Chaotic Invariants of Lagrangian Particle Trajectories for Anomaly Detection in Crowded Scenes

Shandong Wu
Computer Vision Lab
University of Central Florida
sdwu@eecs.ucf.edu

Brian E. Moore
Department of Mathematics
University of Central Florida
bmoore@math.ucf.edu

Mubarak Shah
Computer Vision Lab
University of Central Florida
shah@eecs.ucf.edu

Abstract

A novel method for crowd flow modeling and anomaly detection is proposed for both coherent and incoherent scenes. The novelty is revealed in three aspects. First, it is a unique utilization of particle trajectories for modeling crowded scenes, in which we propose new and efficient representative trajectories for modeling arbitrarily complicated crowd flows. Second, chaotic dynamics are introduced into the crowd context to characterize complicated crowd motions by regulating a set of chaotic invariant features, which are reliably computed and used for detecting anomalies. Third, a probabilistic framework for anomaly detection and localization is formulated.

The overall work-flow begins with particle advection based on optical flow. Then particle trajectories are clustered to obtain representative trajectories for a crowd flow. Next, the chaotic dynamics of all representative trajectories are extracted and quantified using chaotic invariants known as maximal Lyapunov exponent and correlation dimension. Probabilistic model is learned from these chaotic feature set, and finally, a maximum likelihood estimation criterion is adopted to identify a query video of a scene as normal or abnormal. Furthermore, an effective anomaly localization algorithm is designed to locate the position and size of an anomaly. Experiments are conducted on known crowd data set, and results show that our method achieves higher accuracy in anomaly detection and can effectively localize anomalies.

1. Introduction and Related Work

Crowd scene analysis has been attracting increasing attention in the computer vision community. High object densities provide the main source for challenges in this field of research. As seen in Figure 1, a crowd scene may contain hundreds or even thousands of objects. Therefore, conventional techniques [1, 2] aimed at non-crowded scenes involving object detection, segmentation, tracking, etc., suffer from the problems due to severe inter-object occlusion, small object size, similar appearance, etc. Although some traditional methods seem to work for low density crowded scenes [3], they are likely to fail in handling extremely dense crowd motions. To overcome or avoid these difficulties, researchers are developing new regimes for crowded scene analysis by making use of some crowd-specific observations. For example, the objects in a coherent crowd may move with common dynamics and those dynamics can be modeled to characterize the crowd flow. Yet, many scenes contain incoherent or completely random motion of objects, which complicates the analysis of crowds.

Figure 1. Crowd scenarios with different levels of coherency.

Anomaly detection plays an important role in many applications such as crowd surveillance, public place monitoring, security control, etc. Throughout the literature, an anomaly is treated as a context-sensitive term, which is usually studied according to the scenarios of interest [4, 5]. While an anomaly can be defined in different ways, it is found that an appropriate modeling of scenes can be an important basis for supporting effective detection of anomalies [6, 8]. Limiting our review of the literature to work directly related to analysis of crowded scenes, common features used for modeling crowded motions include optical flow [9, 10, 11, 12], gradient [13], spatiotemporal volume [14], etc. One notable point is that merely relying on instantaneous motion features makes it difficult to capture certain long-term motion properties.

Several methods have been reported for anomaly detec-

tion in crowded scenes. Ali and Shah [10] used Lagrangian Coherent Structures (LCS) to segment coherent crowd flows for video segmentation. This method is constrained due to its incapability to segment incoherent motions. Social force model was studied in [9] to explore the human interactions within crowds to detect abnormal events. It is observed that the use of bag-of-words method may have difficulty in localizing anomalies. Kratz and Nishino [13] used hidden Markov model (HMM) to detect abnormal events in crowded scenes, but the cuboid based windowing strategy tends to result in loss of useful information, because the coherently meaningful features may be separated in different cuboids. HMMs were also used in [11] where principal component analysis was employed to build feature prototypes for crowd event detection.

In addition, chaos has been extensively studied in the physics community [15] where the typical research subject is a univariate time series. Ali et al. [7] introduced chaotic embedding to solve action recognition problems. The tracked motion trajectories of individual humans are treated as a time series to reconstruct the chaotic attractor, then several related features are extracted to describe the motions for action recognition.

In this paper, we propose a method for anomaly detection in croweded scenes based on Lagrangian particle dynamics and chaotic invariants, which outperforms the state of the art and is able to handle both coherent and incoherent scenes. Our main contribution is the automatic production of representative trajectories of a crowd, for which chaotic invariants can be computed, whereas previous work assumes the time trajectory data is already available [7]. Figure 2 shows the overall pipeline of our method.

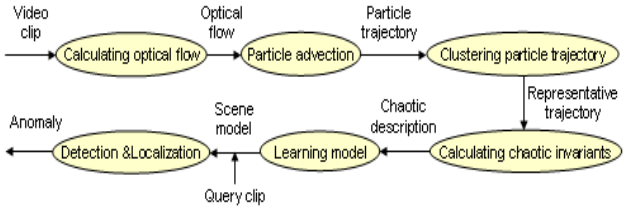


Figure 2. Framework for anomaly detection and localization.

2. Crowd Flow Representation

2.1. Particle Advection

A given video of a crowd scene is first divided into a series of clips with respect to time. Each clip is represented by a matrix of size $T \times W \times H$, where T is the number of Frames and $W \times H$ denotes the frame resolution (width by height). Then the optical flow of each clip is computed so that we obtain the local motion velocity set denoted by

$$\{(U_w^t, V_h^t)|w \in [1, W], h \in [1, H], t \in [1, T-1]\}. \quad (1)$$

Next, following [10], we perform the particle advection to estimate the positions of moving particles employing sub-pixel level optical flow interpolation. Numerically solving the evolution equations using

$$X_w^{t+1} = X_w^t + U_w^t, \quad (2)$$
$$Y_h^{t+1} = Y_h^t + V_h^t, \quad (3)$$

we can obtain the position vector (X_w^t, Y_h^t), of a particle at (w, h) and time t. Thus, this particles trajectory is denoted by $\{(X_w^t, Y_h^t)|t \in [1, T]\}$ and all of the particle trajectories in the clip are denoted

$$\{(X_w^t, Y_h^t)|w \in [1, W], h \in [1, H], t \in [1, T]\}. \quad (4)$$

The system is re-initialized for each clip, enabling the advection process to capture possible updates to scene properties at later times.

2.2. Representative Trajectories

In a crowded scene, overlaid with particles that move with the flow, many trajectories may belong to a single sub-object, and particle trajectories distributed in one or more such sub-object region(s) may be merged to give a representative trajectory that describes the motion therein. Figure 3 shows the particle trajectories for the scenes shown in Figure 1, which are essentially describing the flow in terms of the movement of particles. Since each particle overlays a single pixel, there is a huge number of particle trajectories. In order to achieve a more efficient and manageable scene representation, we cluster trajectories such that each trajectory represents one cluster.

A trajectory cluster is allowed to have flexible size. Comparatively, while a representative trajectory may be equivalent to the corresponding region of a human head, hand, or foot for instance, it may also possibly correspond to an area that involves a whole person, or multiple people,

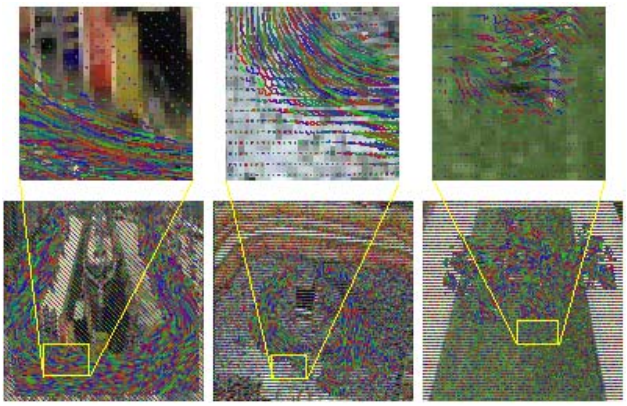


Figure 3. Particle trajectories overlayed on three crowd scenes. Top row shows zoom-in view of parts of each scene.

where the constitutive particles experience close-grained motion. Hence, a representative trajectory in crowds is inherently different from the traditionally tracked trajectory of an object in sparse scenes, where normally a body part or a person as a whole [16, 17] are tracked.

We generate representative trajectories in two steps. First, we remove relatively motionless particles and trajectories that carry minor information due to slight camera motion or noise. For a given trajectory set (4), we discard those with variance smaller than a predefined threshold,

$$\text{var}\left\{(X_w^t, Y_h^t)\right\} < \epsilon. \tag{5}$$

Figure 4 demonstrates the result after this step, showing that the overall motion in the scene is roughly revealed.

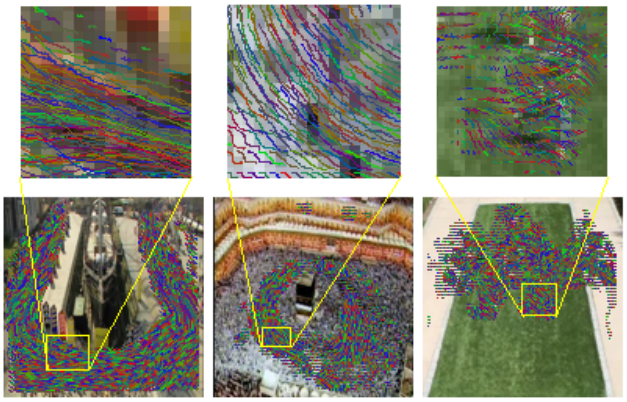


Figure 4. Trajectories after low variance particles are removed. Top row shows zoom-in view of parts of each scene.

Second, we cluster the remaining particle trajectories iteratively by k-means algorithm according to position information. Note that we choose an iterative clustering strategy to determine an optimal number of clusters. Starting with a relatively large number of clusters, we measure the inter-cluster Jensen-Shannon divergence among resulting clusters. The two clusters, with a Jensen-Shannon distance falling below a predefined threshold, are merged, and this process is iterated until the threshold is reached. This way a number of clusters are generated, and the cluster centers are extracted as the representative trajectories. Experiments show that the anomaly detection is basically insensitive to the number of clusters.

Figure 5 demonstrates an intermediate result of the second step, where 202 clusters are generated and each cluster is shown by a rectangle in random colors. The zoomed-in plots show particle trajectories with thin lines, while the thick black lines are representative trajectories of the corresponding clusters. Figure 6 shows the representative trajectories after clustering, giving a compact yet informative representation of a crowd flow.

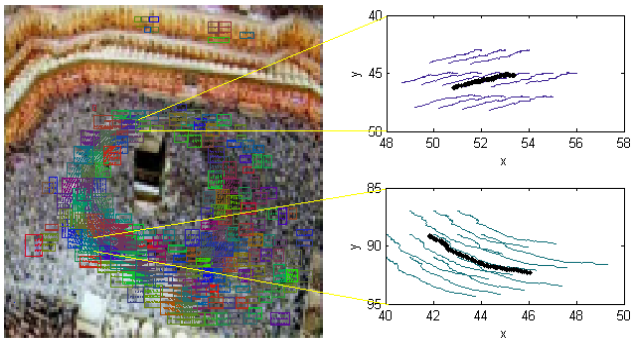


Figure 5. Trajectories clustered according to position information, (left) and representative trajectories for two clusters (right).

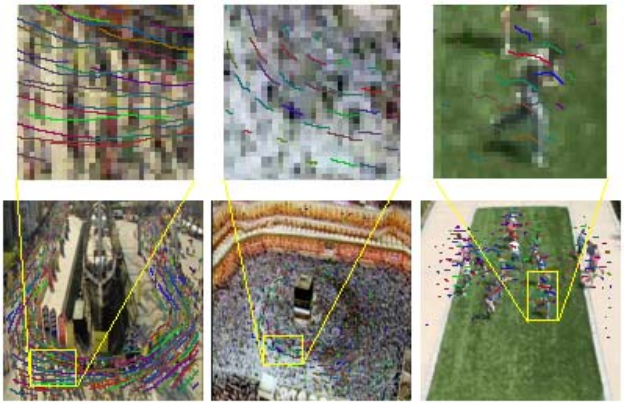


Figure 6. Representative trajectories for three scenes. Top row shows zoom-in view of parts of each scene.

3. Chaotic Invariants

Using representative trajectories, we proceed toward a comprehensive model of the scene through chaotic dynamics, which are capable of handling both coherent and incoherent scenes and offer a description using only two features. A representative trajectory essentially admits the data of two time series (x and y). Based on the Taken embedding principle (see Section 3 in [7] for preliminaries of chaotic embedding), we are able to reconstruct a strange attractor in phase space that has equivalent dynamics to the system revealed by time series. The attractor is made up of orbits and may be quantified by a set of representative chaotic features.

We restrict our focus to two chaotic invariants, known as the largest Lyapunov exponent L and the correlation dimension D, which may be used to characterize a chaotic system. These invariants are the same for two representative trajectories with the same chaotic dynamics independent of their position and magnitude. Nevertheless, positions of representative trajectories are important for detecting and locating anomalies in a consistent flow, and we include the mean M of the x and y coordinates for a representative trajectory

in our feature set $F = \{L, D, M\}$. This is similar to the work of [7], where chaotic invariants are used for human action recognition, but a key difference is that we do not use the correlation sum, since we find that F is sufficient for anomaly detection and localization in crowded scenes.

For a representative trajectory, we calculate chaotic features for two scalar time series in x and y, and then concatenate them. Taking X^t as an example, we use time delay embedding to reconstruct the corresponding orbit in phase space, denoted $\tilde{X}_t = [X^t, X^{t+J}, X^{t+2J}, \ldots, X^{t+(m-1)J}]$, for all $t \in [1, T]$, where J is the time delay, m is the embedding dimension. Then, for a particular reference point $\tilde{X}_j$, we locate its nearest neighbor $\tilde{X}_{\hat{j}}$ by searching for the point that minimizes the Euclidean distance, i.e. $\hat{j} = \text{argmin}\|\tilde{X}_j - \tilde{X}_{\hat{j}}\|$. In addition, we require that nearest neighbors have a temporal separation greater than the mean period (denoted p) of the time series, such that $|\hat{j} - j| > p$. (This is another key difference from [7], which allows nearest neighbors from any point in time.) In this way, we consider each pair of neighbors as nearby initial conditions for different orbits. The largest Lyapunov exponent can then be estimated as the mean rate of separation of the nearest neighbors. The correlation dimension is estimated using the number of nearest neighbors, as higher dimensions allow greater numbers of nearby pairs.

3.1. Largest Lyapunov Exponent

The largest Lyapunov exponent provides quantitative information about orbits that start close together but diverge over time. Under the assumption that the underlying systems of representative trajectories are chaotic, we require that L must satisfy conditions for chaos, i.e. $L > 0$. To this end, we use a method proposed in [18] for estimating L from small and noisy data sets, appropriate for our problem, and the method also ensures that $L > 0$.

According to the definition of L, assuming $d(t)$ is the average divergence at time t and k is a constant that normalizes the initial separation, we have $d(t) = ke^{Lt}$. Then, the jth pair of nearest neighbors can be assumed to diverge approximately at a rate L. We therefore derive

$$d_j(t_i) \approx k_j e^{Lt_i}, \tag{6}$$

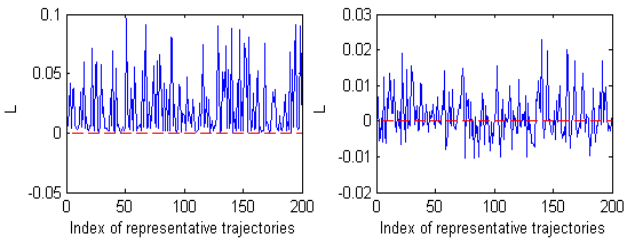


Figure 7. Largest Lyapunov exponents for representative trajectories using our method (left) and the method of [7] (right).

where $t_i = i\Delta t$ and Δt is the sampling period of the time series, k_j is the initial separation, and $d_j(t_i)$ denotes the distance between the jth pair of nearest neighbors after i discrete time steps. Equation (6) can be reformulated as

$$\ln d_j(t_i) \approx \ln k_j + Lt_i. \tag{7}$$

This equation represents a set of approximately parallel lines each with a slope roughly proportional to L. Now, the largest Lyapunov exponent can be accurately approximated by fitting the average line defined by

$$y(t_i) = \frac{\langle \ln d_j(t_i) \rangle}{\Delta t}, \tag{8}$$

where $\langle \cdot \rangle$ denotes the average over all j. Averaging is crucial for computing accurate values of L for a small and noisy data set. This method for computing L has proven to be insensitive to the changes in time delay, embedding dimension, size of data set, and to some extent noise [18]. Figure 7 shows our estimated largest Lyapunov exponents for the x time series of the representative trajectories of a clip of the scene shown in Figure 6 (right) compared to the method of [7]. Note that the exponents computed in our results, are all positive, just as one should expect from a chaotic system.

3.2. Correlation Dimension

The correlation dimension measures the size of an attractor, which defines the chaotic dynamics, and it is most easily estimated using the correlation sum

$$C(\delta) = \frac{2}{Q(Q-1)} \sum_{i \neq j} H(\delta - \|\tilde{X}_i - \tilde{X}_j\|). \tag{9}$$

H denotes the Heaviside step function, δ is a threshold distance, Q is the number of points in the time series. Then the correlation dimension D is approximated using $C(\delta) \approx \delta^D$.

4. Anomaly Detection and Localization

The feature set F_n for $n \in [1, N]$, with N the number of representative trajectories, is used to characterize a crowd flow of a video clip and to detect the change of the dynamics in a flow. Formally, the anomaly is defined as any spatiotemporal change in system dynamics (chaotic or/and positions) among clips. A global anomaly focuses on the entire change of dynamics among clips, while a local anomaly focuses on dynamic changes among clips near particular spatial points. We employ a method in which a normal scene model is learned first, and then a testing clip is identified as normal or abnormal by checking the probability that the feature set for the testing clip belongs to the normality model, and anomalies are localized according to the positions of abnormal representative trajectories.

4.1. Probabilistic detection of anomalies

A scene model of normality is learned for a crowd flow employing F_n and using a training set of video clips of normal crowd flows. We use the Gaussian mixture models (GMM) to describe the probability density function of the normality,

$$P(\Gamma|\Phi) = \sum_{k=1}^{K} w_k N(\Gamma; \mu_k, \Sigma_k), \tag{10}$$

where Γ denotes a four (or six if M features are included) dimensional set of random variables with two features for each time series in x and y. We choose training samples from S clips and K is the number of multivariate Gaussian components $N(\Gamma; \mu, \Sigma)$, where μ is the mean and Σ is the covariance. To achieve optimal model learning (that is, for parameter Φ), we employ the expectation-maximization (EM) algorithm incorporating the iterative pairwise replacement algorithm (IPRA) [22] to iteratively learn an optimal number of Gaussian components. Once the normality scene model Φ_m is learned, we apply the maximum likelihood estimation to compute the probability of an observation using equation (10) to judge if it is normal or abnormal by comparing with a probability threshold L_{thres}.

4.2. Anomaly localization

We further localize the anomaly in terms of position and size, once a scene is identified as abnormal. We tackle this problem by analyzing the likelihoods, contributed by each representative trajectory, of an abnormal scene and localize those trajectories with low likelihoods, which are most probably the anomaly sources. Suppose the normality model is learned using F and the feature set of the abnormal scene is denoted by F_a. We locate all representative trajectories that correspond to abnormal features in F_a and cluster them according to position. Clusters with fewer number of trajectories are filtered out using a threshold N_{thres} and the remaining clusters reveal the abnormal regions of the scene.

5. Experimental Results

5.1. Anomaly Detection

The first experiment is for validating global anomaly detection using the unusual crowd activity dataset from University of Minnesota [23]. The dataset includes eleven video sequences of three different crowd scenarios and each sequence begins with normal behavior, followed by a portion of abnormal panic. Figure 8 shows a number of frames selected from three sequences of the three scenarios, respectively. Since the panic ensues for all people involved, we use this dataset to test the global anomaly detection. For this data set, take clips consisting of 10 frames, then compute representative trajectories and chaotic features for each clip. Since computation of chaotic features requires many points in a time series, we interpolate the representative trajectories to get 500 points in each time series.

We select two sequences from each of the three scenarios (six sequences in total) and each sequence is divided into 10-frame clips. We perform the particle advection on each clip by associating a particle to every 2×2 non-overlapping sub window in the image. Figure 9 shows the representative trajectories (193 in each case) of clips 1, 461, and 531.

Then we calculate the chaotic feature set F for the six sequences and learn three normality models for the three scenarios by randomly picking 3/4 of all of the normal clips as training samples. Figure 10 shows the GMM based normality model learned for scenario 3. The red points represent chaotic feature values and the ellipses represent GMM components (47 components in total).

Then we test the anomaly detection using all of the remaining clips (including normal and abnormal) of a sequence and calculate the probability to judge them as normal or abnormal. As an example, Figure 11 shows the likelihood profile of the clips of a sequence in scenario 1, from which we see (between clips 50 and 58) an obvious decrease, corresponding to an anomaly. This detection agrees with the ground truth.

We test all of the six sequences to measure overall performance. Figure 12 shows the ROC curves of our method (a), and method in [9] (b). As can be seen, the performance of our method is promising, having a higher detection accuracy than social force model [9], and pure optical flow [9] in general. Also, in [9], several false positives can be seen, which, however, rarely happens in our results, as demonstrated in Figure 11.

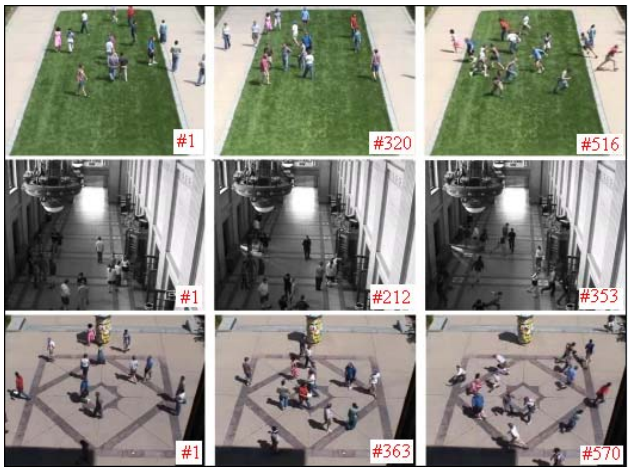


Figure 8. Sample frames from three crowd scenes. The first two frames in each row show normal behavior, and the third frame shows abnormal escape panic.

Figure 9. Representative trajectories for three clips in a sequence, the first one shows normal behavior and the last two are abnormal.

5.2. Anomaly Localization

In this subsection we first demonstrate how to precisely capture where an anomaly occurs and its size in an incoherent scene. Figure 13 shows frames from two clips (29, 30) of a sequence, in which people are clapping, and after a while, two persons enter the scene/field starting to dance while others continue to clap. Our method correctly identified clip 29 as normal and clip 30 as abnormal. From here, we continue to localize the anomaly by locating the persons entering the scene with different dynamics. In Figure 13(b), we show the rough scope of the anomaly using two yellow rectangles where the larger one is the main source of anomaly and the smaller one is a minor source (where a person starts slight dancing from sitting there and changes the dancing behavior). We want to spot the two sources of the anomaly in clip 30.

We find the average of likelihoods of all representative trajectories, and apply threshold to the average to determine if a trajectory is normal or abnormal. The likelihood profile for trajectories in clip 30 is shown in Figure 14. A trajectory whose likelihood is below the average of likelihoods of

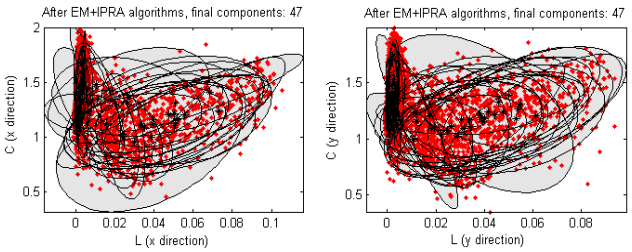


Figure 10. Marginal PDF of two chaotic features of x (left) and y (right) of learned 4-D mixture of Gaussian model.

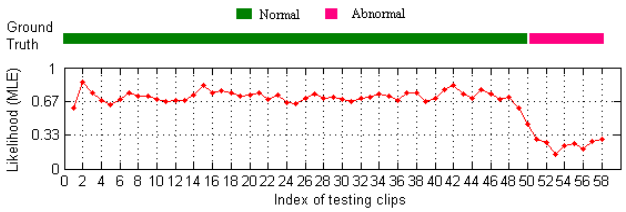
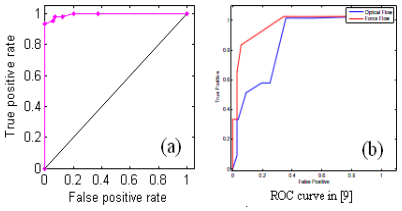


Figure 11. Likelihood profile for testing clips and corresponding ground truth.



Method	Area under ROC
Our method	0.99
Social force model [9]	0.96
Optical flow [9]	0.84

Figure 12. ROC curves for (a) our method, and (b) method of [9].

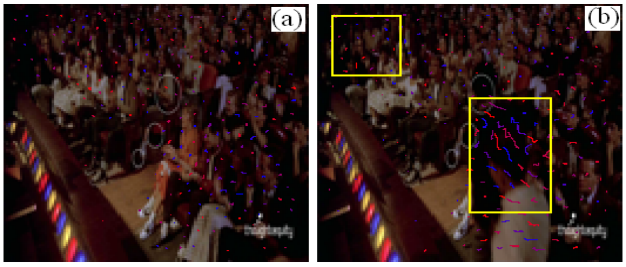


Figure 13. (a) Normal clapping behavior, and (b) introduction of abnormal dancing behavior.

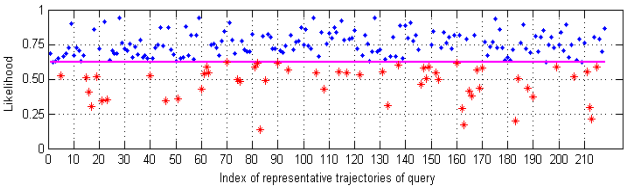


Figure 14. For clip 30 correctly detected anomalies, red points below threshold correspond to abnormal representative trajectories, while blue points above threshold correspond to normal.

all trajectories in a clip is declared as abnormal. Next, we cluster abnormal trajectories based on their location information. This results in 12 clusters, which are shown by thick red rectangles in Figure 15(b). Then some cluster with a fewer trajectories are removed. Finally we obtain the anomaly localization result as shown in Figure 15(c).

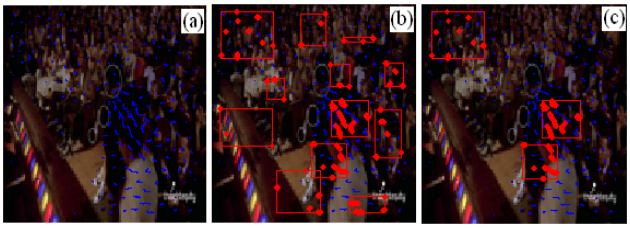


Figure 15. A frame from a clip with abnormal behavior, (a) representative trajectories, (b) candidates for local anomalies, (c) correct localization of anomalies.

Next, we show results on a marathon sequence using the model (Figure 16(a) and (b)) that takes both chaotic and position features into account. For this experiment we created a query clip by synthesizing an anomaly and manually changing the position of a number of representative trajec-

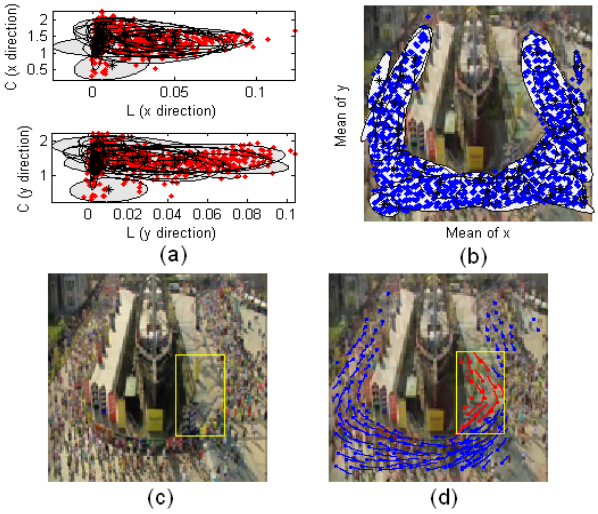


Figure 16. Position-caused anomaly localization.

tories (yellow rectangle, Figure 16(c)). This query clip is correctly detected as abnormal, and we are able to locate the anomaly in one cluster with 21 representative trajectories (plotted in red, Figure 16(d)), which generally agrees with the ground truth. This experiment illustrates the ability of our method to localize the position-caused anomalies in consistent motions, where the change in chaotic dynamics has basically not occurred.

6. Conclusion

We propose a novel method for detecting and localizing anomalies in complicated crowd sequences using a Lagrangian particle dynamics approach, together with chaotic modeling. We define representative trajectories to serve as a compact, yet informative, modeling elements in crowd flows. The method for obtaining representative trajectories is effective, and performs well for a range of extremely crowded to sparse scenes. Representative trajectories also provide a simple means for obtaining time series data, which can effectively be used for chaotic modeling of a scene. We also regulate a representative chaotic feature set to reliably capture the chaotic dynamics of representative trajectories to be used for probabilistic anomaly detection and localization. The promising experiment results outperform the related work.

Acknowledgments: This research was supported by the U. S. Army Research Laboratory and the U. S. Army Research Office under grant number W911NF-09-1-0255.

References

[1] H.T. Nguyen, Q. Ji, and A.W.M. Smeulders. Spatio-Temporal Context for Robust Multitarget Tracking. PAMI, vol. 29, no. 1, pp. 52-64, 2007. 1

[2] M.P. Kumar, P.H.S. Torr, and A. Zisserman. Learning Layered Motion Segmentations of Video. IJCV, 76, 2008. 1

[3] P. Tu, T. Sebastian, G. Doretto, N. Krahnstoever, J. Rittscher, T.Yu. Unified crowd segmentation. ECCV 2008. 1

[4] H. Zhong, J. Shi, and M. Visontai. Detecting Unusual Activity in Video. ICPR 2004, pp. 819-826. 1

[5] O. Boiman and M. Irani. Detecting Irregularities in Images and in Video. ICCV 2005, pp. 462-469. 1

[6] H. Dee and D. Hogg. Detecting Inexplicable Behaviour. BMVC 2004, pp. 477-486. 1

[7] S. Ali, A. Basharat, and M. Shah. Chaotic Invariants for Human Action Recognition. ICCV 2007, Brazil. 2, 3, 4

[8] T. Xiang and S. Gong. Online Video Behaviour Abnormality Detection Using Reliability Measure. BMVC 2005. 1

[9] R. Mehran, A. Oyama, and M. Shah. Abnormal Crowd Behavior Detection using Social Force Model. CVPR 2009, pp. 935-942. 1, 2, 5, 6

[10] S. Ali and M. Shah. A Lagrangian Particle Dynamics Approach for Crowd Flow Segmentation and Stability Analysis. CVPR 2007, Minneapolis, pp. 1-6. 1, 2

[11] E. Andrade, S. Blunsden, and R. Fisher. Modeling Crowd Scenes for Event Detection. ICPR 2006, pp. 175-178. 1, 2

[12] A. Adam, E. Rivlin, I. Shimshoni, and D. Reinitz. Robust Real-Time Unusual Event Detection using Multiple Fixed-Location Monitors. PAMI 30: 555-560, Mar. 2008. 1

[13] L. Kratzm and K. Nishino. Anomaly Detection in Extremely Crowded Scenes using Spatio-Temporal Motion Pattern Models. CVPR 2009, pp 1446-1453. 1, 2

[14] Y. Ke, R. Sukthankar, and M. Hebert. Event Detection in Crowded Videos. ICCV 2007. 1

[15] P. Grassberger and I. Procaccia. Characterization of Strange Attractors. Phys. Rev. Lett. 50, 346-349, 1983. 2

[16] D. Knossow, R. Ronfard, and R.P. Horaud. Human Motion Tracking with a Kinematic Parameterization of External Contours. IJCV, vol. 79, no. 2, pp. 247-269, 2008. 3

[17] J. Sullivan and S. Carlsson. Tracking and Labeling of Interacting Multiple Targets. ECCV 2006, pp. 619-632. 3

[18] M.T. Rosenstein, J.J. Collins, and C.J. de Luca. A Practical Method for Calculating Largest Lyapunov Exponents from Small datasets. Physica D, 65: 1-2, pp. 117-134, 1993. 4

[19] A. Wolf, J.B. Swift, L. Swinney, and J.A. Vastano. Determining Lyapunov Exponents from a Time Series. Physica D, 16: 285-317, 1985.

[20] D.T. Kaplan and L. Glass. Direct Test for Determinism in a Time Series. Phys. Rev. Lett. 68: 427-30, 1992.

[21] E. Ott, W.D. Withers, and J.A. Yorke. Is the Dimension of Chaotic Attractors Invariant under Coordinate Changes? Journal of Statistical Physics, vol. 36, pp. 687-697, 1984.

[22] S.D. Wu and Y.F. Li. Flexible Signature Descriptions for Adaptive Motion Trajectory Representation, Perception and Recognition. Pattern Recognition, 42: 194-214, 2009. 5

[23] Unusual crowd activity dataset made available by the University of Minnesota at: http://mha.cs.umn.edu/movies/crowdactivity-all.avi. 5