Anomalous Crowd Event Analysis Using Isometric Mapping

Aravinda S. Rao, Jayavardhana Gubbi and Marimuthu Palaniswami

Abstract Anomalous event detection is one of the important applications in crowd monitoring. The detection of anomalous crowd events requires feature matrix to capture the spatio-temporal information to localize the events and detect the outliers. However, feature matrices often become computationally expensive with large number of features becomes critical for large-scale and real-time video analytics. In this work, we present a fast approach to detect anomalous crowd events and frames. First, to detect anomalous crowd events, the motion features are captured using the optical flow and a feature matrix of motion information is constructed and then subjected to nonlinear dimensionality reduction (NDR) using the Isometric Mapping (ISOMAP). Next, to detect anomalous crowd frames, the method uses four statistical features by dividing the frames into blocks and then calculating the statistical features for the blocks where objects were present. The main focus of this study is to understand the effect of large feature matrix size on detecting the anomalies with respect to computational time. Experiments were conducted on two datasets: (1) Performance Evaluation of Tracking and Surveillance (PETS) 2009 and (2) Melbourne Cricket Ground (MCG) 2011. Experiment results suggest that the ISOMAP NDR reduces the computation time significantly, more than ten times, to detect anomalous crowd events and frames. In addition, the experiment revealed that the ISOMAP provided an upper bound on the computational time.

Aravinda S. Rao,
The University of Melbourne, Department of Electrical and Electronic Engineering, Parkville Campus, VIC - 3010, Australia. e-mail: aravinda@student.unimelb.edu.au

Jayavardhana Gubbi,
TCS Innovation Labs, Tata Consultancy Services, Bengaluru, Karnataka - 560066, India. e-mail: j.gubbi@tcs.com

Marimuthu Palaniswami
The University of Melbourne, Department of Electrical and Electronic Engineering, Parkville Campus, VIC - 3010, Australia. e-mail: palani@unimelb.edu.au

The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-28658-7_35

1 Introduction

Anomalous event detection is one of the important applications in crowd monitoring. Most video surveillance systems endeavor to detect such events. Video surveillance systems are also used to track targets in multi-camera video surveillance, analyze crowd behavior, provide security for people, and remotely monitor elderly people in home care. Human activity recognition places an important role in many applications, and video surveillance systems facilitate to detect such activities. Different types of human activities exist and examples include “walking,” “running,” “jogging,” “fighting,” “waving,” and other. Surveillance systems are now ubiquitously installed in airports, shopping malls, public transport hubs, stadiums, concerts, etc. [22]. However, the current surveillance applications lack automated analysis of crowd behavior and detecting interested events.

(a) Walking (b) Running (c) Merging (d) Splitting (e) Dispersion (f) Evacuation

Fig. 1 Example frames of crowd events from the PETS 2009 dataset [5].

Automated analysis of interested “events” or “activities” using surveillance systems with intelligent video analytics can deliver a suite of tools to monitor the crowd. However, the algorithms face several challenges in object detection, tracking and detecting human activities. In the case of the human visual system, the eyes and brain coordination are so well-developed that the replication of same process has been a long-standing research in artificial intelligence systems. Many sophisticated algorithms have been proposed over the last four decades, but the algorithms fail in difficult and crowded scenarios. Some of the critical problems faced by computer vision algorithms are the loss of information from 3D to 2D during image formation at the cameras [24], nonuniform illumination, shadows, articulated motions of humans that make algorithms infeasible to track movements, occlusions that mask presence of the objects or object parts.

The quality of the video systems has also improved from standard definition to high definition. The voluminous amount of data along with the problem of human detection and occlusions requires intelligent systems. For real-time crowd monitoring and increased presence of video analytics on cameras, it is critical that video analytics are fast and provide near real-time results. Fast video analytics on cameras saves video bandwidth by not transmitting video to a centralized server, and in addition, avoids unnecessary storage of video footage from every camera in a network. Dimensionality reduction methods provide low-dimensional representation of high-dimensional, complex data.

The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-28658-7_35

In this work, we present a fast approach to detection of anomalous crowd events. Two types of anomalies are being detected: (1) anomalous crowd events and (2) anomalous crowd frames. First, to detect anomalous crowd events, the motion features are captured using the optical flow [7]. A feature matrix of motion information is subjected to NDR using the ISOMAP [18]. The anomalous crowd events are then detected using the hyperspherical clustering [13]. The proposed approach detects the crowd events such as change of directions in people movements, splitting, and dispersion, etc. Next, to detect anomalous crowd frames, the method uses four statistical features (contrast, correlation, energy, and homogeneity) by dividing the frames into blocks and then using the blocks where objects were present. The main contributions of this work include: (1) a new approach to detect anomalous crowd using NDR and (2) analysis of computation time impacting the real-time crowd monitoring and video analytics.

2 Related Work

2.1 Crowd Anomaly Detection

Anomaly detection in crowd events refers to anomalous behavior of the crowd. This is a time-related event, i.e., at certain times crowd events will be normal and other times it will be abnormal (hereafter abnormal,unusual events are called as anomalous) [1, 25]. Outlier detection is one of the main approaches in anomaly detection. A comprehensive survey of anomaly detection based on vision systems is found in [12]. Mahadevan *et al.* [10] consider anomalous events as outliers and use the mixture of dynamic texture (MDT) features. The probability distribution of normal event is learned using the dynamic texture features. Temporal anomalies are detected by first using the Gaussian Mixture Model (GMM) to learn the local intensities and then replacing the foreground intensities with MDT. Spatial anomalies were mapped by computing the discriminant saliency features at each video location using the MDT. The abnormality map of the video was the combined mapping results of both temporal and spatial anomalies.

Adam *et al.* [1] defined small regions in the video for monitoring the flow of moving objects. Optical flow was used to compute the motion vectors. Probability matrix of flow vectors was constructed in the defined region of normal and anomalous events. Chen *et al.* [3] utilized (Lucas-Kanade) optical flow approach to track the feature points. The orientations of the vectors were computed and stored as bins. To determine the dominant force and direction of the motion, force-field model was applied. An anomalous event was detected upon discovering a sudden appearance of force. Wang and Miao [23] used the optical flow vectors to indicate the motion points and the Kanade-Lucas-Tomasi (KLT) corners to track the features. A model was generated by using the motion patterns in different blocks. Anomalous events were detected based on the deviation from the learned normal model. Liao *et al.* [9]

The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-28658-7_35

used four descriptors namely crowd kinetic energy (motion intensity), histogram of motion directions, spatial distribution of motion intensity and localization between two frames for the detection of anomalous events. Liao *et al.* [9] targeted the detection of fighting events. Support Vector Machine (SVM) with Radial Basis Function (RBF) kernel was used to train the 13-dimensional feature vectors and classification of abnormal events.

Tziakos *et al.* [21] used the motion vectors to detect the abnormal events by detecting the motion vectors and then representing them on a low-dimensional manifold using the Laplacian Eigenmaps (LE) [2]. The unusual events were then classified using the supervised Mahalanobis classifier. Thida *et al.* [19] proposed to use the blocks of Histogram of Optical Flow (HOOF) feature corresponding to each frame as features to detect anomalous events. A new NDR method termed as Spatio-Temporal Laplacian Eigenmap (ST-LE), was proposed based on LE. A Novelty classifier was trained using the test samples in the low-dimensional space to classify the anomalous events.

2.2 Dimensionality Reduction

The dimensionality reduction methods can be broadly categorized as linear methods and nonlinear methods. Linear methods assume the existence of features on a linear subspace, whereas nonlinear methods approximate the nonlinear feature space into a linear subspace and then find low-dimensional representations. The dimensionality reduction method can be described as: given a set (data) $X = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]$ with $\mathbf{x}_i \in \mathbb{R}^m$, find a set $Y = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n]$ with $\mathbf{y}_i \in \mathbb{R}^d$ that represents X such that $d \ll m$.

Principal Component Analysis (PCA) [8] reduces the data dimensions by assuming the linear subspace. The objective function used by the PCA is to maximize the variance such that the reconstruction error from the reduced space is minimized in a least square sense. Classical Multidimensional Scaling (MDS) [20] preserves the distances between points from high-dimensional space to lower dimensions. Euclidean distance is the commonly used distance metric between points.

ISOMAP [18] finds the low-dimensional representation in three steps. At first the distance $d(i, j)$ between the neighboring points (i, j) in the input space ($X \in \mathbb{R}^m$) is calculated and a weighted graph is constructed. Next, the distances $d(i, j)$ are used to calculate the geodesic distance between all the points of the input space to find the shortest path. In the last step, classical MDS is applied to weighted graph to construct a d-dimensional embedding from the m-dimensional input space X, where $d \ll m$. ISOMAP considers the global isometry of points between the input feature space and embedded space. ISOMAP is the nonlinear analogue of PCA.

LLE [16] is similar to the ISOMAP but constructs the neighborhood graph based on the local linear properties as opposed to the global approach considered by the ISOMAP. The same localness is maintained in the low-dimensional embedded space. The local linearity is based on the neighboring points in the feature space.

The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-28658-7_35

Lapalacian Eigenmaps (LE) [2] is similar to LLE, however, the weighted graph is constructed based on neighbors defined by a heat kernel. Hessian Eigenmaps (HE) [4] construct the hessian of feature space in a given neighborhood. Then a low-dimensional embedding is found based on the Hessian matrix.

3 Methodology

3.1 Preprocessing and Feature Extraction

Video frames from CCTV cameras contain high-frequency noises due to the nature of the formation of the images. In order to remove these noises, the frames are converted to grayscale. The high-frequency noise is eliminated by Gaussian applying a 2D Gaussian filter with $\sigma = 0.5$ and a block size of 5×5, which is a low-pass filter. The Gaussian filter parameters were chosen based on the method presented in [14]. The feature vector F is a matrix of size $R^{m \times n}$, where m indicates the feature vector length for each frame and n indicates the number of frames in a given video sequence.

3.1.1 Anomalous Crowd Events

After removing the noise, the motion information between frames is computed using the optical flow approach [7]. Let (x,y) denote a pixel in a frame $I(x,y)$. x and y indicate the horizontal and vertical axes, respectively. The optical flow in [7] assumes the brightness constancy in calculating the apparent motion. $E(x,y,t)$ denotes the brightness at point $I(x,y)$, where t represent the time parameter. Then the optical flow vectors are given by [7]

$$E_x u + E_y v + E_t = 0 \tag{1}$$

where $u = \frac{dx}{dt}$ and $v = \frac{dy}{dt}$ are the velocities in horizontal and vertical directions. Feature matrix F is constructed by combining row vectors u and v into a single row vector $[u|v]$, as a row vector for each frame.

3.1.2 Anomalous Crowd Frames

To detect anomalous frames, the proposed method first tracks the objects and determines the centroid in each frame. Then, each frame is divided into blocks of different sizes (8×8, 16×16, 32×32 and 48×64). Now the method compares objects' centroids with block centers and calculate the statistical features for the block where the object centroids were found. The method calculates the four statistical features for the block where the object centroids were found. Statistical features

The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-28658-7_35

such as contrast, correlation, energy, and homogeneity [6], have been found to be useful in many crowd monitoring applications [17]. Thus the feature matrix F is a sparse matrix in which the nonzero blocks indicate one of the four statical features, number of blocks is determined by the block sizes chosen, and feature matrix rows equal the number of frames in a given video.

3.2 Nonlinear Dimensionality Reduction (NDR)

The feature matrix $F \in R^{m\times n}$ is provide as an input to the ISOMAP. The ISOMAP first calculates the $d_F(i,j)$ distances between video frames based on the feature points, i and j, in the input feature space. A weighted graph G will be constructed based on the neighborhood k. During the construction of the graph, the distance between the nodes (between the feature points in this case) are based on the geodesic distance. The geodesic distance $d_M(i,j)$ is the shortest distance between two points. The weighted neighborhood graph is assumed to be isometric to lower-dimensional embedding. Next MDS is applied to the $D_G = \{d_G(i,j)\}$ to find the embedding graph whose dimensions will be d such that the intrinsic geometry is preserved. The low-dimensional embedding is given by vectors y_i such that the following cost function is minimized:

$$E = \left\| \tau(D_G) - \tau(D_Y) \right\|_{L_2}, \tag{2}$$

where τ is an operator that calculates the inner product between nodes and D_Y is the Euclidean distance between frames in the embedded space.

3.3 Anomaly Detection

In this work, we use the anomaly detection scheme devised for environmental sensing applications [13]. The low-dimensional data from the ISOMAP is used to reduce the computational complexity of the data. The anomalous frames are detected based on the following parameters that are iteratively updated: (1) at first, clustering of feature points in the low-dimensional space is achieved with fixed width clustering w, (2) next, the clusters formed by the neighboring featuring points are clustered into a larger cluster. Based on the average inter-point distance, clusters are merged if the inter-point distance is above τ and the distance between points are greater than ψ standard deviation, and (3) finally, the anomalous clusters are identified based on the k-nearest neighbors.

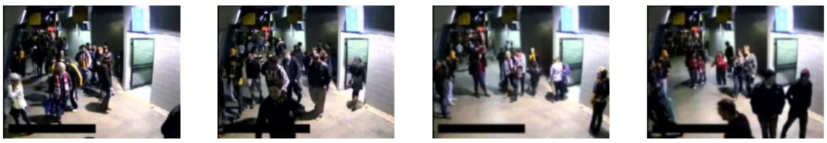
The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-28658-7_35

4 Experiment

The proposed method was implemented in MATLAB 2014 on a 64-bit Windows 7 equipped with 4 GB RAM and Intel® i7 – 2600 CPU running at 3.4 GHz. The anomaly detection algorithms were implemented in Java.

4.1 Dataset

The proposed method has been tested on two datasets: (1) the PETS 2009 [5] and (2) MCG 2011 [14]. The PETS 2009 [5] dataset has three different video sequences (S1, S2, and S3) and each sequence contain different sets : L1, L2 and L3. Each set comes with different timings (such as $13-57$, $14-16, 14-27, 14-31$ and $14-33$). The timings refer to the hour ("hh")–minute ("mm") of data collection. There are eight different views $(001, 002, \ldots, 008)$ for each dataset time. This work uses S1.L1 (with time 13-57, View001) and S1.L2 (with time 14-06, View001) for detecting anomalous frames, and S3 (with timings 14-16, 14-31, and 14-33) for detecting anomalous crowd events, respectively. The PETS 2009 dataset has frames of size 576×768 and in color Joint Photographic Experts Group (JPEG or JPG) format. The PETS 2009 data were collected using Axis cameras. The MCG dataset was collected on four Australian Football League (AFL) matches that were held at the MCG in 2011 and Fig.2 shows the sample frames. The data were collected using six cameras (C1-C6) at 30 fps. MCG data have a frame size of 640×480 and were in RGB color mode collected in Advanced Systems Format (asf). They were then converted to JPEG. In this work data from camera C5 were used and have highly crowded scenes, and the video length is about 20 minutes that makes it ideal to evaluate the computer vision algorithms. In addition, MCG dataset is a natural dataset (i.e., it is not a simulated environment) that provides better data to analyze crowd behavior and events.



(a) (b) (c) (d)

Fig. 2 Sample frames from MCG dataset [14].

To the best of our knowledge, crowd events such as walking, running etc. that can be clearly distinguished have been captured only by the PETS 2009 dataset. Therefore, only the PETS 2009 dataset is used detect anomalous crowd events. To

The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-28658-7_35

detect anomalous frames, this work uses both PETS 2009 and MCG datasets. To track objects, ground truth information provided in [11] was used.

4.2 Results and Discussion

The feature vectors computed were of size $u+v=3942$. The number of frames in the dataset 14-16, View-001, for the walking and running events were 222. Similarly, for 14-31 and 14-33, the number of frames were 130 and 377 respectively. The frame rate was 7 frames per second for the PETS 2009 dataset. Therefore, the size of the feature vectors for the three datasets (14-16, 14-31, 14-33) would be 222×3942, 130×3942 and 377×3942 respectively. In all the experiments, the parameters k and w were set to 3 and 1; merging threshold was set to $\tau = \frac{1}{2}w$ and $\psi = 1$, respectively.

The anomalous crowd frames resulted in video frames people where there were no people or frames which completely dissimilar compared with neighboring frames. Experiment was conducted on PETS 2009 (S1.L1, 13-57, View001 and S1.L2, 14-06, View001) and MCG (16S1, C5) datasets. The feature matrix sizes change depending on the block sizes (8×8, 16×16, 32×32 and 48×64) and number of frames. For PETS 2009 dataset, the feature matrix sizes are: 220×6912, 220×1728, 220×144, and 220×144 for 8×8, 16×16, 32×32 and 48×64 block sizes, respectively. For the MCG dataset, feature matrix sizes are: 31375×1200, 31375×300, and 31375×100 for 16×16, 32×32, and 48×64 block sizes, respectively.

Computation time is an important aspect in the real-world scenarios. Table 1 provides the details of the dataset along with feature matrix size, computational time with and without using the ISOMAP [18] for anomalous crowd event deletion using hyperspherical clustering [13]. Table 1 shows that to detect walking and running events (14-16, View001), the computational time is about 2 minutes and 40 seconds. The time for detecting anomalous events in dataset 14-31, View001, is about 1 minute and 35 seconds and in 14-33,View001 dataset, it is about 4 minutes and 36 seconds. On the other hand, the detection of same anomalous events using the ISOMAP, the computational times are about 1.05, 1.08 and 1.5 seconds, respectively. Table 1 also shows the computational time to detect anomalous crowd frames. In the case of PETS datasets (S1.L2), we see that clustering without using the ISOMAP is almost 10 times higher for block sizes 8×8. In addition, we also see that the clustering computational time (without ISOMAP) reduces as the block sizes are increased. In the case of 32×32 the clustering computational time (without ISOMAP) approximately equals clustering with ISOMAP. In the case of 48×64 block size, the clustering computational time is better than using the ISOMAP. For the MCG dataset, the number of frames makes the clustering algorithm (without ISOMAP) to take longer times (nearly double). This is clearly evidenced from the experiment results provided in Table 1. For example, the contrast features of block sizes 16×16 require 130 seconds for clustering algorithm directly, whereas using

The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-28658-7_35

Table 1 Table provides the computation time recorded to detect anomalous events. It also indicates the number of frames, frame size, and computational time required with and without using the ISOMAP. (Note: the detailed breakdown of computation time (clustering and ISOMAP)) could not be provided due to space limitations)

Dataset	Feature (Video)	Matrix size	Block size	Computational time (in milliseconds) Clustering [13] Without ISOMAP	Clustering [13] With ISOMAP
PETS 2009	Optical Flow (14-16)	222×3942	-	2,040	1,050
	Optical Flow (14-31)	130×3942	-	1,350	1,080
	Optical Flow (14-33)	377×3942	-	4,360	1,500
PETS 2009 (S1.L2 14-06)	Contrast	220×6912	8×8	25,623	2,259
		220×1728	16×16	2,212	2,206
		220×432	32×32	452	2,682
		220×144	48×64	121	2,290
	Correlation	220×6912	8×8	23,997	2,329
		220×1728	16×16	2,057	2,103
		220×432	32×32	366	2,701
		220×144	48×64	113	2,160
	Energy	220×6912	8×8	24,074	2,289
		220×1728	16×16	2,075	2,166
		220×432	32×32	390	4,645
		220×144	48×64	153	2,496
	Homogeneity	220×6912	8×8	24,123	2,313
		220×1728	16×16	2,068	2,110
		220×432	32×32	356	3,030
		220×144	48×64	135	2,129
MCG 2011	Contrast	31375×1200	16×16	130,000	16,369
		31375×300	32×32	13,000	518
		31375×100	48×64	2,870	302
	Correlation	31375×1200	16×16	2,934,000	15,205
		31375×300	32×32	758,000	559
		31375×100	48×64	300,870	312
	Energy	31375×1200	16×16	1,387,080	15,354
		31375×300	32×32	12,499	507
		31375×100	48×64	2,525	311
	Homogeneity	31375×1200	16×16	1,255,480	15,473
		31375×300	32×32	12,494	499
		31375×100	48×64	2,496	319

The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-28658-7_35

ISOMAP, the computational time reduces to 17 seconds. Because of computational intensiveness and difficulties (memory problems) in calculating large feature matrices, the 8×8 have not been reported in this work.

The proposed approach detected events such as people changing the direction, people starting to run, etc (not shown due to limited space). In addition, the proposed approach was also able to detect events such as splitting and dispersion. The ISOMAP computes the low-dimensional embedding based on the global geometry, which is important to maintain the connections between the frames. Another important aspect is that the clustering [13] algorithm endeavors to classify the events in a distributed manner, i.e. the algorithm [13] can be used for detecting anomalies at each camera rather than transmitting the entire video to a centralized video analytics unit. This also assists real-time and large-scale crowd monitoring [13]. The distributed hyperspherical clustering has a computational complexity of $\mathscr{O}(nN_c)$, where n is the number of feature vectors and N_c is the number of clusters for each camera. From Table 1, it is evident that as the n increases, the computational time increases drastically. Therefore, hyperspherical clustering by itself would not be a fast solution to detect anomalies (crowd events and frames) at camera level. From Table 1, we also observe that the anomaly detection using the ISOMAP has nearly a constant time bound. The ISOMAP has a computational complexity of $\mathscr{O}(n^3)$ and therefore, serves as an upper bound for large feature matrices (large n). This is critical and highly important for large-scale, real-time video analytics.

In [15], authors had detected anomalous frames and objects using the Gray Level Co-occurrence Matrix (GLCM) and object paths respectively. It was reported that the frames where there is a relative change compared to others were detected such as scene becoming completely empty after the exist of the entire crowd. In [19], authors proposed to use HOOF to capture the crowd events and then used supervised classifier. However, in this work, we instead use the motion features and iterative clustering algorithm to detect the anomalous frames based on the crowd events. The advantage of this approach is there is no need for training. Only the neighborhood parameter k and cluster width w is required for clustering [13]. From the experiment, we found that for these particular datasets and features, changing of k or w did not have any impact on the detection of anomalous events. In [15], it was also reported that there were computational challenges because of large feature matrices. This problem of computationally demanding task was addressed in this work by the use of ISOMAP. From Table 1, it is evident that the nonlinear manifold algorithms like the ISOMAP can greatly reduce the computational cost. The ISOMAP also preserves the intrinsic geometry of the data, which is very critical in detecting the anomalous events that are temporally related.

In this work, the ISOMAP was used as it is a global dimensionality reduction approach and maintains the inter-point distances. In addition, the study used the ISOMAP parameters, such as the neighborhood parameter $k = 7$ and number of lower-dimensional representation to be equal to ten, and these parameters were determined based on the empirical knowledge. However, further work is required to determine the ISOMAP input parameters and also to examine whether these parameters could be determined automatically without manual intervention. Furthermore,

The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-28658-7_35

more investigation into the performance of whether NDR algorithms that are global, local or a combination of both is required.

5 Conclusion

Anomalous event detection is one of the important applications in crowd monitoring. The detection of anomalous crowd events requires feature matrix to capture the spatio-temporal information to localize the events and detect the outliers. However, feature matrices often become computationally expensive with large number of features. In this work, a fast approach was presented to detect anomalous crowd events and frames using the ISOMAP to reduce feature space and hyperspherical clustering to detect anomalies, respectively. The main focus of the study was to understand the effect of feature matrix size on detecting the anomalies with respect to computational time. In addition, the study revealed that the ISOMAP provides a bound on the computational time and therefore, hyperspherical clustering would also provide a bound in detecting the anomalies. The proposed approach is highly relevant to crowd monitoring and for situations where deploying analytics at camera level is required.

References

1. Adam, A., Rivlin, E., Shimshoni, I., Reinitz, D.: Robust real-time unusual event detection using multiple fixed-location monitors. IEEE Transactions on Pattern Analysis and Machine Intelligence **30**(3), 555–560 (2008). DOI 10.1109/TPAMI.2007.70825
2. Belkin, M., Niyogi, P.: Laplacian eigenmaps and spectral techniques for embedding and clustering. In: NIPS, vol. 14, pp. 585–591 (2001)
3. Chen, D.Y., Huang, P.C.: Dynamic human crowd modeling and its application to anomalous events detcetion. In: 2010 IEEE International Conference on Multimedia and Expo (ICME), pp. 1582–1587. IEEE (2010). DOI 10.1109/ICME.2010.5582938
4. Donoho, D.L., Grimes, C.: Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data. Proceedings of the National Academy of Sciences **100**(10), 5591–5596 (2003)
5. Ferryman, J.: PETS 2009 Benchmark Data (2009). URL `http://www.cvg.rdg.ac.uk/PETS2009/a.html`
6. Haralick, R.M.: Statistical and structural approaches to texture. Proceedings of the IEEE **67**(5), 786–804 (1979). DOI 10.1109/proc.1979.11328
7. Horn, B.K.P., Schunck, B.G.: Determining optical flow. Artificial Intelligence **17**, 185–203 (1981)
8. Hotelling, H.: Analysis of a complex of statistical variables into principal components. Journal of Educational Psychology **24**(6), 417–441 (1933). DOI 10.1037/h0071325
9. Liao, H., Xiang, J., Sun, W., Feng, Q., Dai, J.: An abnormal event recognition in crowd scene. In: 2011 Sixth International Conference on Image and Graphics (ICIG), pp. 731–736. IEEE (2011). DOI 10.1109/ICIG.2011.66
10. Mahadevan, V., Weixin, L., Bhalodia, V., Vasconcelos, N.: Anomaly detection in crowded scenes. In: 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1975–1981. IEEE (2010). DOI 10.1109/CVPR.2010.5539872

The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-28658-7_35

11. Milan, A.: Ground Truth - PETS 2009 (2015). URL http://www.milanton.de/data.html
12. Popoola, O.P., Kejun, W.: Video-based abnormal human behavior recognition—a review. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews **42**(6), 865–878 (2012). DOI 10.1109/tsmcc.2011.2178594
13. Rajasegarar, S., Leckie, C., Palaniswami, M.: Hyperspherical cluster based distributed anomaly detection in wireless sensor networks. Journal of Parallel and Distributed Computing **74**(1), 1833 – 1847 (2014). DOI http://dx.doi.org/10.1016/j.jpdc.2013.09.005
14. Rao, A.S., Gubbi, J., Marusic, S., Palaniswami, M.: Estimation of crowd density by clustering motion cues. The Visual Computer pp. 1–20 (2014). DOI 10.1007/s00371-014-1032-4. URL http://dx.doi.org/10.1007/s00371-014-1032-4
15. Rao, A.S., Gubbi, J., Rajasegarar, S., Marusic, S., Palaniswami, M.: Detection of anomalous crowd behaviour using hyperspherical clustering. In: 2014 International Conference on Digital lmage Computing: Techniques and Applications (DlCTA), pp. 1–8. IEEE (2014). DOI 10.1109/DICTA.2014.7008100
16. Roweis, S.T., Saul, L.K.: Nonlinear dimensionality reduction by locally linear embedding. Science **290**(5500), 2323–2326 (2000). DOI 10.1126/science.290.5500.2323
17. Srivastava, S., Ng, K.K., Delp, E.J.: Crowd flow estimation using multiple visual features for scenes with changing crowd densities. In: 2011 8th IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS), pp. 60–65. IEEE (2011)
18. Tenenbaum, J.B., Silva, V.d., Langford, J.C.: A global geometric framework for nonlinear dimensionality reduction. Science **290**(5500), 2319–2323 (2000). DOI 10.1126/science.290.5500.2319
19. Thida, M., Eng, H.L., Dorothy, M., Remagnino, P.: Learning Video Manifold for Segmenting Crowd Events and Abnormality Detection, *Lecture Notes in Computer Science*, vol. 6492, book section 34, pp. 439–449. Springer Berlin Heidelberg (2011). DOI 10.1007/978-3-642-19315-6_34. URL http://dx.doi.org/10.1007/978-3-642-19315-6_34
20. Torgerson, W.: Multidimensional scaling: I. theory and method. Psychometrika **17**(4), 401–419 (1952). DOI 10.1007/BF02288916
21. Tziakos, I., Cavallaro, A., Xu, L.Q.: Event monitoring via local motion abnormality detection in non-linear subspace. Neurocomputing **73**(1012), 1881–1891 (2010). DOI http://dx.doi.org/10.1016/j.neucom.2009.10.028. URL http://www.sciencedirect.com/science/article/pii/S0925231210001487
22. Valera, M., Velastin, S.A.: Intelligent distributed surveillance systems: a review. In: IEE Proceedings Vision, Image and Signal Processing,, vol. 152, pp. 192–204. IET (2005)
23. Wang, S., Miao, Z.: Anomaly detection in crowd scene. In: 2010 IEEE 10th International Conference on Signal Processing (ICSP), pp. 1220–1223. IEEE (2010). DOI 10.1109/ICOSP.2010.5655356
24. Yilmaz, A., Javed, O., Shah, M.: Object tracking: A survey. Acm computing surveys (CSUR) **38**(4), 13 (2006)
25. Zhao, B., Fei-Fei, L., Xing, E.P.: Online detection of unusual events in videos via dynamic sparse coding. In: 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3313–3320. IEEE (2011). DOI 10.1109/CVPR.2011.5995524

The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-28658-7_35