore, current stereo matching [6] methods which fully utilize target appearance feature to reduce matching ambiguity would fail here. Besides, advanced multiple target tracking methods [7–9] were efficient to monitor a few numbers of targets, but they would lead to extremely high computational costs when tracking large and variable numbers of interacting individuals. Most recently, some researchers have attempted to overcome these difficulties: [10] have successfully reconstructed the 3D coordinates of large starling flocks, but acquiring the complete trajectory of each individual still remains a challenge to be addressed; [11] proposed a method for tracking emerging bats in three dimensions, but the accuracy and capacity of their method for handing particle systems with complex motion is unclear.

We present here an automatic and efficient method to address the above challenges by using two video cameras. As shown in Fig. 2, given the detected targets throughout the video sequence of each camera, our algorithm measures 3D trajectory of each individual in a large group by solving three linear assignment problems (LAP). For each video sequence, the first LAP obtains 2D tracks of moving targets and is able to maintain target identities in the presence of occlusions; the second one establishes the correspondences of visually similar targets across two views via a novel technique named maximum epipolar co-motion length (MECL); the last one links 3D track segments into complete trajectories via computing global assignment based on temporal and kinematic cues. MECL, which encodes both the geometric and motion features of the whole track, is not only able to effectively reduce matching ambiguity but also further diminish the influence of frequent occlusions. Our method is computationally efficient when

#138377 - $15.00 USD Received 22 Nov 2010; revised 16 Jan 2011; accepted 23 Jan 2011; published 6 Apr 2011

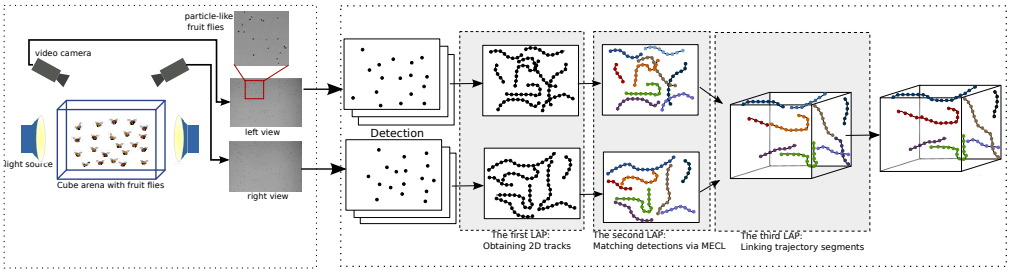

Fig. 2. Experimental setup and the diagram of the proposed method. When matching tracks across various camera views in the second step, two tracks with the same color correspond to a matched pair.

handling large numbers of targets.

We first validated the accuracy and robustness of the proposed method on simulated particle swarms with various particle densities. We then applied it to track large fruit fly (*Drosophila melanogaster*) group flying in 3D space. Recently, the flight behaviour of *Drosophila* in the group context has attracted much attention of scientists [12], and two machine vision systems proposed most recently [13, 14] were both designed to monitor walking fruit fly groups by using single video camera. Many existing automatic 3D fruit fly tracking systems [15–18] were designed to track single or at most a few numbers of subjects. Acquiring and analyzing 3D flight paths of large *Drosophila* group may provide new insights to their flight behaviours, such as collision avoidance, in group context. Using the proposed system, we successfully measured the 3D flight paths of a *Drosophila* group comprising more than five hundred flying individuals, which, to our knowledge, is the first quantitative data on such a large *Drosophila* group. We evaluated the accuracy and robustness of the proposed method by comparing the acquired results with manually generated ground truth. This method is general-purpose and can be easily adapted to a wide range of other 3D tracking tasks in group behaviour research. It can also potentially be used for tracking microparticle groups in 3D [19, 20].

This paper is organized as following: in section 2, after discussing the target detection algorithm, we present the formulation and the corresponding numerical solution to three steps of the proposed approach. Section 3 first evaluates the performance of proposed method on simulated particle swarms with different particle densities and then demonstrates the 3D tracking results of large *Drosophila* group. Discussions and conclusions are presented in the last section.

2. Problem and method

Prior to capturing the particles moving in 3D space, two video cameras we used here must be geometrically calibrated and temporally synchronized. With the recorded video sequences, we first detect the target positions by using background subtraction [21]. Taking in these detections as input, the proposed method tracks 3D motion trajectories by solving three linear assignment problems (LAPs) as illustrated in Fig. 2. The following subsections detail the formulation of the LAPs.

2.1. Two-dimensional tracking via spatial assignment

Tracking a large number of moving targets in 2D video sequences of each camera is realized through two steps. The first step is to track targets by state prediction and estimation, and the second step is to associate detections with the tracks. We use Kalman filter [22] to realize state prediction and estimation, and formulate the state-measurement association as an LAP problem.

#138377 - $15.00 USD Received 22 Nov 2010; revised 16 Jan 2011; accepted 23 Jan 2011; published 6 Apr 2011

The details are explained as follows.

Target state prediction and estimation can be achieved by Bayesian filters which compute the successive posterior densities $p(\mathbf{x}(t)|\mathbf{z}(1:t))$ from the set of detected measurements $\mathbf{z}(1:t) = \{\mathbf{z}(1),\cdots,\mathbf{z}(t)\}$, where $\mathbf{z}(t) \in \mathbb{R}^2$ is the noisy measurement of a hidden state $\mathbf{x}(t)$ at time t. In our implementation, $\mathbf{x}(t) \in \mathbb{R}^4$ contains the 2D position and 2D velocity vector. Based on first order Markov assumption, we model the problem as a linear Gaussian dynamic system in the following way:

$$\mathbf{x}(t) = \mathbf{F}\mathbf{x}(t-1) + \mathbf{w}(t) \tag{1}$$

$$\mathbf{z}(t) = \mathbf{H}\mathbf{x}(t) + \mathbf{u}(t) \tag{2}$$

where $\mathbf{F}$ is the state transition matrix from $t-1$ to t, $\mathbf{H}$ is the observation matrix. $\mathbf{w}(t)$ and $\mathbf{u}(t)$ are the process noise and measurement noise, which are assumed to be zero-mean white Gaussian noises. Let $\hat{\mathbf{x}}(t)$ be the predicted state based on $\mathbf{z}(1:t-1)$. Given the following assumptions in linear Gaussian system above: (1) the process noise and state are mutually independent; (2) the process noise and measurement noise are mutually uncorrelated and independent, the optimal solution that minimizes the mean-squared error $\mathbb{E}[||\mathbf{x}(t) - \hat{\mathbf{x}}(t)||^2]$ is the Kalman filter [22]. After state initialization, Kalman filter is constructed by two steps in each frame: state prediction and estimation. States in current frame are predicted by using the estimated states in previous frame:

$$\hat{\mathbf{x}}(t) = \mathbf{F}\mathbf{x}(t-1), \quad \mathbf{F} = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3}$$

where Δt denotes the time period of the sampling interval. After employing data association detailed below, each predicted state is assigned to one measurement. The current state is then estimated by:

$$\mathbf{x}(t) = \hat{\mathbf{x}}(t) + \mathbf{G}[\mathbf{z}(t) - \mathbf{H}\hat{\mathbf{x}}(t)] \tag{4}$$

where $\mathbf{G}$, known as Kalman gain [22], is updated recursively.

We here adopt a simplified version of Kalman filter: α-β filter [23], of which $\mathbf{G}$ is expressed as:

$$\mathbf{G} = \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \\ \frac{\beta}{\Delta t} & 0 \\ 0 & \frac{\beta}{\Delta t} \end{bmatrix}, 0 \le \alpha, \beta \le 1, \quad \mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \tag{5}$$

α and β determine the fraction of measurement error applied to the predicted position and velocity (namely the predicted state), respectively. Compared to Kalman filter, it is more flexible to choose gain parameters α and β : one can use fixed coefficients, or can vary them according to the situation of noises. Generally, the large the amount of noise is, the smaller their values should be. In our experiment below, α-β filter performed effectively with fixed parameters. This simplification, with limited loss of performance, reduces the computational costs, and thus makes it more efficient for tracking large numbers of moving particles.

Before estimating particle states, a central problem is how to determine the optimal correspondences between detected measurements and predicted states in each frame, which is known as data association problem. The simplest suboptimal method is nearest neighbor association [24], which associates the measurement with the target whose predicted position is closest. Although simple and efficient when dealing with a small number of targets, it is ineffective to track a large number of interacting targets. We here formulate this problem as a linear

#138377 - $15.00 USD Received 22 Nov 2010; revised 16 Jan 2011; accepted 23 Jan 2011; published 6 Apr 2011

assignment problem (LAP) [25] whose computationally efficient solution provides a globally optimal association result.

The goal of LAP, given a cost matrix C, is to find the assignment matrix A between the detected measurements and predicted states by minimizing the sum of assignment costs. C is non-negative and defined by the assignment costs between measurements and predicted states in each frame. We here use Euclidean distance to define the assignment cost:

$$C(i,j) = \|\mathbf{z}^i(t) - \mathbf{H}\hat{\mathbf{x}}^j(t)\| \tag{6}$$

where $\mathbf{z}^i(t)$ and $\hat{\mathbf{x}}^j(t)$ denote the ith and jth element of $\mathbf{z}(t)$ and $\hat{\mathbf{x}}(t)$, respectively. A is a binary matrix and the entry $A(i,j)$ is determined as follows:

$$A(i,j) = \begin{cases} 1 & \text{if } \mathbf{z}^i(t) \text{ is assigned to } \mathbf{H}\hat{\mathbf{x}}^j(t) \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

A is found by solving the following optimization problem:

$$A = \arg\min_{A} \quad \sum_{i=1}^{m}\sum_{j=1}^{n} C(i,j)A(i,j) \tag{8}$$

$$\text{subject to} \quad \sum_{j=1}^{n} A(i,j) = 1, \quad \sum_{i=1}^{m} A(i,j) = 1 \tag{9}$$

where m and n denote the number of elements in $\mathbf{z}^i(t)$ and $\hat{\mathbf{x}}^j(t)$, respectively. Equation (9) guarantees that each predicted state could include at most one assigned measurement and each measurement could be assigned to at most one predicted state. In order to prohibit physically impossible assignments, any element in cost matrix C whose cost exceeds a threshold C_{thr} is set to ∞. This threshold can be selected based on prior kinematic information. The LAP in each frame can be solved by the Hungarian algorithm [26].

Assignment selection is easy when the number of particles is small and constant. However, when dealing with large numbers of moving targets, the number of measurements is often variable from frame to frame because of: (1) frequent occlusions; (2) newly entering particles; (3) temporarily disappearing particles due to missed detections; (4) particles moving out of the camera's field of view; (5) false positive detections resulting from noises, etc. As a result, not all the predicted states will be assigned to corresponding measurements, and vice versa. In order to adapt the above LAP to handle these difficulties, we implement it in the following way:

1. States that find their matched measurements in current frame are updated according to Eq. (4). All these updated states are labeled as active.

2. States without matches, which might result from temporary/permanent disappearance or occlusions, are associated with dummy measurements and updated by simply assigning $\hat{\mathbf{x}}(t)$ to $\mathbf{x}(t)$. All these states are labeled as active and unmatched status. Active states remaining unmatched status for more than T_{pred} frames, which indicates the higher probability of permanent disappearance, will be labeled as inactive and terminated.

3. Measurements without matches, which may be the newly entering targets or false positive detections, are initialized as new states and labeled as active. Generally, the false positive detections last for single or at most several successive frames, therefore, they can be easily removed by discarding the tracks whose lengths are less than a threshold.

In each frame, measurements only associate with active states. T_{pred} , which mainly depends on occlusion severity, is set to $2\sim5$ frames in our experiments. As shown in the experiment part, the above strategy can effectively handle the difficulties resulting from high particle density condition.

#138377 - $15.00 USD Received 22 Nov 2010; revised 16 Jan 2011; accepted 23 Jan 2011; published 6 Apr 2011

2.2. Track matching via MECL

In order to retrieve the 3D positions of moving particles in a group, we need to establish the correspondences of detected targets across camera views. As shown in Fig. 3(a), provided the relative translation and rotation of two cameras have been determined via camera calibration [28], the set of possible matches for point p_1 in the 1*st* view is constrained to lie on the associated epipolar line [28] l in the 2*nd* view. Finding a corresponding point along epipolar line can be simplified to be a 1D searching problem after image rectification [29], which is implemented by projecting two views onto a common plane so that the epipolar lines map to horizontally aligned lines in the transformed views as shown in Fig. 3(a) . To reduce the matching ambiguities along the rectified epipolar line, existing stereo correspondence algorithms [6] fully utilize the image texture feature. These methods would fail here because the particle-like targets are indistinguishable in appearance. The following fact, as illustrated in Fig. 3(b), suggests that the acquired 2D tracks encode rich information to solve the particle matching problem: if a particle swarm moving in 3D space is captured by two geometrically calibrated and temporally synchronized cameras with common field of view, 2D projections of the same particle at the same time step will submit to epipolar constraint. Thus particle correspondence problem can be solved by matching 2D tracks across views. This can be again formulated as an LAP, whose solution provides an optimal track assignment by minimizing the sum of a defined matching cost.

Based on the above fact, we here define a novel matching cost named Maximum Epiplor Co-motion Length (MECL). Let $\Gamma_k = \{\Gamma_k^1, \ldots, \Gamma_k^{N_k}\}$ denote the trajectory set of the kth view, where N_k is the number of the trajectories. Each element in Γ_k^i is expressed as: $\Gamma_k^i(t) = (x_k^i(t), y_k^i(t), t)$, where$(x_k^i(t), y_k^i(t))$ is the tracked position in the tth frame. For convenience, let $\mathscr{T}_k^i$ be the frame index sequence of Γ_k^i, of which the first and the last elements are denoted by $\mathscr{T}_k^i(b)$ and $\mathscr{T}_k^i(e)$, respectively. After rectifying each frame of two video sequences, the frame index sequences of the matched point pairs of Γ_1^i and Γ_2^j, as shown in Fig. 3(c), are defined as:

$$\mathscr{M}_{\Gamma_1^i,\Gamma_2^j} = \{t \big| \|y_1^i(t) - y_2^j(t)\| \leq \varepsilon, t \in \{\mathscr{T}_1^i \cap \mathscr{T}_2^j\}\} \tag{10}$$

where ε is the matching error tolerance and $\{\mathscr{T}_1^i \cap \mathscr{T}_2^j\}$ is the common time span of Γ_1^i and Γ_2^j. We subsequently replace $\mathscr{M}_{\Gamma_1^i,\Gamma_2^j}$ with $\mathscr{M}$ for simplicity. $\mathscr{M}$ is said to be contiguous if all the point pairs are matched in the common time span. Generally, errors that occur in tracking module will make it noncontiguous as shown in Fig. 3(c). In such situation, $\mathscr{M}$ will be split into several contiguous subsets, each of which is one frame index sequence of the corresponding matched segment. This can be formulated as follows:

$$\mathscr{M} = \{\mathscr{M}^1, \mathscr{M}^2, \ldots, \mathscr{M}^{N_\mathscr{M}}\} \tag{11}$$

$$\mathscr{M}^p = \{\mathscr{M}^p(b), \mathscr{M}^p(b)+1, \ldots, \mathscr{M}^p(e)\} \tag{12}$$

$$\mathscr{M}^{p+1}(b) - \mathscr{M}^p(e) \geq 2 \tag{13}$$

where $N_\mathscr{M}$ denotes the number of subsets in $\mathscr{M}$. Equation (13) means that a point pair remains unmatched for at least one frame. Let $|\mathscr{M}^p|$ denote the epipolar motion length, namely the cardinality of $\mathscr{M}^p$, MECL is expressed as:

$$\text{MECL}(\Gamma_1^i, \Gamma_2^j) = \max\{|\mathscr{M}^p|\} \quad p = \{1, \ldots, N_\mathscr{M}\} \tag{14}$$

Other motion cues such as velocity information can also be incorporated into the definition of MECL. The trajectory matching cost between Γ_1^i and Γ_2^j is defined to be the negative MECL:

$$C(\Gamma_1^i, \Gamma_2^j) = -\text{MECL}(\Gamma_1^i, \Gamma_2^j) \tag{15}$$

#138377 - $15.00 USD Received 22 Nov 2010; revised 16 Jan 2011; accepted 23 Jan 2011; published 6 Apr 2011

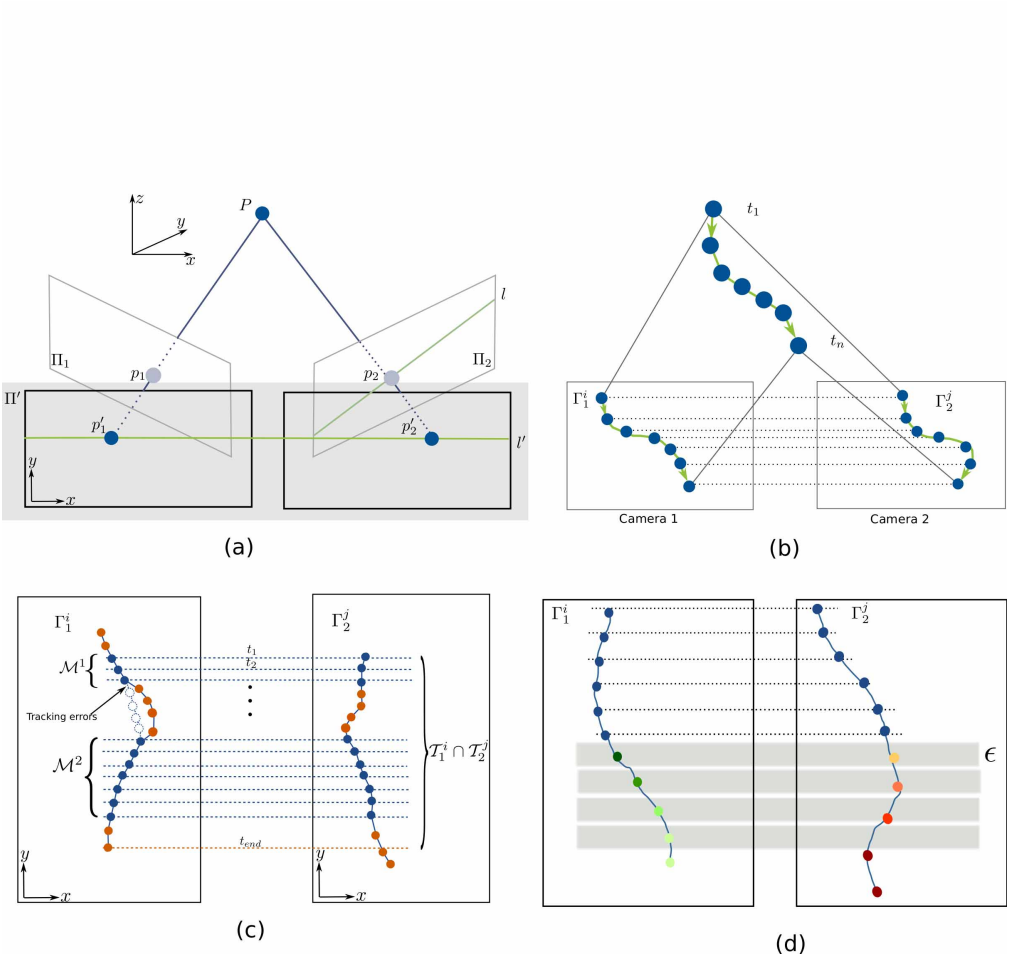

Fig. 3. (a): Epipolar constraint: given a calibrated image pair, matching candidates for point p_1 in 1*st* view must lie on the associated epipolar line l in the 2*nd* view. Image rectification is the process of projecting the planes of two views, Π_1 and Π_2, onto a common plane Π' so that the epipolar lines l map to horizontally aligned lines l' in the rectified views. The following image pairs are supposed to be calibrated and rectified. (b): Projections of the same particle on 2D image planes will submit to epipolar constraint. Γ_1^i and Γ_2^j denote two tracks obtained from two video sequences and the dashed horizontal line represents the epipolar line. (c): 2D tracks matching. Each matched point pair is marked by the same color. Unmatched point pair is generally caused by tracking errors. (d): Grey band denotes the matching tolerance band, and any point pair falling in this band is considered to be matched. Note that near the head or tail of two matched tracks, two truly unmatched points may still fall in the tolerance band for several frames.

#138377 - $15.00 USD Received 22 Nov 2010; revised 16 Jan 2011; accepted 23 Jan 2011; published 6 Apr 2011

In order to make the matching cost non-negative as required by the definition of LAP, MECL is normalized by:

$$\mathrm{MECL}(\Gamma_1^i,\Gamma_2^j)=\max\{|\mathscr{M}^p|\}\cdot(\frac{1}{\mathscr{T}_1^i}+\frac{1}{\mathscr{T}_2^j})\quad p=\{1,\ldots,N_{\mathscr{M}}\} \tag{16}$$

The trajectory matching cost is then transformed into:

$$C(\Gamma_1^i,\Gamma_2^j)=2-\mathrm{MECL}(\Gamma_1^i,\Gamma_2^j) \tag{17}$$

By solving the LAP with the above cost matrix, each track in one view will find its correspondence in the other view. As mentioned above, tracks may be partially matched due to association errors resulting from the first LAP. To diminish the influence of these errors and match the tracks as complete as possible, we perform the track matching process iteratively. For a matched track pair Γ_1^i,Γ_2^j, let $\tilde{\mathscr{M}}$ be the subset of $\mathscr{M}$ whose cardinality corresponds to $\mathrm{MECL}(\Gamma_1^i,\Gamma_2^j)$. Then each track can be decomposed into three segments:

$$\Gamma_k^a=\Gamma_k^i(\mathscr{T}_k^i(b):\tilde{\mathscr{M}}(b)-1+\tau) \tag{18}$$

$$\Gamma_k^b=\Gamma_k^i(\tilde{\mathscr{M}}) \tag{19}$$

$$\Gamma_k^c=\Gamma_k^i(\tilde{\mathscr{M}}(e)+1-\tau:\mathscr{T}_k^i(e)) \tag{20}$$

$$k\in\{1,2\},\tau\in\mathbb{Z},\tau\geq 0 \tag{21}$$

where τ is the parameter to adjust the overlap length between Γ_k^b and Γ_k^a , Γ_k^c. The value selection for τ will be discussed below. All the Γ_k^b are collected and labeled as matched, while all the Γ_k^a and Γ_k^c are labeled as unmatched and gathered to go through the matching process iteratively. In order to make the matching algorithm more robust, unmatched tracks of length less than a threshold L_{thr} are discarded, since the possibility of being erroneously matched will be very high if a track is too short.

The defined MECL differs the matching cost in [27] which was used to reconstruct 3D surface in the following ways: first, the common epipolar length we used here encodes the particle motion information during the whole common time span while the maximum distance error term used in [27] weighs only on the single time step with maximum epipolar error. And thus finding corresponding tracks in our way is more reliable as shown in experiment section. Second, compared the method in [27], MECL is able to not only handle the tracking errors but also yield more complete and correct matched track pairs by operating iteratively.

2.2.1. Selection of overlap parameter

Due to the association errors and relative motion between two particles, as shown in Fig. 3(d), truly unmatched points may still remain matched for several frames near the head or tail of one matched segment. Consequently, the resulting reconstruction errors in 3D space may be so large as to make 3D tracks linking, the last LAP of our approach, very difficult or even failed. To overcome this difficulty, we set a parameter τ to adjust the overlap length between Γ_k^b and Γ_k^a, Γ_k^c. $\tau=0$ means that there is no overlap; in real scenario, its value should be a small positive integer to generate a short overlap region.

2.3. 3D track linking via temporal assignment

With the obtained 2D matched track pairs, 3D trajectories are then reconstructed by performing triangulation via Direct Linear Transformation algorithm [28]. However, partially matched track pairs resulting from tracking errors will make the reconstructed 3D trajectories broken

#138377 - $15.00 USD Received 22 Nov 2010; revised 16 Jan 2011; accepted 23 Jan 2011; published 6 Apr 2011

into segments. The last LAP is designed to link these tracklets (trajectory segments) into complete motion trajectories. We use $\Gamma = \{\Gamma^1, \Gamma^2, \ldots, \Gamma^{N_\Gamma}\}$ to denote the 3D tracklet sets where $\Gamma^k(t) = \{x^k(t), y^k(t), z^k(t), t\}$, $\mathscr{T}^k$ to denote the frame index sequence of Γ^k. Suppose that $\tilde{\Gamma}^j = \{\Gamma^a \rceil \Gamma^b \rceil \Gamma^c\}$ is a complete trajectory broken into three tracklets where the symbol $\rceil$ represents the linking relation, then one way to recover $\tilde{\Gamma}^j$ is through determining the pairwise linking of trackets, namely, $\Gamma^a \rceil \Gamma^b$ and $\Gamma^b \rceil \Gamma^c$. Consequently, this problem is reduced to an LAP, that is how to find the optimal pairwise tracklet assignment among Γ given a linking cost matrix C.

The definition of tracklet linking cost is the key and we here incorporate temporal and kinematic information into linking cost. First, temporal cost C_t, which is used to remove the impossible candidates, is defined as follows:

$$C_t(i,j) = \begin{cases} 1 & 1 \le \mathscr{T}^j(b) - \mathscr{T}^i(e) \le \delta \\ 1 & 0 \le \mathscr{T}^i(e) - \mathscr{T}^j(b) \le \tau \\ \infty & \text{otherwise} \end{cases} \tag{22}$$

As discussed in previous subsection, in order to handle tracking errors, an adjustable overlap region will be generated when removing the matched part of one track. Consequently, Γ^i may overlap linking candidate Γ^j for at most τ frames. Meanwhile, Γ^i may also link Γ^j if the latter is initialized at most δ frames later. Second, kinematic cost captures motion information of two tracklets. As shown in Fig. 4, if two tracklets overlap each other, we choose the mean pairwise Euclidean distance in overlap region as the kinematic cost, that is:

$$C_k(i,j) = \frac{\sum_{t=\mathscr{T}^j(b)}^{\mathscr{T}^i(e)} \|r^i(t) - r^j(t)\|}{\mathscr{T}^i(e) - \mathscr{T}^j(b) + 1} \tag{23}$$

where $r^i(t) = \{x^i(t), y^i(t), z^i(t)\}$. Otherwise, if tracklet Γ^j is initialized at most δ frames after Γ^i terminates, the positions of Γ^i in broken time region are forwardly predicted by constant velocity model, and Γ^j are backwardly predicted in the similar way. The kinematic cost is then computed in the following way:

$$C_k(i,j) = \frac{\sum_{t=\mathscr{T}^i(e)}^{\mathscr{T}^j(b)} \|\tilde{r}^i(t) - \tilde{r}^j(t)\|}{\mathscr{T}^j(b) - \mathscr{T}^i(e) + 1} \tag{24}$$

where $\tilde{r}^i(t)$ denotes the predicted 3D particle positions.

The final linking cost function is the product of temporal cost and kinematic cost, namely $C(i,j) = C_t(i,j) \times C_k(i,j)$. Unlike the cost definitions of two previous LAPs, linking cost is asymmetric, that is $C(i,j) \neq C(j,i)$, because linking relation is directional and irreversible. With the obtained cost matrix, Hungarian algorithm is employed to find the optimal assignments. The complete trajectories can be obtained by tracing the resulting assignment matrix.

3. Experimental results

We first tested the performance of the proposed system by tracking simulated 3D particle swarms with different particle densities. We then applied our system to acquire the 3D motion trajectories of a *Drosophila melanogaster* (fruit fly) group comprising hundreds of flying individuals. The proposed method was implemented by MATLAB and all the experiments were conducted on a PC running Intel Dual-core 2.53 GHz Processor and 2G RAM.

#138377 - $15.00 USD Received 22 Nov 2010; revised 16 Jan 2011; accepted 23 Jan 2011; published 6 Apr 2011

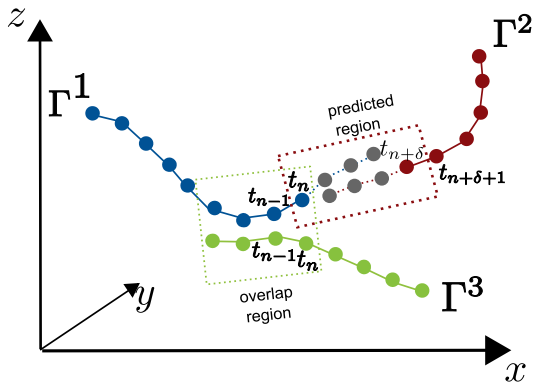

Fig. 4. Linking 3D tracklets. Tracklet Γ^1 has two linking candidates, Γ^2 and Γ^3. Γ^2 starts several frames after Γ^1 terminates, and Γ^3 overlaps Γ^1 several frames. The linking cost between Γ^1 and Γ^2, denoted by $C(1,2)$ is computed according to Eq. (22) and Eq. (24), while $C(1,3)$ is computed according to Eq. (22) and Eq. (23). Grey dots denote the predicted particle positions.

3.1. Object detection

To validate the object detection algorithm mentioned at the beginning of section 2, we compared the detection results with the ground-truth generated by human visual examination. We chose 30 frames from the whole sequences with resolution 960×540 which recorded the flying *Drosophila* group as illustrated below, and detected the targets manually. The correlation between the detection number and the ground-truth was 0.989. The discrepancy mainly resulted from the occlusions, but our data association method, as discussed in section 2.1, is able to effectively handle these situations.

3.2. Results on simulated experiments

3.2.1. Simulated particle swarm settings

We first evaluated the proposed framework on simulated particle swarms, which provided a controllable setting with known ground truths to simulate the challenges mentioned in introduction section. Simulated particles moving in a cube of edge length 2 *m* were initialized with a random position **p**, velocity **v** and a small random perturbation **n**. Particle velocity was updated as follows:

$$\mathbf{v}(t+1) = \theta \mathbf{v}(t) + \mathbf{n} \tag{25}$$

where $\theta \in [0,1]$ was used to modulate the randomness of particle motion, and $\mathbf{n} \sim \mathcal{N}(0, 0.05\mathbf{I})$ obeyed a normal distribution (with mean zero and covariance $0.05\mathbf{I}$). Particles moved randomly when $\theta = 0$ and moved with constant velocity when $\theta = 1$. In the experiment, it was randomly generated from 0.7 to 0.9. If a particle hits a boundary at the next time step, the component of velocity parallel to that boundary is unchanged, and the component of velocity perpendicular to that boundary is reversed. Let the sampling time interval be 0.005 second, the initial velocity magnitudes were randomly distributed between 1.5 and 3.5 m/s. At each sampling interval, 3D positions of all particles were projected onto two image planes of 500×500 resolution from different views, simulating two cameras. CCD video camera noises were not simulated here because the resulting false positive detections generally had very short tracks (less than 3 frames), which could be easily removed. Each particle was rendered by OpenGL as a sphere with radius of 0.02 *m* (1% of side length), and the corresponding 2D projection on image plane

#138377 - $15.00 USD Received 22 Nov 2010; revised 16 Jan 2011; accepted 23 Jan 2011; published 6 Apr 2011

resembled a circular blob with radius of at most 5 pixels. Occlusions occurred when projections of two particles overlapped each other on 2D image sequences.

By increasing the particle number, simulated challenges would be similar or even harsher compared to our real world experiment below. The particle number varied from 20 to 100, and 150 frames were used in the experiments. On average, 2 to 10 particles occluded each other in each frame.

3.2.2. Performance evaluation metrics

To evaluate the method quantitatively, we associated the obtained trajectories Γ_{EXP} with ground-truth trajectories Γ_{GT} by using the approach proposed by [30]. Let $O(i,j)$ denote $\mathscr{T}^i_{EXP} \bigcap \mathscr{T}^j_{GT}$, namely the overlap region of these two trajectories. The distance between Γ^i_{EXP} and Γ^j_{GT} was defined by:

$$D(\Gamma^i_{EXP}, \Gamma^j_{GT}) = \frac{1}{|O(i,j)|} \sum_{t \in O(i,j)} ||r^i_{EXP}(t) - r^j_{GT}(t)| \tag{26}$$

where $r^i(t) = \{x^i(t), y^i(t), z^i(t)\}$. D measures the average error between the acquired positions and ground-truth in overlap region. With the above trajectory distance matrix, association matrix A could be obtained and $A(i,j) = 1$ when the obtained trajectory Γ^i_{EXP} is associated with Γ^j_{GT}. In order to guarantee high accuracy of the acquired trajectories, associations with trajectory distance larger than a threshold D_{thr} would be forbidden. D_{thr}, the error tolerance between Γ_{EXP} and Γ_{GT}, was set to be 0.01. Let $\tilde{A}(\Gamma^j_{GT}) = \{\Gamma^i_{EXP} | A(i,j) = 1\}$, two evaluation metrics, which were used to evaluate the tracking algorithms in [30], were adopted here, that is trajectory fragmentation factor (TFF) and trajectory completeness factor (TCF):

$$\text{TFF} = \frac{\sum_j |\tilde{A}(\Gamma^j_{GT})|}{|\{\Gamma^j_{GT} | \tilde{A}(\Gamma^j_{GT}) \neq \emptyset\}|}; \quad \text{TCF} = \frac{\sum_j \sum_{i|A(i,j)=1} |O(i,j)|}{\sum_j |\mathscr{T}^j_{GT}|} \tag{27}$$

TFF, of which the ideal score is 1, measures on average the number of acquired trajectories used to match one ground-truth trajectory. The larger this value is, the worse the performance on maintaining particle identity is. TCF measures on average the ratio of one ground truth trajectory length covered by the reconstructed trajectories. TCF equals to 1 when all the motion trajectories are accurately obtained. The smaller its value is, the larger missed part is.

3.2.3. Results on simulated particle swarms

We here validated the performance of each LAP of the proposed method. We first detected the particle positions in two simulated image sequences by the object detection method detailed in section 2 and then took in the detections to acquire 3D motion trajectories. We referred our method as **TraMaL** since three LAPs were designed to **Tra**ck particles in video sequences, **Ma**tch particles across various views and **L**ink the 3D track segments, respectively. Model parameter settings of the proposed method were shown in Table 1, and compared methods used to test three modules of our framework were shown in Table 2.

Table 1. Model parameter settings

	α	β	T_{pred}	ε	L_{thr}	τ
Simulated particle tracking	0.9	0.8	3	18	20	5
Fruit fly tracking	0.8	0.7	4	6	8	2

#138377 - $15.00 USD Received 22 Nov 2010; revised 16 Jan 2011; accepted 23 Jan 2011; published 6 Apr 2011

Table 2. Compared methods

Method	Tracking	Matching	Linking
TraMaL (ours)	proposed	proposed	proposed
TraRemL	proposed	REM [27]	proposed
TraMa	proposed	proposed	None
CnMaL	**C**onstant velocity tracking and **N**earest neighbor association [24]	proposed	proposed

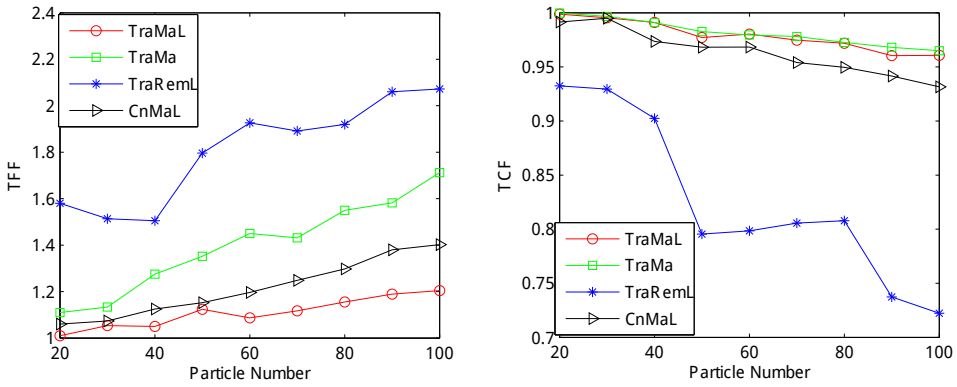

Fig. 5. Performance of compared methods on two evaluation metrics: (a) TFF and (b) TCF. One can see that TraMaL, the proposed method, performs best.

Each experiment was performed 5 times in each particle density and the average values of TCF and TFF were used for comparison. As shown in Fig. 5, our method performed best in terms of both TCF and TFF. The performance of our tracking module was better than that of CnMaL which tracked particles using constant velocity model and associated the detections with tracks via nearest neighbor algorithm [24].

Compared with TraRemL which adopted REM [27] as the trajectory matching cost, our results were remarkably better in both TFF and TCF. This is because MECL, which combines motion and geometric cues into the cost definition, is able to handle tracking errors and match track segments as complete and accurate as possible. This ability was demonstrated in Fig. 6. If only tracking and matching module were used (TraMa), TFF of TraMa was larger than that of TraMaL, and the discrepancy was more obvious when particle number increased. When the particle number was 100, TFF of TraMa was 1.71 while TFF of the proposed method was only 1.18, which indicated that the linking module could make final trajectories more complete. Results as shown in Fig. 6 demonstrated how the matching and linking module worked collaboratively.

Acquired trajectories of simulated particles using the proposed framework were shown in Fig. 7.

3.3. Results on fruit fly swarm

The flight behavior of fruit fly in group context has captured the interests of many scientists and most current experiments were carried out by tracking single or at most several subjects in 3D space as mentioned in introduction part. We applied our system to acquire the 3D flight paths of a large fruit fly (*Drosophila melanogaster*) group consisting of hundreds of flying individuals.

#138377 - $15.00 USD Received 22 Nov 2010; revised 16 Jan 2011; accepted 23 Jan 2011; published 6 Apr 2011

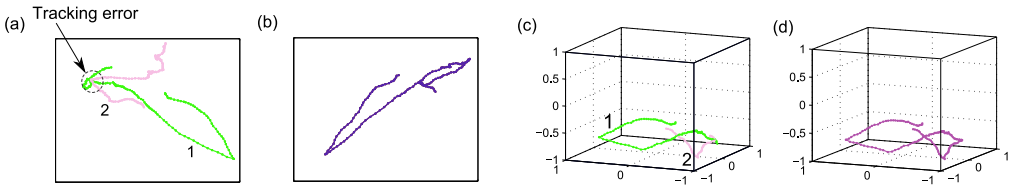

Fig. 6. Matching and linking results of simulated particle swarms. For clarity, only partial results are shown here. (a) and (b) represent the left and right camera view, respectively. Because of tracking errors, parts of two tracks in (a), namely 1 and 2, are both matched with one track in (b). As a result, two 3D trajectory segments are obtained as shown in (c). After performing tracklet linking, a correct and complete trajectory is acquired as shown in (d).

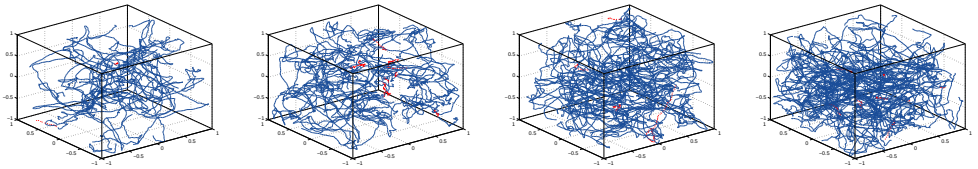

Fig. 7. 3D trajectory acquisition results of the proposed method on simulated particle swarms under different particle densities. From left to right, the particle number is 40, 60, 80 and 100, respectively. The missed and erroneous trajectories are displayed by red color.

3.3.1. Experimental setup

The experiments setup is illustrated in Fig. 2. About six hundred *Drosophila* flied freely in a transparent acrylic box of size 35cm×35cm×25cm, where the background was illuminated by white plane lights. Two Sony HVR-V1C camcorders, synchronized by the method proposed by [31] and calibrated by the algorithms developed by [32], were used to capture the scene from two different views at 20 frame per second. During experiment, both camcorders were still and the lighting was constant. As shown in Fig. 2, fruit flies appeared like dark particles on a relatively bright background. 200 frames with resolutions of 960×540 were captured totally and used to acquire the 3D motion trajectories.

3.3.2. Experimental results

Before performing tracking module, we first detected fruit flies in each frame by the object detection method detailed at the beginning of section 2 and validated in section 3.1. On average, 159.8 and 179.5 targets were detected in each frame in left and right cameras views, respectively. We observed that the *Drosophila* group here exhibited complex motion: (1) Freely flying *Drosophila* performed "body saccades", in which they rapidly turned their heading. The projected positions in sequences were thus more complicated to track. (2) The motion properties among individuals were different: some *Drosophila* flied rapidly while some moved slowly along the box edge. Facing these challenges, our framework still worked efficiently and achieved satisfying results. As shown in Fig. 8, our data association method was able to reliably maintain the target's identity under frequent occlusions. We obtained 489 and 561 2D tracks in two video sequences, respectively, whose lengths were more than 8 frames. After matching these 2D tracks via normalized MECL, we obtained 547 matched pairs. After building the linking cost function with the reconstructed 3D tracklets, the solution to the third LAP linked

#138377 - $15.00 USD Received 22 Nov 2010; revised 16 Jan 2011; accepted 23 Jan 2011; published 6 Apr 2011

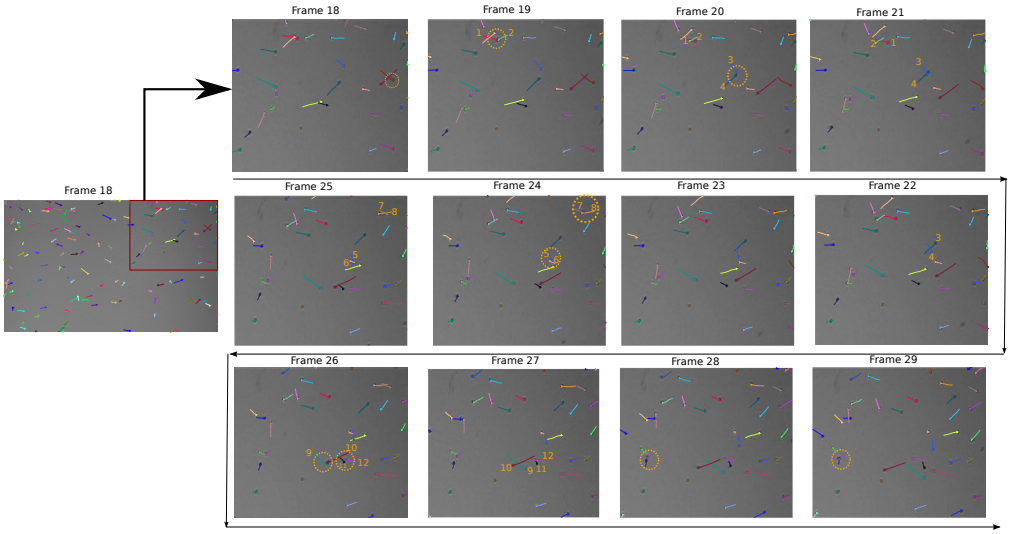

Fig. 8. Tracking fruit flies through occlusions (zooming in for more details). Enlarged parts of frame 18-29 of left view are shown here. Dashed circles suggest the occurrence of occlusions. One can see that the identities of fruit fly were correctly maintained, although the occlusions last for several successive frames. Different colors represent different tracks. Triangle represents the location in current frame and fruit flies failing to be detected because of occlusions etc. are denoted by squares.

120 fragments into 54 complete trajectories as shown in Fig. 9. The mean trajectory lengths of these segments were 36.9 frames, and by contrast the mean trajectory lengths of the linked ones were 84.0 frames. By visual inspection, one can see that these linked trajectories are seamlessly smooth at the adjoining points. As shown in Fig. 10(a), we finally acquired 458 trajectories in total. 22 trajectories of length longer than 150 frames were colored by velocity magnitude and shown in Fig. 10(c).

3.3.3. Comparison with ground truth

Because the proposed method acquires the trajectories automatically, it is important and necessary to compare the results with ground truth. A direct way to collect the ground truth 3D tracks is manually detecting the targets, establishing detection correspondences across two views frame by frame and then associating the 3D positions between consecutive frames to form complete trajectories. Given that each frame consisted of almost two hundred *Drosophila*, this process would be prohibitively labor-intensive and error prone. We accomplished this task in a semi manual way: firstly, we manually corrected detection and tracking errors from the results obtained by the proposed method; secondly, we associated the 2D tracks across two cameras views by human visual inspection; finally, with the matched 2D track pairs and the cameras parameters obtained from camera calibration, we computed the 3D trajectories, which were projected on two image planes again to check the accuracy of manual tracking. Using these trajectories as ground truth data, we finally collected 505 3D trajectories of length 27860 frames in total, while the proposed algorithm obtained 481 3D trajectories of length 27598 frames, among which 458 trajectories of length 27010 frames were correctly reconstructed, that is, 96.9% of the ground truth (the corresponding TCF was 0.969) were correctly recovered. Among the acquired trajectories, 23 trajectories of length 588 frames were incorrect, which only occupied 2.1% of all the acquired trajectories. The comparison results are summarized in

#138377 - $15.00 USD Received 22 Nov 2010; revised 16 Jan 2011; accepted 23 Jan 2011; published 6 Apr 2011

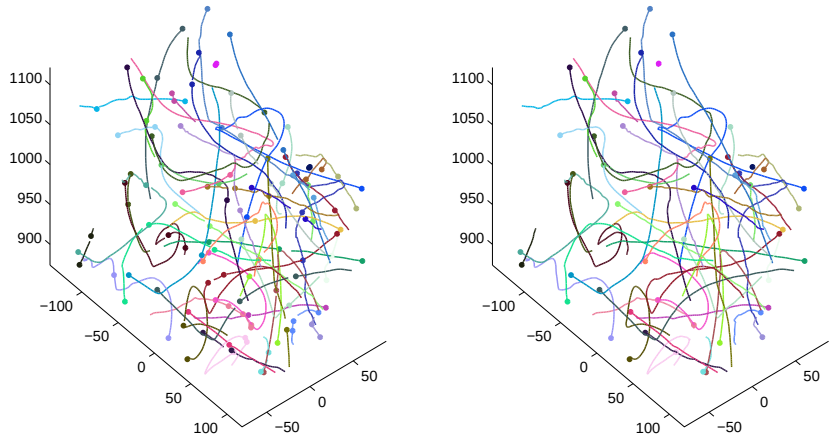

Fig. 9. Linking 3D track segments into complete trajectories (axis units: mm). **Left:** 120 linking candidates in all. **Right:** After performing linking module, 54 complete trajectories are obtained. One can see that these linked trajectories are seamlessly smooth at the adjoining points. Trajectory beginning point is marked by a round dot.

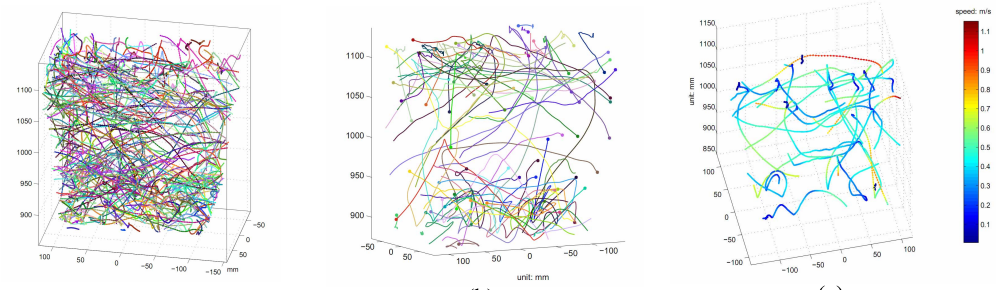

Fig. 10. (a):458 3D trajectories of the *Drosophila* group were finally acquired. (b): 73 trajectories which are longer than 100 frames are demonstrated here. Trajectory of the same fruit fly is marked by the same colour. (c): 22 trajectories which are longer than 150 frames are coloured by velocity magnitude. (axis unit: mm).

#138377 - $15.00 USD Received 22 Nov 2010; revised 16 Jan 2011; accepted 23 Jan 2011; published 6 Apr 2011

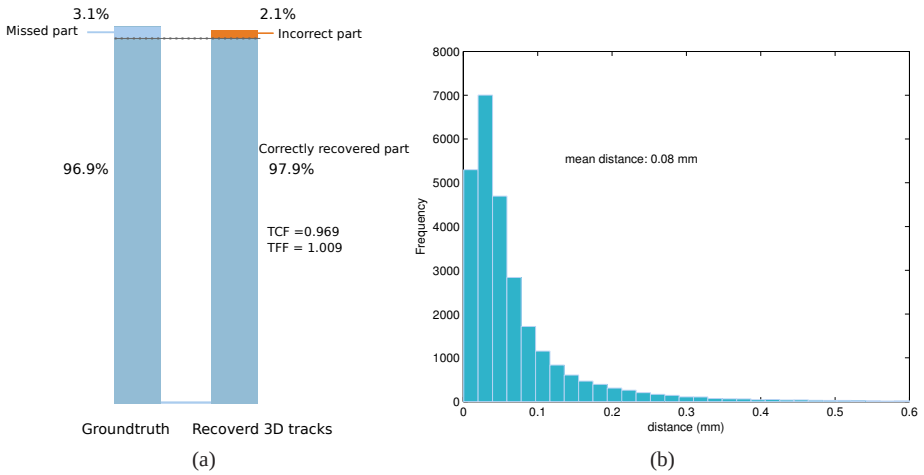

Fig. 11. (a): Comparison between the ground truth and acquired trajectories. The proposed method successfully recovered 96.9% of the ground truth. (b): Frequency distributions of the errors between ground truth and acquired trajectories. The mean error is 0.08 mm.

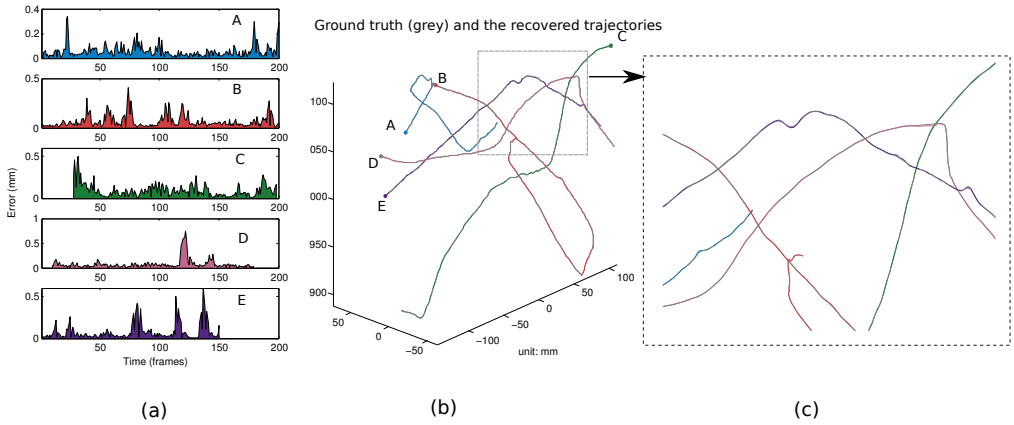

Fig. 12. (a) : Errors between 5 acquired trajectories and the ground truth shown in (b). To detail the difference between them, (c) shows the zooming-in part of the rectangle region in (b).

Fig. 11(a). As for the track completeness, only 4 trajectories in ground truth were broken into 8 track segments in the acquired results, and the TFF was 1.009 here, which indicated that the proposed method was reliable to maintain the target identity in 3D space. We used the distance in each frame between ground truth and the 458 correctly recovered trajectories to validate the tracking accuracy. The distance distribution was shown in Fig. 11(b), and the average distance was 0.08 mm. Some examples were shown in Fig. 12. The whole process took about 196 seconds on the platform illustrated in experiment section. These results indicate that the proposed approach is accurate and efficient to track large groups of moving particles in 3D space.

#138377 - $15.00 USD Received 22 Nov 2010; revised 16 Jan 2011; accepted 23 Jan 2011; published 6 Apr 2011

4. Discussions and conclusions

We presented an automatic and efficient algorithm for accurately measuring 3D motion trajectories of large numbers of moving particles using two video cameras. This method uses one compact LAP framework and is computationally efficient by employing Hungarian algorithm with cubic complexity. Three LAPs work collaboratively to reliably track large numbers of moving particles in 3D space, and the proposed stereo matching technique, MECL, is able to not only match visually similar objects across various camera views but also diminish the influence of tracking errors. The proposed system accurately measured 3D flight paths of a flying *Drosophila* group comprising hundreds of individuals, the further analysis of which may provide new insights to their flight behaviours in group context.

One limitation of the proposed approach is that it requires the collection of the whole video sequences in order to acquire the 3D motion trajectories. This off-line tracking fashion, however, applies for many scenarios in group behavior research where real-time monitoring is not desired. The proposed method, which achieved satisfying experimental results by using two video cameras, can be straightforwardly extended to multi-view scenario when more than two cameras are used. One intuitive way is by implementing the proposed algorithm in each pair of views. Our algorithm is also applicable to a broad spectrum of other 3D tracking tasks. When dealing with flocking birds or schooling fish, for example, one is only required to provide the video sequences captured from calibrated and synchronized cameras. The proposed method is also potential to be used for tracking microparticle systems.

Acknowledgements

The research work presented in this paper is supported by National Natural Science Foundation of China, Grant No. 60875024; Education Commission of Shanghai Municipality Grant No. 10ZZ03; Science and Technology Commission of Shanghai Municipality, Grant No. 09JC1401500; and Shanghai Leading Academic Discipline Project, Project Number B114. We would like to thank Linguo Li and Wei Li (Fudan University, China) for providing fruit flies in the experiments. We are also grateful to Iain Couzin (Princeton University, USA) for his helpful feedback. We also acknowledge two anonymous reviewers for their valuable comments.

#138377 - $15.00 USD Received 22 Nov 2010; revised 16 Jan 2011; accepted 23 Jan 2011; published 6 Apr 2011