Chapter 9
Crowd Flow Segmentation Using Lagrangian Particle Dynamics

Saad Ali and Mubarak Shah

Abstract A crowd of people is composed of *groupings* that arise due to interdependence among its members. Advanced visual surveillance and monitoring capabilities for crowded scenes can make use of this inherent group-based composition of human crowds to understand its global motion dynamics and to compartmentalize it into sub-parts for detailed analysis. In this chapter we propose an algorithm that uses motion information to locate such distinct crowd groupings in terms of flow segments in videos of *large dense crowds*. The flow segments are located using a particle-based representation of the motion in the video. This representation enables detection of boundaries between dynamically distinct crowd groupings.

9.1 Introduction

A crowd of people is composed of *groupings* that arise due to interdependence among its members [2, 14]. This interdependence could be a result of a social relationship (e.g. members of the same family or close circle of friends), a common purpose (e.g. to walking towards the same exit door) or an act of participating in a collective activity (e.g. running in a marathon). Advanced visual surveillance and monitoring capabilities for crowded scenes can make use of this inherent group-based composition of human crowds to understand its global motion dynamics and to compartmentalize it into sub-parts for detailed analysis. In this chapter we

S. Ali (✉)
Center for Vision Technologies, SRI International, 201 Washington Road, Princeton, NJ, USA
e-mail: saad.ali@sri.com

M. Shah
Center for Research in Computer Vision, University
of Central Florida, 4000 Central Florida Blvd., Orlando, FL 32816, USA
e-mail: shah@crcv.ucf.edu

S. Ali et al. (eds.), *Modeling, Simulation and Visual Analysis of Crowds*, The International Series in Video Computing 11, DOI 10.1007/978-1-4614-8483-7_9,

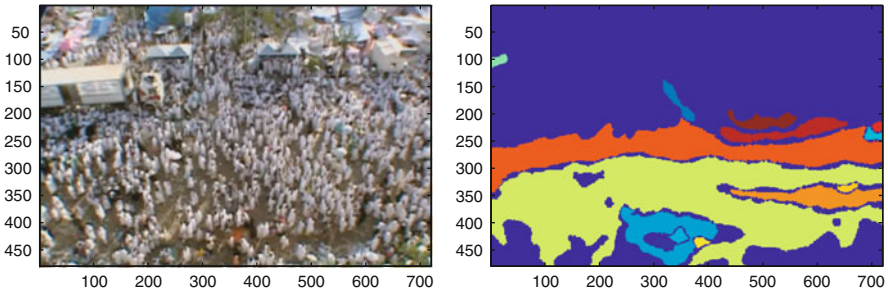

Fig. 9.1 *Left*: A frame depicting groups of people walking in multiple directions. *Right*: The crowd groupings (segments) located by the proposed algorithm

propose an algorithm that uses motion information to locate (or segment) such distinct crowd groupings in terms of flow segments in videos of *large dense crowds*. Figure 9.1 show an instance of a crowded scene where the proposed algorithm discovered various flow segments that belong to distinct crowd groupings in the scene.

Flow segment-based, and in turn group-based, visual analysis of crowded scenes provides multiple benefits: (i) enables a more elaborate and clutter free visualization of various moving groups of people in the scene; (ii) overcomes shortcomings of traditional 'detection and tracking' surveillance approaches that rely on accurate detection of each individual in the scene; (iii) mitigates influence of number of pixels on an individual person and is able to provide reasonable insight into motion of large crowds even at low resolutions.

Lagrangian Particle-based Representation: For segmenting crowd flows, the key idea developed in this chapter is a particle-based representation of the motion in the video. This representation enables detection of boundaries between various dynamically distinct crowd groupings. These boundaries, which are otherwise invisible or imperceivable to human eye, naturally emerge when people walk in different directions or at different speeds.

The proposed particle-based representation consists of particle trajectories that are obtained by examining a cloud of particles (usually in the form of a regular grid) as it mixes and gets transported over time under the action of optical flow generated by the crowd motion. The process of particle propagation using optical field (or motion field in general) is called 'advection'. If we assume that this optical flow field is generated by a certain underlying dynamical system (whose exact form and description is unknown) then one can use these trajectories to reveal representative characteristics of the phase space of this dynamical system where phase space is defined as a space of variables using which all possible states of a dynamical system are represented. The characteristics can include locations of the barriers, mixing properties, location of sources, and sinks in the phase space. Under our assumption the phase space is directly related to the flow field of the crowd, these characteristics can be mapped directly to physical properties of the crowded scene. For example, a barrier in the phase space maps either to a physical obstacle in the scene or to a boundary between crowd groups moving in different directions.

Formally if we assume that the underlying dynamical system is a non-autonomous dynamical system then the barriers are the invariant manifolds of the phase space and are often called *Coherent Structures (CS)* [4]. Generally speaking, Coherent Structures (CS) are separatrices/material lines (i.e. a boundary having two different types of flows on opposite sides) that influence the kinematics of the particle cloud over finite time intervals, and they divide the flow, and in turn the phase space, into dynamically distinct regions where all the particles within the same region have a similar fate or, in other words, coherent behavior. Intuitively speaking, coherent structure is to optical flow data what "edge" is to image data. Note that when coherent structures are studied in terms of quantities derived from particle trajectories, they are named as *Lagrangian Coherent Structures* (LCS).

Note that there are two approaches by which the field of motion can be described: (i) Lagrangian, and (ii) Eulerian. In the Lagrangian approach, properties of the flow are gathered along the path taken by a particle, while in the Eulerian approach properties of the flow are observed at a fixed spatial location. Since in our case particles are allowed to move under the influence of the optical flow, we call our representation a 'Lagrangian particle-based representation'.

LCS Detection: In order to develop an algorithm for detection of LCS (or boundaries between distinct crowd groupings) we make use of several advances in the areas of nonlinear dynamical systems [5, 12], fluid dynamics, [4, 6, 17] and turbulence theory [7,11]. In these disciplines several approaches have been proposed to compute LCS based on whether the underlying dynamical system is periodic [15], aperiodic [3], or quasi-periodic. The crowd movements are generally aperiodic (i.e. time dependent) in a generic setting as there are no or little prior constraints on its speed and direction over longer durations of time. In this chapter we employ the *Lyapunov Exponent* (LE) approach to locate LCS of the phase space. The LE measures the exponential rate of convergence or divergence between two particle trajectories. For a given crowd video, we use a grid that covers the optical flow field of the video and compute the finite-time estimate of Lyapunov Exponents (LEs) for trajectories starting at each point of the grid. This process returns a finite-time scalar Lyapunov Exponent (FTLE) field over the phase space. We use the result by Haller [4] that show coherent structures appearing as ridges in the FTLE field. In turn these ridges can be used as the boundary between various dynamically distinct crowd groupings for segmentation purposes.

We compute two types of LCS: (1) attracting LCS and (2) repelling LCS. The attracting LCS, represented by a forward FTLE field, are computed by advecting the particle grid forward in time, while the repelling LCS, represented by a backward FTLE field, are computed by advecting the particle cloud grid in time. The two FTLE fields are combined to generate a single scalar field that is segmented using an image segmentation algorithm (e.g., a watershed segmentation algorithm in this case). The steps involved crowd flow segmentation are summarized in the block diagram in Fig. 9.3.

Assumptions: Motion of crowds can exhibit a wide range of behaviors and can be captured using a variety of camera setups (e.g. pole mounted or a ground-based camera). Therefore, it is pertinent to layout the assumptions and constraints on the

type of motion and scenes that an processed using the proposed algorithm. Some of these are listed next:

- Crowded scene is viewed from a distance by a camera installed over a tall structure. This constraint results from the abstraction of crowd (or people) as particles. If a scene is viewed from a closer distance, then the algorithm requires a top-down view where only heads of individuals are visible, thereby minimizing artifacts resulting from in-dependent movement of other body parts. Side views of the scene are least preferable within the particle based framework.
- The density of the crowd varies from 3 person per meter square to 7 meter per second square.
- The crowd is formally structured and focused on some collective activity. This constraint results from the fact that LCS detection algorithm exploits in some sense the 'common fate' principle (i.e. trajectories belonging to the same group have the same destination) to localize boundaries between trajectories moving in different directions. If the crowd motion is random or haphazard this may no longer be true.
- Each spatial location in the scene supports one dominant motion. That is, for a any fixed spatial location the distribution of direction and speed of optical flow vectors cannot be multi-modal. This is necessary as algorithm assumes analysis is done only at one time scale and during that time only one type of dominant motion is expected at a location.
- It should be noted that crowd behavior is dynamic in nature and can change drastically. Therefore, in order to perform any video based analysis of crowd motion a sliding window based approach should be adopted. The temporal extent of the window can be kept constant or can be dynamically adopted based on level of activity in the scene. Approaches summarized in this chapter adhere to this principal and performs flow segmentation within of a sliding temporal window.

Chapter Organization The remaining portion of the chapter is organized as follows: Sect. 9.2 provides a overview of the background material and formal definition of various concepts. Section 9.3 discusses the crowd segmentation algorithms and walks the reader through various intermediate steps. Section 9.4 describes experimental setup and presents qualitative results.

9.2 Background, Definitions and Notations

Key background concepts, mathematical notations, and formal definitions are provided in this section. The nomenclature of Shadden et al. [18] is used for this purpose.

Let a compact set $D \subset \mathbb{R}^2$ be the domain of the phase space under study. This domain corresponds to the 2D-spatial extent of the video depicting crowd motion.

Next, define a time-dependent optical flow field $\mathbf{v}(\mathbf{x},t)$ on D that satisfies C^0 and C^2 continuity in time and space, respectively. The C^0 and C^2 assumptions are required to keep the optical flow field smooth. Here, t corresponds to the t-th frame of the video. Then a particle trajectory $\mathbf{x}(t : t_0, \mathbf{x}_0)$, starting at point $\mathbf{x}_0$ at time t_0 can be defined as a solution of

$$\dot{\mathbf{x}}(t; t_0, \mathbf{x}_0) = \mathbf{v}(\mathbf{x}(t; t_0, \mathbf{x}_0), t), \tag{9.1}$$

$$\mathbf{x}(t_0; t_0, \mathbf{x}_0) = \mathbf{x}_0, \tag{9.2}$$

where $\dot{\mathbf{x}}$ is the time derivative. It can also be observed that a trajectory, $\mathbf{x}(t : t_0, \mathbf{x}_0)$, of a particle depends on the initial position $\mathbf{x}_0$ and the initial time t_0. From the above mentioned continuity constraints of optical flow, $\mathbf{v}(\mathbf{x},t)$, it follows that the particle trajectory, $\mathbf{x}(t : t_0, \mathbf{x}_0)$, will be C^1 in time and C^3 in space.

As the goal is to analyze the transport properties (using particle trajectories) of the phase space and, in turn, the underlying crowd, the solution of Eq. (9.1) can be viewed as a transport device or map that takes particles from their initial position $\mathbf{x}_0$ at time t_0 to their position at time t. Formally, this solution is referred as a "flow map," denoted by $\phi_{t_0}^t$, and that satisfies:

$$\phi_{t_0}^t : D \to D : \mathbf{x}_0 \mapsto \phi_{t_0}^t(\mathbf{x}_0) = \mathbf{x}(t; t_0, \mathbf{x}_0). \tag{9.3}$$

In addition, the flow map $\phi_{t_0}^t$ satisfies the following properties:

$$\phi_{t_0}^{t_0}(\mathbf{x}) = \mathbf{x}, \tag{9.4}$$

$$\phi_{t_0}^{t+s}(\mathbf{x}) = \phi_s^{t+s}(\phi_{t_0}^s(\mathbf{x})) = \phi_t^{t+s}(\phi_{t_0}^t(\mathbf{x})). \tag{9.5}$$

These properties follow directly from the existence and uniqueness theorem that allows one to conclude that there exists only one solution to a first-order differential equation that satisfies the given initial condition. Next we describes the key concept of FTLE field and discuss the steps involved in its computation from the flow map ϕ.

9.2.1 Finite Time Lyapunov Exponent Field

As mentioned earlier crowd segments/groupings are located using LCS, and the localization of LCS in turn requires computation of the FTLE field. The Lyapunov exponent is an asymptotic quantity that measures the extent to which an infinitely-close pair of particles separate in an infinite amount of time. In the theory of dynamical systems, it is used as a tool for measuring the chaoticity of the system under consideration by measuring the rate of exponential divergence between the neighboring trajectories in the state/phase space. Traditionally, for any given

dynamical system, $\dot{x} = f(x)$, the maximum Lyapunov characteristic exponent is defined as $\gamma = lim_{t\to\infty}\chi(t)$, with

$$\chi(t) = \frac{1}{t} ln \frac{\mid \xi(t) \mid}{\mid \xi(0) \mid}, \tag{9.6}$$

where $\xi(t)$ is the current state of the system, while $\xi(0)$ is the initial state of the given system. These states are usually obtained by solving the differential equation controlling the evolution of the system.

When the Lyapunov exponent analysis is performed over a grid of particles over finite times, it generates a FTLE field. In our formulation, the state of the system is defined as the maximum possible separation between a particle and its neighbors. Essentially, this means that the Lyapunov exponent now can be defined as a ratio of the initial separation to the maximum possible separation between the particle and its neighbors. Using this definition of the Lyapunov exponent, FTLE field $\sigma_T(\mathbf{x}_0, t_0)$ can be computed using the flow map $\phi_{t_0}^{t_0+T}$, which contains the final locations of the particles at the end of particle advection. The flow map, as mentioned earlier, quantifies the transport properties of the phase space by taking a particle from the initial position, $\mathbf{x}_0$, at time t_0 to its later position at time $t_0 + T$.

One important point to note is that the FTLE does not capture the instantaneous separation rate, but rather measures the average, or integrated, separation rate between trajectories. This distinction is important because, in time-dependent complex crowd flows, the instantaneous optical flow is not very informative. However, by accounting for the integrated effect of the crowd-flow using particle trajectories in the FTLE field, we hope to extract information that is more indicative of the actual transport behavior.

The formal derivation of the expression of FTLE proceeds as follows [7, 18]. Consider a particle $\mathbf{x} \in D$ at initial time t_0 (Fig. 9.2). Following advection, the position of the particle after a time interval T is $\mathbf{x} \mapsto \phi_T^{t_0+T}(\mathbf{x})$. Now, when advected through the flow, any arbitrary particle that is infinitesimally close to $\mathbf{x}$ at time t_0 will behave in a manner similar to $\mathbf{x}$ locally in time. However, as the advection time increases the distance between these neighboring particles will change. Now, if we represent the neighboring particle by $\mathbf{y} = \mathbf{x} + \delta\mathbf{x}(0)$ (Fig. 9.2), where $\delta\mathbf{x}(0)$ is an arbitrarily-oriented unit vector, then after a time interval T, the distance between them becomes:

$$\delta\mathbf{x}(t_0 + T) = \phi_{t_0}^{t_0+T}(\mathbf{y}) - \phi_{t_0}^{t_0+T}(\mathbf{x}) \tag{9.7}$$

$$= \frac{d\phi_{t_0}^{t_0+T}(\mathbf{x})}{d\mathbf{x}} \delta\mathbf{x}(0) + O(\|\delta\mathbf{x}(0)\|^2). \tag{9.8}$$

Since the distance $\delta\mathbf{x}(0)$ is infinitesimally small, we can drop the higher order terms in the Taylor series expansion of the flow map around the location $\mathbf{x}$. The magnitude, $\| \delta\mathbf{x}(t_0 + T) \|$, of the final separation can be computed by taking the standard L_2 norm

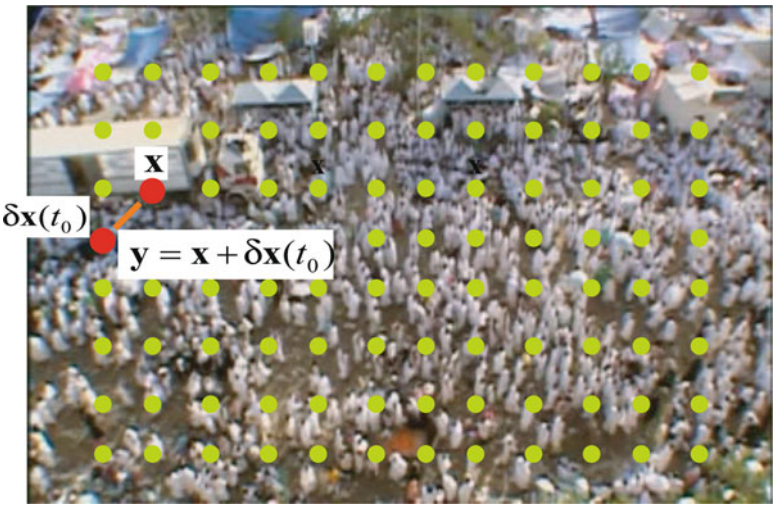

Fig. 9.2 Computation of FTLE. The initial separation between particle $\mathbf{x}$ and $\mathbf{y} = \mathbf{x} + \delta\mathbf{x}(0)$ is $\delta\mathbf{x}(0)$. In order to compute the FTLE between them, we need to find out the magnitude of the final separation after a time interval T

$$\|\delta\mathbf{x}(t_0+T)\|_2 = \left\| \frac{d\phi_{t_0}^{t_0+T}(\mathbf{x})}{d\mathbf{x}} \delta\mathbf{x}(0) \right\|_2. \tag{9.9}$$

We are interested in finding out the maximum possible separation between the particle, $\mathbf{x}$, and all its neighbors, which, in other words, means that we seek to maximize $\| \delta\mathbf{x}(t_0+T) \|_2$ over all possible choices of $\delta\mathbf{x}(0)$:

$$\| \delta\mathbf{x}(t_0+T) \|_2 = \max_{|\delta\mathbf{x}(0)|=1} \left\| \frac{d\phi_{t_0}^{t_0+T}(\mathbf{x})}{d\mathbf{x}} \delta\mathbf{x}(0) \right\|_2. \tag{9.10}$$

Using the operator norm, the above equation can be written as:

$$\| \delta\mathbf{x}(t_0+T) \|_2 = \max_{|\delta\mathbf{x}(0)|=1} \left\| \frac{d\phi_{t_0}^{t_0+T}(\mathbf{x})}{d\mathbf{x}} \delta\mathbf{x}(0) \right\|_2 = \left\| \frac{d\phi_{t_0}^{t_0+T}(\mathbf{x})}{d\mathbf{x}} \right\|_2. \tag{9.11}$$

The right-hand side of the above equation is the matrix L_2 norm that can be computed simply by using the standard property that states that, for any matrix A, the matrix L_2 norm is the square root of the maximum eigenvalue of the positive definite symmetric matrix $A^T A$. If we consider $A = \frac{d\phi_{t_0}^{t_0+T}(\mathbf{x})}{d\mathbf{x}}$, then $A^T A$ is

$$\Delta = A^T A = \frac{d\phi_{t_0}^{t_0+T}(\mathbf{x})}{d\mathbf{x}}^* \cdot \frac{d\phi_{t_0}^{t_0+T}(\mathbf{x})}{d\mathbf{x}}, \tag{9.12}$$

where superscript '*' refers to the transpose operator. It is interesting to note that Δ is also known as the finite time version of the Cauchy-Green deformation tensor. The quantity $\frac{d\phi_{t_0}^{t_0+T}(\mathbf{x})}{d\mathbf{x}}$ is the spatial gradient tensor of the flow map. The maximum eigenvalue of Δ is represented by $\lambda_{max}(\Delta)$.

Now, knowing the magnitude of the maximum possible separation, $\lambda_{max}(\Delta)$, and the initial separation, $\delta\mathbf{x}(0)$, between the particle and its neighbors, we can compute the FTLE field, σ, with a finite integration time T corresponding to point $\mathbf{x} \in D$ at time t_0 as:

$$\sigma_{t_0}^T = \frac{1}{T} \ln \sqrt{\lambda_{max}(\Delta)}. \tag{9.13}$$

Since, $\delta\mathbf{x}(0)$ is a unit vector, we eliminated it from the above equation. The above quantity is computed for each $\mathbf{x} \in D$ to obtain the entire FTLE field at time t_0.

9.2.2 Lagrangian Coherent Structures

The LCS corresponds to the boundaries between the crowd flows of distinct dynamics. They appear as ridges in the FTLE field of the video. The relationship between ridges in the FTLE field and the LCS can be explained in the following way. If two regions of a phase space have qualitatively different dynamics, then we expect a coherent motion of particles within each region, and, therefore, the eigenvalues of Δ will be close to 1, an indication that the fate of nearby particles is similar inside the region. At the boundary of the two regions, particles will move in incoherent fashion, and, therefore, will create much higher eigenvalues. These higher values will make the ridge prominent in the FTLE field and point to the locations of the LCS.

We compute two types of LCS, namely "Attracting Lagrangian Coherent Structures" (ALCS) and "Repelling Lagrangian Coherent Structures" (RLCS). The former will emphasize those boundaries between the crowds from which, in a given time interval (*forward in time*), all nearby particle trajectories separate; the later will emphasize those boundaries between the crowds from which in a given time interval (*backward in time*), all nearby particle trajectories separate. For the computation of ALCS, the particle grid is initialized at the first optical flow field and advected forward in time, followed by the computation of forward FTLE field. For the computation of RLCS, the particle grid is initialized at the last optical flow field and advected backward in time, followed by the computation of backward FTLE field.

9.3 Crowd Segmentation: The Algorithm

In this section, we bring together all the concepts explained so far and describe the algorithmic steps that involved in carrying out the segmentation of crowd into dynamically distinct groupings. A block diagram in Fig. 9.3 provides a higher-level view of the algorithmic steps.

9.3.1 Optical Flow Computation

Given a video sequence, the first task is to compute the optical flow between the consecutive frames of the video. We employ two different schemes for this purpose. The first scheme consists of a block-based correlation in the Fourier domain. The process starts by selecting a square patch centered at the same pixel location of two consecutive frames F_1 and F_2, of the given video. The pixel values in both blocks are mean normalized, and a correlation surface is constructed by performing cross correlation in the frequency domain. The peaks are located in the correlation

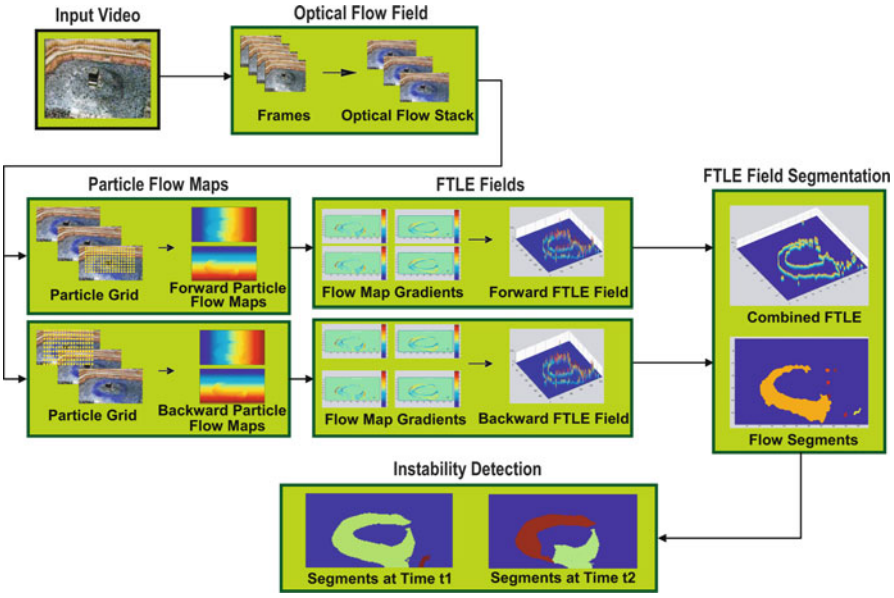

Fig. 9.3 Block diagram of the crowd-flow segmentation algorithm. (1) The input is a video of a crowded scene. (2) Computation of optical flow from the frames of the video. (3) Forward and backward advection of particle grid resulting in forward and backward particle flow maps. (4) Computation of respective FTLE fields from the forward and backward particle flow maps. (5) Fusion of forward and backward FTLE fields and label assignment using the watershed segmentation algorithm. (6) Detection of abnormal events (or crowd-flow instabilities)

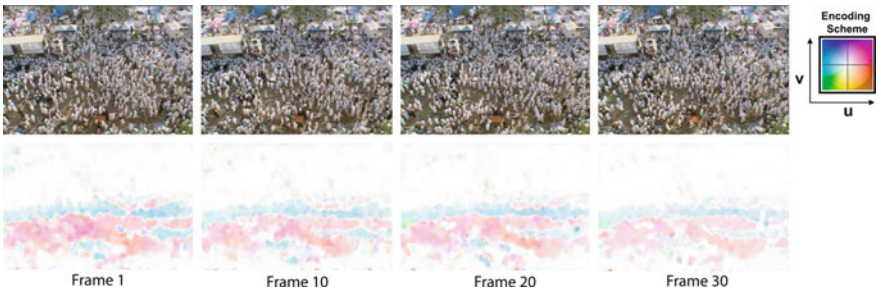

Fig. 9.4 (Color online) Examples of optical flow fields computed by using the algorithm of [1]. *Top Row*: Frames of the video. *Bottom Row*: Color-coded optical flow for the corresponding frames

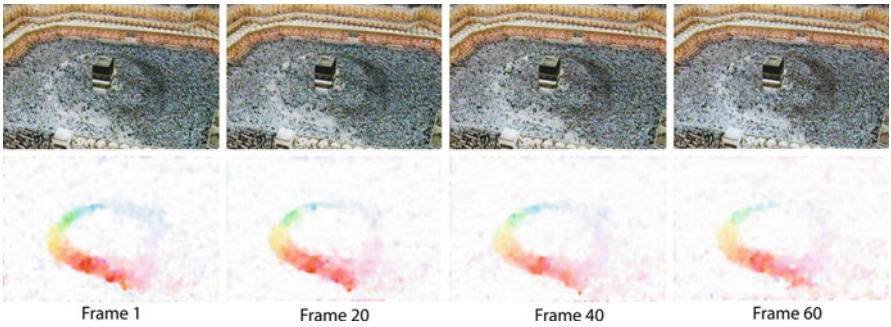

Fig. 9.5 (Color online) Examples of optical flow fields computed by using the algorithm of [1]. *Top Row*: Frames of the video. *Bottom Row*: Color-coded optical flow for the corresponding frames

surface and are used to calculate the displacement. Note that all the pixels inside a block are assigned the same displacement value. The process is repeated for all possible blocks in the given frame. Local outliers in the displacement vectors are replaced in a post-processing step, by using adaptive local median filtering. The removed vectors are filled by interpolation of the neighboring velocity vectors. A typical size of the block employed in our experiments is 16×16 pixels. The second scheme that we used is proposed in [1] where grey value constancy, gradient constancy, smoothness, and multi-scale constraints were used to estimate a high-accuracy optical flow.

To analyze the crowd-flow in a given interval of T frames, we pool the optical flow fields, $\mathbf{v}(1), \mathbf{v}(2), \ldots, \mathbf{v}(T)$, to generate a 3D volume of optical flows. To simplify the notation, we have removed the dependence of $\mathbf{v}$ on location $\mathbf{x}$. This 3D volume of optical flow is used to advect the particles, where parameter T is used as the integration time. we use the symbol B_t^{t+T} to represent a the 3D volume of optical flow fields $\mathbf{v}(t), \mathbf{v}(t+1), \ldots, \mathbf{v}(t+T)$. Figures 9.4–9.7 show color-coded optical flows computed from different sequences in our data set.

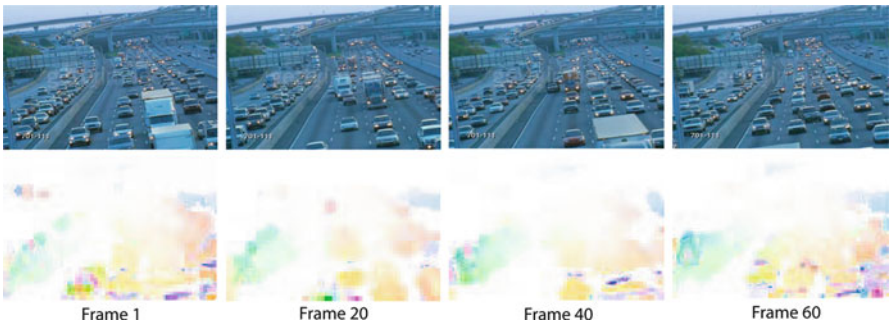

Fig. 9.6 (Color online) Examples of optical flow fields computed by using the block-based correlation algorithm. *Top Row*: Frames of the video. *Bottom Row*: Color-coded optical flow for the corresponding frames

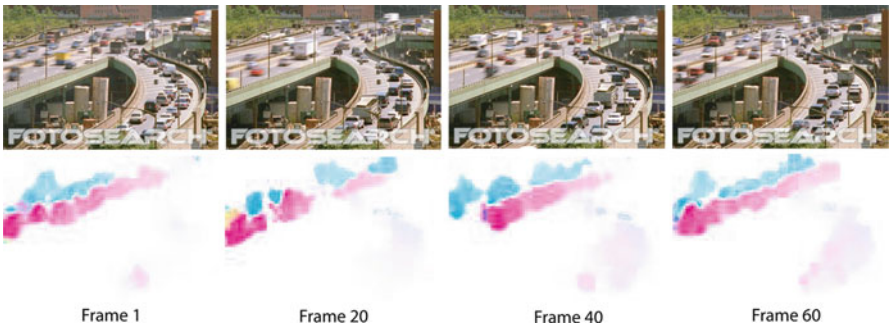

Fig. 9.7 (Color online) Examples of optical flow fields computed by using the block-based correlation algorithm. *Top Row*: Frames of the video. *Bottom Row*: Color-coded optical flow for the corresponding frames

9.3.2 Particle Advection

The next step is to advect a grid of particles through the 3D volume of flow fields, B_t^{t+T}, that corresponds to the time interval t to $t+T$. we start by launching a grid of particles over the first optical flow field, $\mathbf{v}(t)$, in B_t^{t+T}. Ideally, the resolution of the grid should be the same as the number of pixels in each frame of the video. An example of this Cartesian mesh of particles placed over the flow field of a crowd video and the trajectories of particles are provided in Fig. 9.8.

Next, the Lagrangian trajectory $[x(t+T;t,x_0,y_0),y(t+T;t,x_0,y_0)]$ corresponding to a particle at grid location (x_0,y_0) is computed by solving the ordinary differential equations numerically:

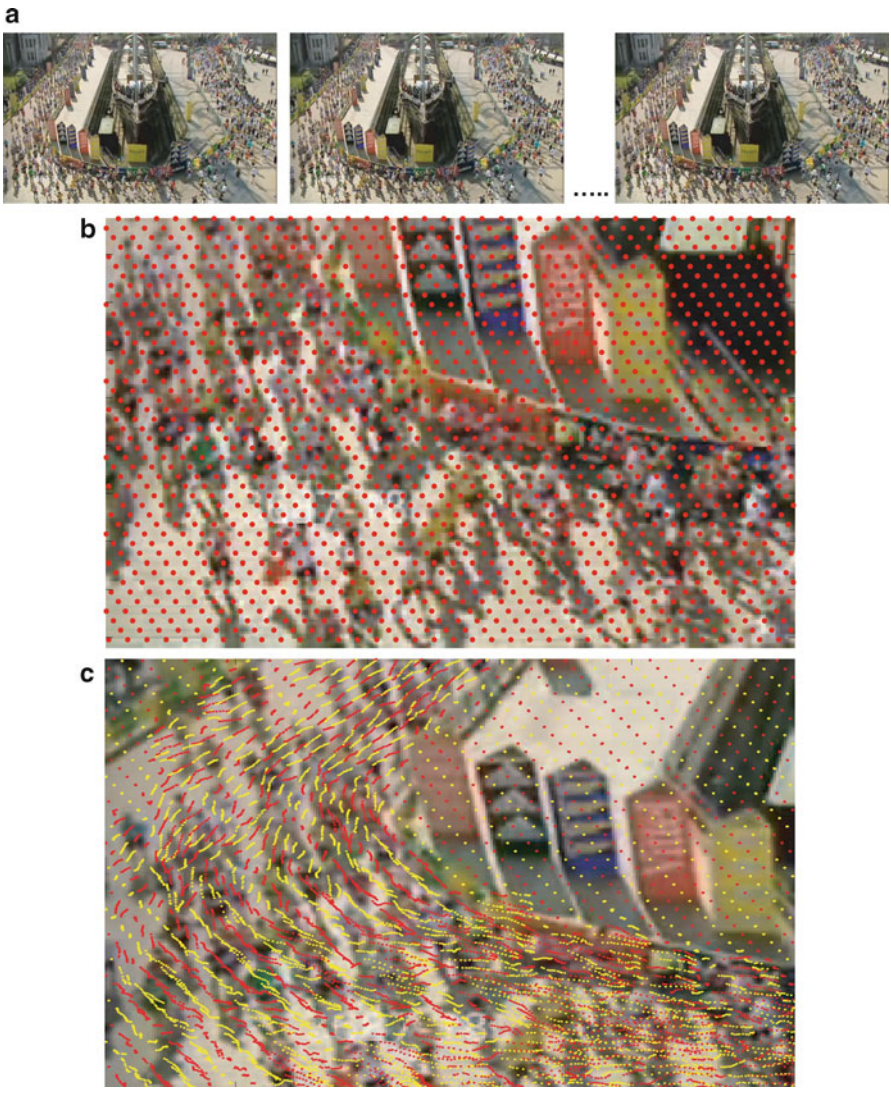

Fig. 9.8 The particle advection process. (**a**) Frames from the input video. (**b**) A grid of particles is overlaid on the flow field of the input sequence. (**c**) Trajectories of the particles are obtained by advecting them through the flow field

$$\frac{dx}{dt} = u(x,y,t), \frac{dy}{dt} = v(x,y,t), \tag{9.14}$$

subject to the initial conditions $[x(0),y(0)] = (x_0,y_0)$. $t+T$ represents the time up-till which we want to compute the trajectory. we use the fourth order Runge-Kutta-Fehlberg algorithm along with cubic interpolation [13] of the velocity field

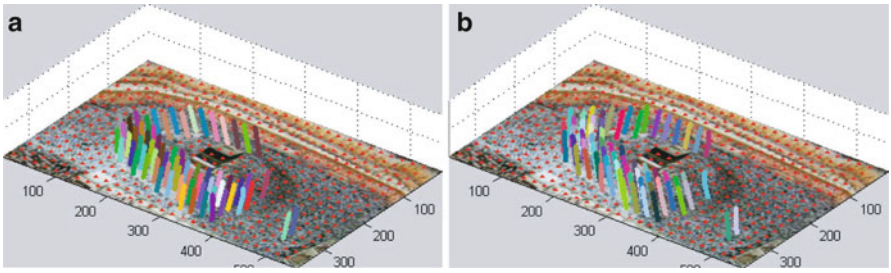

Fig. 9.9 (**a**) The Lagrangian trajectories obtained by forward integration. (**b**) The Lagrangian trajectories obtained by backward integration

to solve this system. The backward particle advection is carried out by initializing the grid of particles over the last optical flow field $\mathbf{v}(t+T)$ in the 3D volume of optical flow fields B_t^{t+T}. The direction of the optical flow vectors is reversed for the backward integration. Figure 9.9a provides a visualization of the Lagrangian trajectories obtained by forward integration, while Fig. 9.9b provides the visualization of the Lagrangian trajectories obtained by the backward integration. The length of integration, $T = 50$, was used for this purpose.

Note that, in our case the domain D is not closed and trajectories can leave the domain. The particles that leave the domain are not advected anymore, and their last available positions are kept in the flow map. That is, we do not perform any re-seeding of the particles if they leave the domain.

9.3.3 Particle Flow Maps and FTLE Field

During forward and backward integration, a separate pair of flow maps, namely ϕ_x and ϕ_y, is maintained for the grid of particles. These flow maps are used to relate the initial position of each particle to its later position obtained after the advection process. This way, the particle flow maps integrate the motion over longer durations of time, which is lacking in the instantaneous optical flow. Here, the first map, ϕ_x, keeps track of how the x coordinate of particles is changing, and, similarly, ϕ_y keeps track of the y coordinate of particles. we use notation ϕ_x^f and ϕ_y^f to refer explicitly to forward flow maps, and ϕ_x^b and ϕ_y^b to refer explicitly to backward flow maps. When the explicit references are not important, we omit the superscripts.

At the start, these maps are populated with the initial positions of the particles, which are the pixel locations at which the particle is placed. The particles are then advected under the influence of B_t^{t+T} using the method described in Sect. 9.3.2. The positions of the particles are updated until the end of the integration time length T.

The computation of the FTLE field from the particle flow maps requires computation of the spatial gradients of the particle flow maps, i.e., $\frac{d\phi_x}{dx}$, $\frac{d\phi_x}{dy}$, $\frac{d\phi_y}{dx}$, and $\frac{d\phi_y}{dy}$. This step is accomplished by using a finite differencing approach for taking

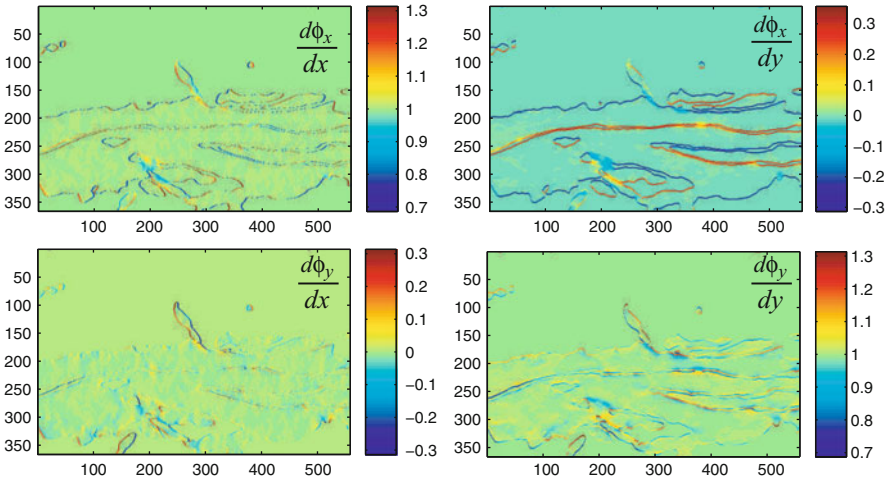

Fig. 9.10 The spatial gradients of the particle flow maps for the sequence shown in Fig. 9.4

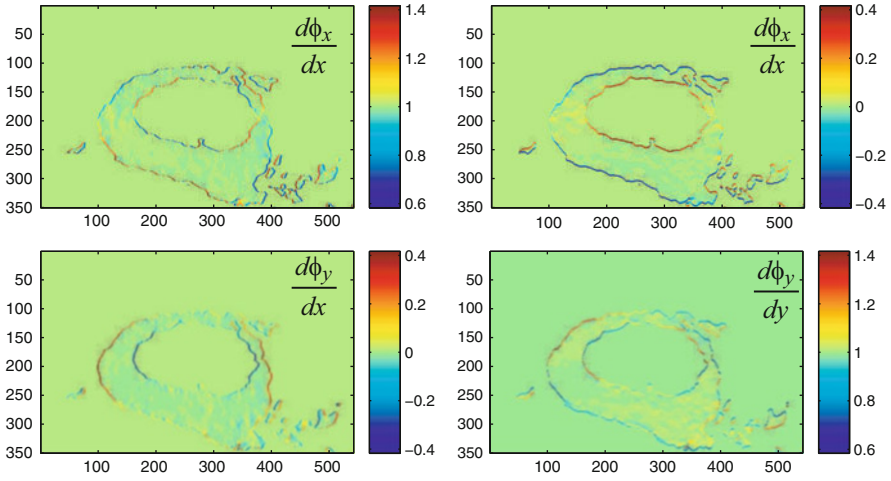

Fig. 9.11 The spatial gradients of the particle flow maps for the sequence shown in Fig. 9.5

derivatives. Figures 9.10 and 9.11 show spatial gradients of particle flow maps for two different sequences in the data set. It can be observed that a high gradient is present where the neighboring particles are behaving differently over the length of the integration. The Cauchy-Green deformation tensor is computed by substituting the spatial gradients of the particle flow maps in Eq. (9.12). Finally, the FTLE field is computed by finding the maximum eigenvalue of the Cauchy-Green deformation tensor and plugging it in Eq. (9.13). Figures 9.12–9.15 show a number of FTLE fields corresponding to different crowd sequences in our data set. In these examples, the combined FTLE field is obtained by adding the forward and backward FTLE

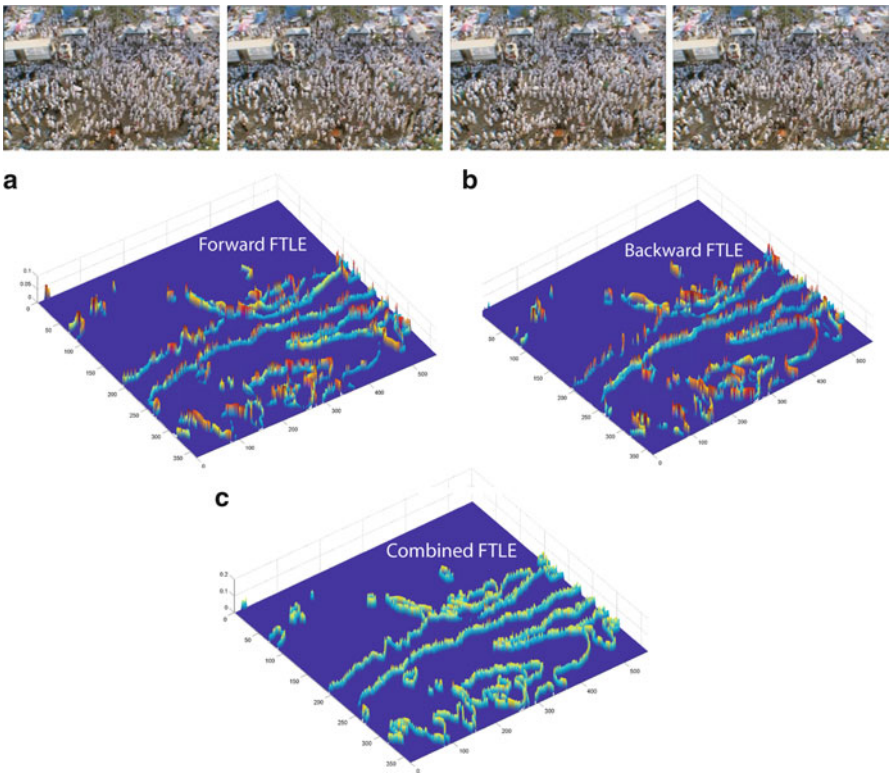

Fig. 9.12 FTLE field for the sequence shown at the *top*. The sequence has multiple groups of people intermingling with each other. The ridges are prominent at the locations where the neighboring crowd groups have dynamically distinct behavior. (**a**) The forward FTLE field obtained by the forward integration of particles. (**b**) The backward FTLE field obtained by the backward integration of particles. (**c**) The combined FTLE field

fields. It can be observed that ridges in these fields (Figs. 9.12–9.15), which point to the location of LCS, are very prominent, and, therefore, can be used to separate regions of the crowd-flow that are dynamically distinct from each other.

The utility of computing forward and backward FTLE fields becomes obvious from the analysis of the FTLE fields shown in Fig. 9.13. In this video sequence traffic from the ramp is merging onto the main highway. When the particles are advected forward in time, no LCS appear at the intersection of the ramp and the main highway (Fig. 9.13a). The reason is that the particles at the intersection move forward coherently in time as the destinations of the underlying traffic flow on the ramp and the main highway are the same. But when these particles are advected backward in time, the LCS appear at the intersection (Fig. 9.13b) since the particles at the intersection do not have the same destination backward in time because the underlying traffic is originating from different locations. In other words, by backward integration, we am able to take into account the origin of the flow

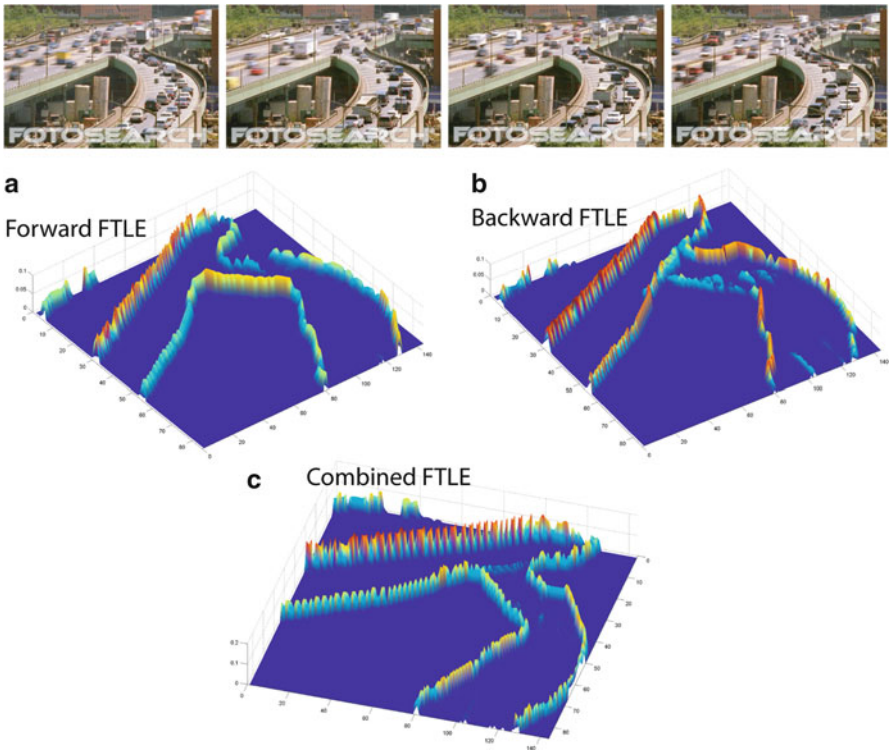

Fig. 9.13 FTLE field for the sequence shown at the *top*. The sequence has multiple lanes of traffic, and the traffic from the ramp is merging onto the main highway. (**a**) The forward FTLE field obtained by the forward integration of particles. Note that no LCS are present at the intersection of the ramp and the highway. (**b**) The backward FTLE field obtained by the backward integration of particles. Note that LCS have now appeared at the intersection of the ramp and the highway. (**c**) The combined FTLE field

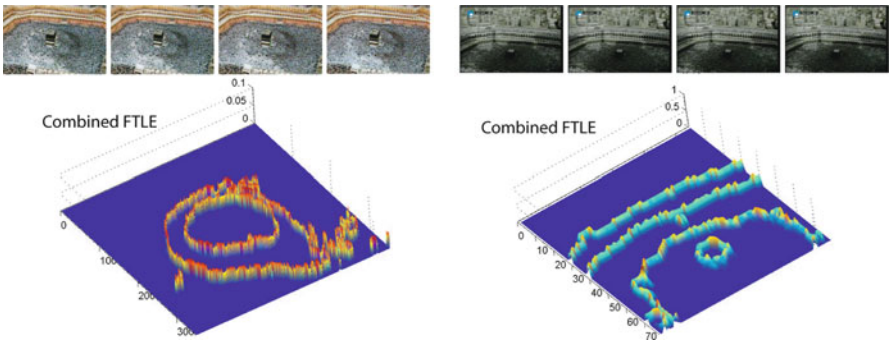

Fig. 9.14 The combined FTLE fields for the sequences shown at the *top*

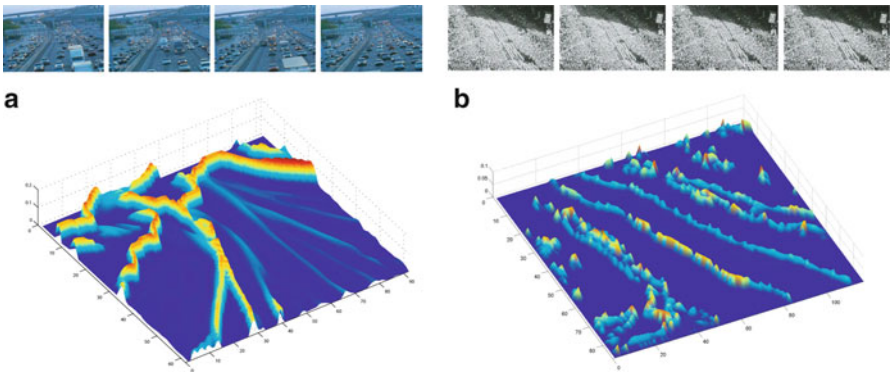

Fig. 9.15 The combined FTLE fields for the sequences shown at the *top*

in addition to its destination. This capability is important to completely resolve different crowd-flow segments present in the scene. This point will become clearer when we present the segmentation results in a later section.

9.3.4 FTLE Field Segmentation

The LCS in the FTLE field can be treated as the watershed lines dividing individual catchment basins. Each catchment basin represents the distinct crowd grouping that is present in the scene. The catchment basins are homogeneous in the sense that all the particles belonging to the same catchment basin have the same origin and destination. To generate a distinct labeling for each catchment basin, we employ the watershed segmentation algorithm [16]. The final segmentation map is created by removing those segments where the magnitude of the flow is zero. we call such segments "vacuum segments." Note that, due to the unique strength of the FTLE field based representation, we do not have to pre-specify the number of crowd-flow segments. This way, we are able to overcome the problem of specifying the number of segments or clusters which is common in most of the clustering and segmentation algorithms [19].

9.4 Experiments and Discussion

This section discusses the experimental setup and the data sets used in the experiments. It also presents the segmentation results along with a discussion of the interpretation of the results.

Fig. 9.16 Example of sequences used in our experiments

9.4.1 Datasets and Experimental Setup

We have tested our approach on videos taken from the stock footage web sites (Getty-Images [9], Photo-Search), and Video Google [10] which are now part of UCF Crowd data set [8]. Two types of crowded scenarios are covered in these videos: the first scenario consists of scenes involving the high-density crowds, while the second scenario consists of high-density traffic scenes. Traffic scenes can be treated as a close approximation of the motion of crowds of people and, therefore, provides us with useful data for testing the performance of the proposed algorithm. Another set of videos were taken from the National Geographic documentary, entitled "Inside Mecca," which covers the yearly ritual of Hajj performed by close to two million people. Therefore, this event provides a unique opportunity for capturing data about the behavior of large gatherings of people in a realistic setting. Figure 9.16 shows key frames from some of these sequences.

For each video, the optical flow is computed by using the algorithms previously described in Sect. 9.3.1. The computation of the optical flow is performed at a coarser resolution than the resolution of the image to reduce the computational cost. Next, a grid of particles is placed over the flow field. The resolution of the grid is kept the same as the number of pixels on which the flow field is computed. The forward and backward particle flow maps are generated using the advection algorithm described in Sect. 9.3.2. The corresponding FTLE fields are computed from the spatial gradient tensor of the flow maps using Eq. (9.13). The backward

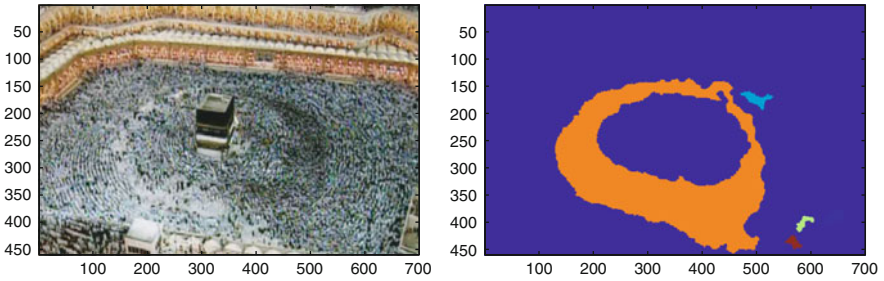

Fig. 9.17 The flow segmentation result on a video taken from the National Geographic documentary "Inside Mecca." *Left*: A frame from the video. *Right*: The crowd-flow segmentation mask

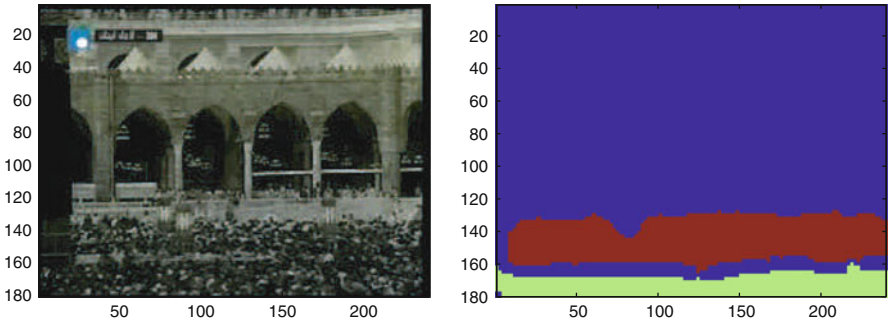

Fig. 9.18 The flow segmentation result on a video from "Video Google." *Left*: A frame from the video. *Right*: The crowd-flow segmentation mask

and forward FTLE fields are fused to generate a combined FTLE field. The fusion is carried out by adding the values of both fields. Finally, the segmentation is performed using the watershed segmentation algorithm.

9.4.2 Segmentation Results

This section presents qualitative analysis of the results obtained on different video sequences. Figures 9.17–9.25 show the segmentation results on all the sequences in the data set.

The first sequence, shown in Fig. 9.17, are extracted from the National Geographic documentary entitled "Inside Mecca". The sequence depicts thousands of people circling the Kabba in a counter-clockwise direction. In this case, the group of people circling in the center is part of the same flow segment because of its common dynamics and desirable goal. The optical flow field of the crowd motion offers a unique challenge as one can observe from the color-coded optical flow shown in Fig. 9.4. The different colors emphasize that the flow vectors along the circular path

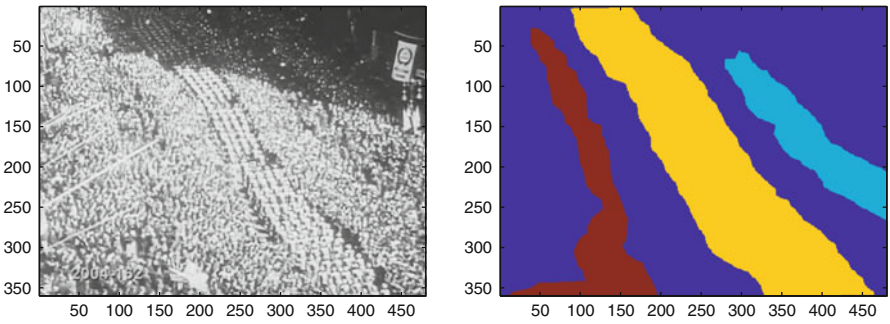

Fig. 9.19 The flow segmentation result on a video taken from the stock footage web site "Getty Images." *Left*: A frame from the video. *Right*: The crowd-flow segmentation mask

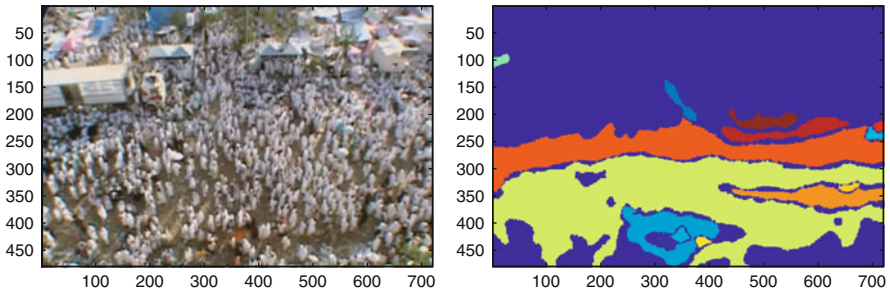

Fig. 9.20 The flow segmentation result on a video taken from the National Geographic documentary "Inside Mecca." *Left*: A frame from the video. *Right*: The crowd-flow segmentation mask

have different directions and magnitudes. This means that a simple clustering of these vectors will not allow us to assign these vectors to the same cluster when, in fact, they all belong to one cluster. The result is shown in Fig. 9.26a, where mean-shift clustering was used to cluster the optical flow vectors extracted from the instantaneous optical flow field. The clustering results are shown for different choices of the band-width parameter. But even with different values of bandwidth, the mean-shift is not able to correctly localize the circular segment. However, using our method where we integrate the motion information over longer durations of time, we are able to correctly segment the complex crowd motions (Fig. 9.17). The LCS structures previously shown in Fig. 9.14a, show that the dynamic behavior of the crowd moving in a circle is preserved by emphasizing the boundaries of the coherent flow regions. Another result of a similar type of motion is presented in Fig. 9.21. In this case, there is an additional group of people that is walking on top of the roof. Our method is able to localize this additional crowd-flow segment as well.

The next result that we would like to discuss is shown in Fig. 9.20. This sequence contains complex motion dynamics as there are several groups of people that are intermingling with each other and moving in various directions. The challenges

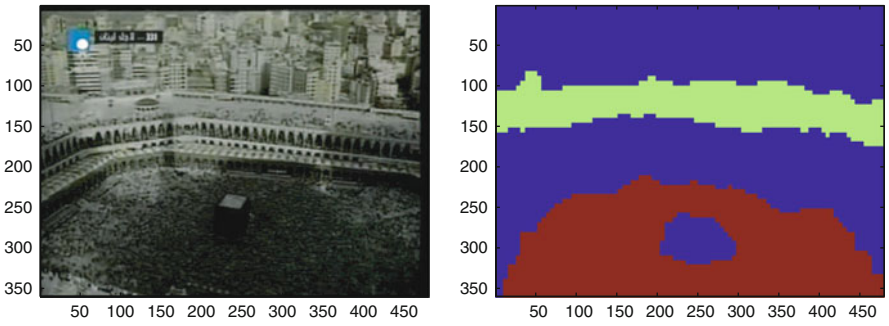

Fig. 9.21 The flow segmentation result on a video from "Video Google." *Left*: A frame from the video. *Right*: The crowd-flow segmentation mask

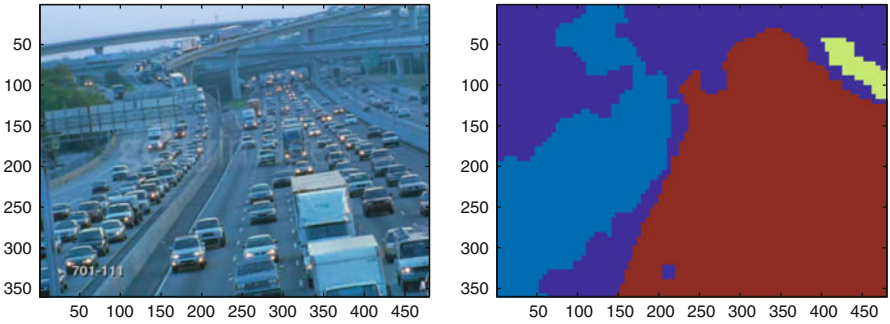

Fig. 9.22 The result of the flow segmentation on a high-density traffic scene. This segmentation was obtained by using both the forward and backward FTLE fields. *Left*: A frame from the video. *Right*: The crowd-flow segmentation mask

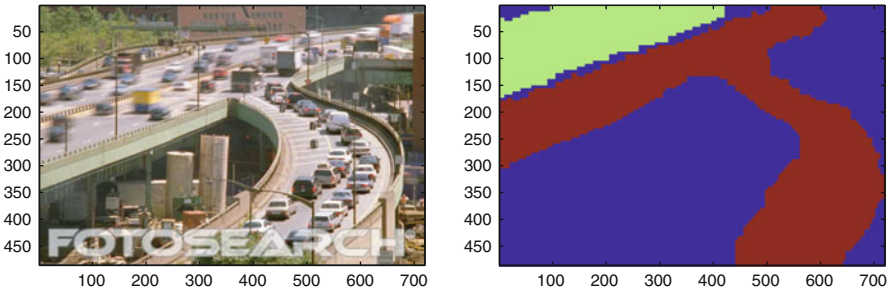

Fig. 9.23 Result of the flow segmentation on a high-density traffic scene. The segments correspond to group of cars that are behaving dynamically different from each other

posed by this sequence are different in that the mixing barriers between various crowd groupings must be correctly located. The segmentation result shown in Fig. 9.20 demonstrate that we am able to localize most of the distinct crowd groupings that were present in the scene. The discovered barriers between the crowd

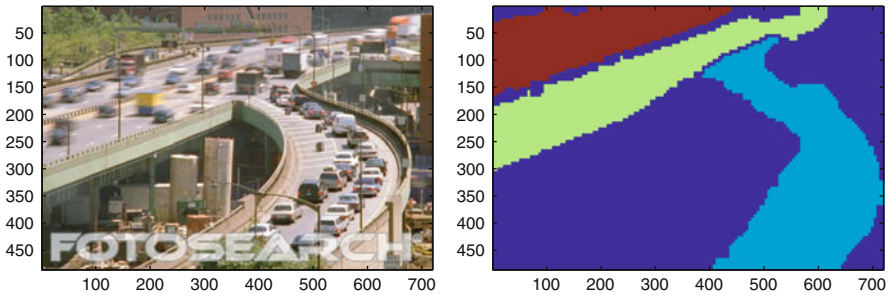

Fig. 9.24 The result of the flow segmentation on a high-density traffic scene. This segmentation was obtained by using only the forward FTLE field. *Left*: A frame from the video. *Right*: The crowd-flow segmentation mask

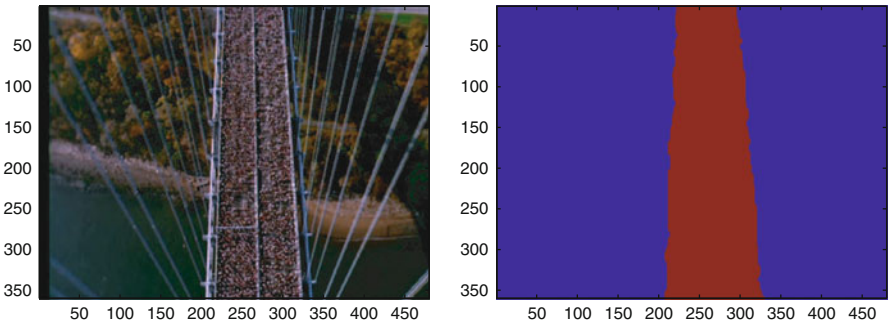

Fig. 9.25 The result of the crowd-flow segmentation on a marathon sequence. *Left*: A frame from the video. *Right*: The crowd-flow segmentation mask

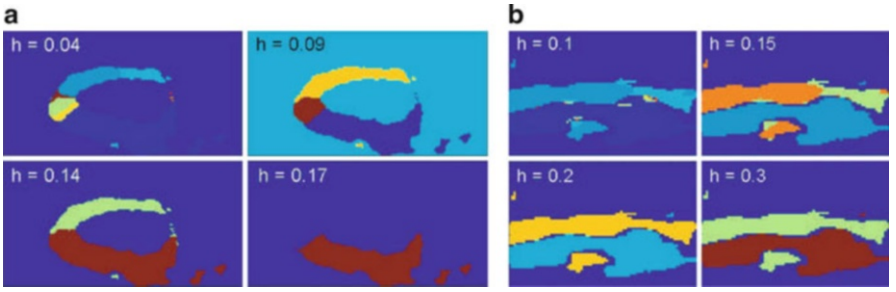

Fig. 9.26 A comparison with respect to the mean shift segmentation. (**a**) The segmentation obtained for the sequence shown in Fig. 9.17. (**b**) The segmentation obtained for the sequence shown in Fig. 9.20

groupings can be observed in the combined FTLE field shown in Fig. 9.12c. The barriers which appear in the form of ridges in the FTLE field, encapsulate each crowd group. A comparison is again performed with the mean-shift clustering approach (Fig. 9.26b), but, again, the mean shift is not able to localize all the crowd-flow

segments. This again points to the utility of integrating motion information over longer periods of time, which helps to get a better picture of the crowd motion. Some other example results on sequences involving groups of people are presented in Figs. 9.18, 9.19, and 9.25.

Next, we discuss segmentation results on a high-density traffic sequence (Fig. 9.23). The results on this sequence highlight the utility of using both forward and backward integration of particles through the 3D volume of optical flows. In this sequence, vehicles are moving in two opposite directions on the main highway, while a flow of traffic is merging onto the main highway from the ramp. The challenge in this sequence is to find the right membership of the flow generated by the traffic on the ramp by resolving its origin and destination. If we only use the forward integration, it is obvious that all the particles initialized over the ramp will have the same fate as the particles on the main highway. This means that the traffic on the ramp will become part of the flow generated by the lane on the right-hand side of the highway. Another way to look at the forward integration is from the viewpoint of flow continuity, where out-going flux on the ramp is equal to the additional flux received by the highway at this location. The segmentation result shown in Fig. 9.23 validates the above observation where same labeling is being assigned to the ramp and to the right lane of the main highway. This ambiguity can be resolved by the addition of the backward integration of particles. Since they are considered backwards in time, the particles on the two sections of the road do not share the same origin or, in other words, the outgoing flux is not equal to the flux received by the two sections of the road. The segmentation result shown in Fig. 9.24 demonstrates that by using both forward and backward integration of particles, a flow segmentation that is more refined is obtained. The result on another traffic sequence is shown in Fig. 9.22.

9.5 Summary

This chapter has developed an algorithm for segmenting scenes of crowds of people into ‘crowd groupings’ that are dynamically distinct. For this purpose, the spatial extent of the video is treated as a phase space of a non-autonomous dynamical system in which transport from one region of the phase space to the other is controlled by the optical flow. Next, a grid of particles is advected forward and backward in time through this phase space and the amount by which the neighboring particles diverged is quantified by using a Cauchy-Green deformation tensor. The maximum eigenvalue of this tensor is used to construct a Finite-Time Lyapunov Exponent (FTLE) field, which revealed the time-dependent invariant manifolds of the phase space called Lagrangian Coherent Structures (LCS). The LCS in turn divided the crowd-flow into regions of different dynamics.

The strength of this approach lies in the fact that it bypasses the need for low-level detection of individual objects altogether, which will be impossible in a high-density crowded scene, and generates a concise representation of the complex mechanics of human crowds using only the global analysis.

References

1. Brox, T., Bruhn, A., Papenberg, N., Weickert, J.: High accuracy optical flow estimation based on a theory for warping. In: ECCV, Prague (2004)
2. Cartwright, D., Zander, A.: Group Dynamics: Research and Theory, 3rd edn. Tavistock Publications, London (1968)
3. Chrisohoides, A., et al.: Experimental visualization of Lagrangian coherent structures in aperiodic flows. Phys. Fluids **15**, L25 (2003)
4. Haller, G.: Distinguished material surfaces and coherent strcutures in three-dimensional fluid flows. Phys. D **149**(4), 248–277 (2001)
5. Haller, G.: Finding finite-time invariant manifolds in two dimensional velocity data. Chaos **10**(1), 99–108 (2000)
6. Haller, G.: Lagrangian strcutures and the rate of strain in partition of two dimensional turbulence. Phys. Fluids **13**(11), p. 3365 (2001)
7. Haller, G., Yuan, G.: Lagrangian coherent structures and mixing in two dimensional turbulence. Physica D **147**(3–4), 352–370 (2000)
8. http://crcv.ucf.edu/data/crowd.php
9. http://www.gettyimages.com
10. http://www.video.google.com
11. Lapeyre, G., et al.: Characterization of finite-time Lyapunov exponents and vectors in two-dimensional turbeulence. Chaos **12**(3), 688–698 (2002)
12. Lekien, F., et al.: Dynamically consistent Lagrangian coherent structures. Am. Inst. Phys.: 8th Exp. Chaos Conf. D **742**, 132–139 (2004)
13. Lekien, F., et al.: Tricubic interpolation in three dimensions. J. Numer. Methods Eng. **63**(3), 455–471 (2005)
14. Lewin, K.: In: Cartwright, D. (ed.) Field Theory in Social Science; Selected Theoretical Papers. Harper & Row, New York (1951)
15. Malhotra, N., et al.: Geometric structures, lobe dynamics and Lagrangian transport in flows with a periodic time dependence, with applications to Rossby wave flow. J. Non-Linear Sci. **8**, 401 (1998)
16. Meyer, F.: Topographic distance and watershed lines. Signal Process. **38**, 113–125 (1994)
17. Poje, A.C., Haller, G.: The geometry and statistics of mixing in aperiodic flows. Phys. Fluids **11**(10), 2963–2968 (1999)
18. Shadden, S.C., et al.: Definition and properties of Lagrangian coherent structures from finite-time Lyapunov exponents in two-dimensional aperiodic flows. Physica D, Nonlinear Phenomena, **212**(3–4), 271–304 (2005)
19. Shi, J., Malik, J.: Normalized cuts and image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. **22**(8), 888–905 (2000)