# A Computer Vision Tool Set for Innovative Elder Pedestrians Aware Crowd Management Support Systems

Sultan Daud Khan[1], Giuseppe Vizzari[2], and Stefania Bandini[2,3]

Makkah Techno Valley,[1]
sdkhan@gistic.org
Complex Systems and Artificial Intelligence Research Center,[2]
University of Milano-Bicocca, Milan, Italy
viz,bandini@unimib.it
Research Center for Advanced Science and Technology,[3]
The University of Tokyo, Japan

**Abstract** As the population of world is increasing, and even more concentrated in urban areas, ensuring public safety is becoming a taunting job for security personnel and crowd managers. Mass events like sports, festivals, concerts, political gatherings attract thousand of people in a constrained environment, therefore adequate safety measures should be adopted. The ageing of the population further increases the urgency of computer supported crowd management support systems especially considering the fragility of elder pedestrians. In recent years, researchers developed several models for simulating crowd dynamics. These models should be properly calibrated and validated by means of data acquired in the field. In this paper, we will describe a computer vision tool set that can provide support to information needs of an integrated crowd management support system.

**Keywords:** ageing, crowd flow analysis, computer vision

## 1 Introduction

Crowd phenomena related to mass events related to sports, festivals, concerts, political gatherings, and so on, are mostly observed in urban areas, which sometimes attract even hundreds of thousands people in addition to normal inhabitants. Pedestrian and crowd modeling research context regards events in which a large number of people may be gathered or bound to move in a limited area; this can lead to serious safety and security issues for the participants and the organizers. Despite all safety measures, crowd disasters still occur frequently. A summary of different recent incidents of crowd disaster can be found in Table 1: although it is not possible to provide a detailed breakdown of the causes of the different casualties (due to the heterogeneity of the events and the involved participants, and sometimes due to the lack of accessible documents analyzing the events),

**Table 1.** Crowd disasters and casualties

| Year | Place | Deaths |
|---|---|---|
| 2015 | Mina, Saudi Arabia | > 2000 |
| 2011 | Stadium, Bamako(Mali) | > 36 |
| 2011 | Pilgrimage, Kerala(India) | 102 |
| 2010 | Loveparade, Germany | 21 |
| 2010 | Water festival, Combodia | > 375 |
| 2006 | Stadium, Yemen | > 51 |
| 2005 | Religious procession, Iraq | > 640 |
| 1990 | Pilgrimage, Saudi Arabia | 1426 |
| 1982 | Stadium, Russia | 340 |

most of these cases were characterized by conflicting flows of relevant number of pedestrians in a constrained space.

The understanding of the dynamics of large groups of people is very important in the design and management of any type of public events. In addition to safety and security concerns, also the comfort of event participants is another aim of the organizers and managers of crowd related events. Large people gatherings in public spaces (like pop-rock concerts or religious rites) represent scenarios in which crowd dynamics can be quite complex due to different factors: for instance, the large number and heterogeneity of participants, their interactions, the process in which they are set, and also exogenous factors like potential dangerous situations that can arise. Such crowding phenomena poses serious challenges to public safety and crowd management. Another important aspect which needs to be highlight here, and obvious from the Table 1, that most of crowd disasters occurred in religious festivals, for example *Hajj*, where more than two million Muslims around the globe came to perform the religious duty at the same time in a constrained environment. Many of the participants in religious events are old and aged people. We will now focus on the Hajj to present statistical data motivating the urgency of the development of computer supported crowd management support systems especially considering elder pedestrians. In particular, in the remainder of the paper, we will describe a computer vision tool set that can provide support to information needs of an integrated crowd management support system.

## 2 Ageing and Hajj

Ageing of population is currently one of the most relevant demographic component in industrialized nations, where it is going to produce significant modifications from the economic, social and cultural perspective. This phenomena should not be considered as the cause of negative consequences, but invested to highlight relationships, needs and potentialities that an ageing society is able to express. In particular, it is necessary to reflect on how the social inclusion of elderly people will be guaranteed in future and how to improve their mobility. Mobility is essential for general independence as well as ensuing good health and quality of life, and one of the most relevant and important activities of daily living for maintaining independence. Although Saudi Arabia is not facing the problems of ageing society but every year they have huge gathering of aged people during hajj as shown in 3 and 4.

There are couple of reasons of aged people coming for hajj. The first reason is that, most of the pilgrims came for performing hajj from developing countries like Egypt, Pakistan, India, etc. as apparent from Figure 2. The per capita income of these countries is very low and most of population is living below poverty level. The population of these countries is very high and usually there is only one bread winner, supporting 5 to 6 members of the family. Under these circumstances, people are not financially stable enough to go for the hajj at the early stages of their lives. They usually safe the money for whole of their lives, so at the end, they could go for the hajj. The second reason is that most of them think, although not true from religious point of view, that if they die during performing hajj, which normally happened due the health problems related to ageing discussed above, they would go the heavens.

### 2.1 Mortality Rate in Hajj

The number of pilgrims are increasing every year and since hajj involves unique migration of large number of people moving from one place to another in extreme hot weather within a constraint environment. Such huge migration of people from one place to another while performing rituals often leads to accidents, such as stampedes and failures of crowd control. In most of the cases stampede occurs due to movement of conflicting flows (moving in opposite directions), for example, the group of people after finishing stoning the devil ritual return and come in conflict with the group of people going to perform the same ritual. Hence panic spreads among the pilgrims in order to avoid being trampled, and many pilgrims died as a result. The number of pilgrims per year died during stampedes is shown in Figure 5.

Beside stampedes the rate of natural deaths among pilgrims is high, since most of the pilgrims are from developing countries are old and often with poor health. Most deaths are due to the cardiovascular and respiratory diseases. Over the past few years, cardiovascular disease become an significant cause of deaths of most of pilgrims. For example, more than 60% of the Intensive care units (ICUs) of hospitals in Mina, Arafat came from cardiovascular reasons. The percentage of pilgrims admitting to hospitals during hajj specific days is higher as illustrated in Figure 1. The percentage of cardiovascular diseases was very high during the hajj 2002, 13.8% admitted to hospitals due to respiratory problems as shown in Table 2.

Analysis of the age distribution revealed that admission to hospitals is often dominated by the pilgrims older than 40 years [1]. Age, in fact, is by far the

[1] Khan N.A., Ishag A.M., Ahmad M.S., El-Sayed F.M., Bachal Z.A., Abbas T.G. Pattern of medical diseases and determinants of prognosis of hospitalization during 2005 Muslim pilgrimage Hajj in a tertiary care hospital. A prospective cohort study, Saudi Medical Journal. 2006;27(9):1373-1380

**Table 2.** Significant causes of death at Hajj

| **Diseases** | **Mortality** |
|---|---|
| Cardiovascular | 45.8% |
| Respiratory | 13.8% |
| Traffic accidents | 6.4% |
| Cerebrovascular | 3.4% |

**Table 3.** Pilgrims age vs Mortality

| **Age(years)** | **Mortality** |
|---|---|
| Less than 20 | 0.0% |
| 20-39 | 3.5% |
| 40-59 | 2.02% |
| 60-79 | 67.5% |
| Greater than 80 | 8.8% |

most important factor in the development of cardiovascular diseases, and since a significant portion of pilgrims comes from countries in which the average age expectancy is lower than in the Western world.

The analysis of crowd dynamics cannot directly reduce the number of casualties related to preexisting health conditions, but it can help detecting and preventing situations in which the density of pilgrims could represent a problem. Researchers from different communities like sociology, civil, physics and computer science are studying crowding phenomena from different angles. Besides these efforts, computer vision research community developing algorithms that can automatically understand the crowd dynamics in the real-world scenes. Despite these efforts, computer vision research community have not achieved the desired level of applicability and robustness. This is due to the fact that the algorithms are based on particular assumptions which are often violated in real-world environment.

During Hajj, every year government of Saudi Arabia deployed more than 100,000 security personnel. In high density crowded areas, surveillance cameras are generally installed in different locations that can even cover the whole crowd scene. Detecting specific activities in real-time videos is the task of analysts sitting in surveillance room and watching over multiple TV screens. Such manual analysis of high density crowds is a tedious job and usually prone to errors. For instance, more than 5,000 surveillance cameras are mounted on different locations in Mina. Still it could not help in preventing the disaster of 2015. Therefore we need automatic analysis of the crowd which can reliably estimate the density of the crowd and detect specific activities. Creating such kind of virtual analyst has become the focus of many researchers. This research has a wide range of application domain in crowd management, public space design, underwater fishes analysis (and animal behavior studies in general), and cell population analysis.

In this paper, we propose a computer vision based tool set, the goal of which is to compute important measurements related to the crowd dynamics. This tool set currently includes the following functionalities: (1) *Crowd flow segmentation*

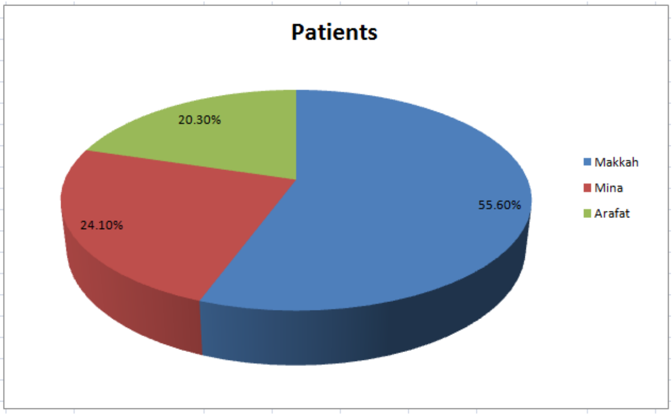


**Figure 1.** Patients admitted to hospitals

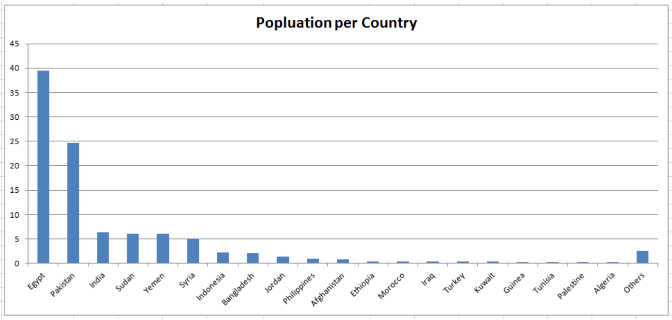


**Figure 2.** Distribution of Hajj pilgrim population per country

*and crowd counting*; (2) *Crowd behavior understanding by identifying source and sink locations*; (3) *Group detection and tracking in crowds*; This tool set can be helpful in initializing and validating crowd simulation models and also provide support to crowd management centers.

## 3 Crowd Flow Segmentation and Crowd Counting

In this section, we discuss an important contribution that this tool set can give to the pedestrian and crowd safety is to localize large and conflicting flows in crowds. Such kind of conflicting motion patterns may lead to the congestions which ultimately ends with crowd disasters. Therefore, early detection of such kind of motion patterns and more importantly, counting the number of people in these situations are important steps for decision makers and crowd managers. The purpose of crowd flow segmentation is to locate those groups in crowds that are distinct and spatio-temporally dominant. In order to achieve the goal of crowd flow segmentation, we generate global representation of the scene by localizing

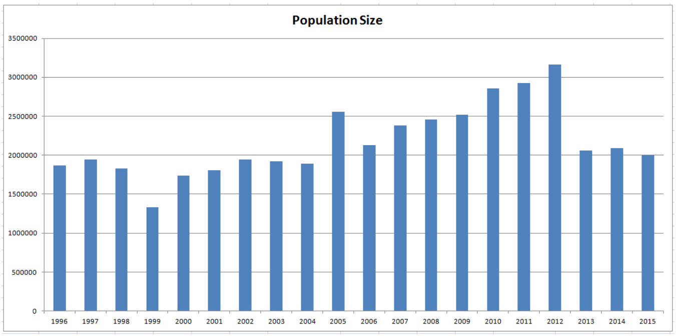


**Figure 3.** Number of Hajj pilgrims per year

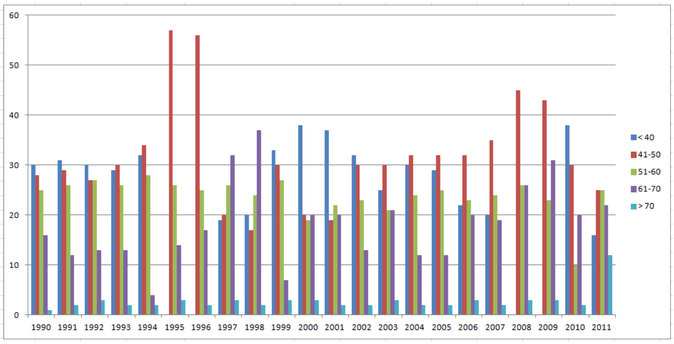


**Figure 4.** Distribution of age for Hajj pilgrims

all the distinct regions/segments in the scene. we extract global representation of the scene by computing dense optical flow that computes a change at every pixel. such kind of global representation of scene makes us independent of detection and tracking of individuals in the scene. Detection and tracking pedestrians are the traditional methods for crowd analysis. But these traditional methods works well in low density situations but robustness of these methods becomes very low when applied to high density scenarios.

In high density situations, the researchers usually extract global information from the scene by using optical flow. Like [1] proposed a dynamical system for crowd flow segmentation by detecting lagrangian coherent structures in the phase space. The work introduced in [9] detects dominant flows by detecting and tracking of SIFT features, whereas in [4] a spectral clustering technique for crowd flow segmentation by computing sparse optical flow is employed. The work described in [12] extract multiple visual features for crowd flow estimation. We observe that after flow segmentation, the above methods fail to detect small flows and unclear boundaries among different flows. Moreover, the above methods are computationally expensive and can not be applicable in real time. A relatively fast method is proposed in [7], where crowd flow is segmented by using derivative

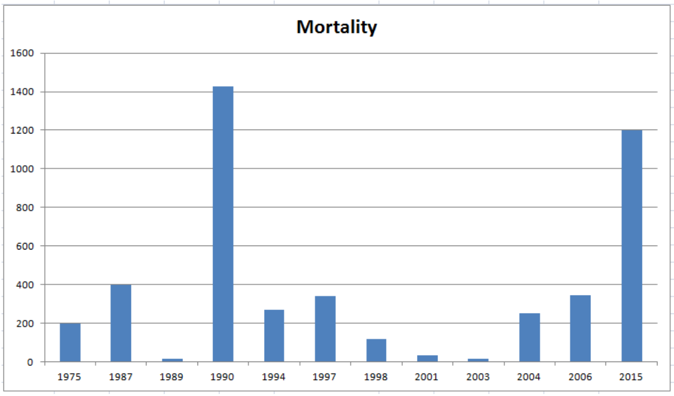


**Figure 5.** Mortality rate due to stampedes and crowd related accidents

curve of the histogram of angle matrix. Since this methods considers only the peaks of the histogram curve, therefore it loses a lot of meaningful information about the crowd flows. In order to capture whole motion information in the scene, we compute dense optical flow followed by the K-means clustering. After K-means clustering small blobs appears at the boundaries of distinct flows, which is removed by our blob absorption approach. Comparing to the state-of-the-art methods, our methods can detect small as well as large flows and by employing blob absorption approach, we detect clear boundaries among distinct flows. After segmentation, we count the number of people in each segment. The overall framework is presented in the next section.

### 3.1 Framework

Our crowd flow segmentation and counting framework is composed of four processing blocks as shown in Figure, Foreground extraction, Flow segmentation, Blob absorption and Counting. In the following, we give details of each processing block.

**Foreground Extraction** Extracting foreground objects is the most important pre-processing step and therefore forms the basis of our framework. Foreground extraction is useful for detection, tracking and understanding the behavior of the object. A survey on motion detection techniques can be found in [8], In traditional visual surveillance, usually with a fixed camera, research use background subtraction methods, where the moving objects are detected, if the intensity values of pixels in the current frame deviate significantly from the background. These kind of methods are prone to noise, a small change in illumination can be detected as foreground. We use adaptive Gaussian mixture model (GMM) to generate a foreground mask, $f_{g(x,y,t)}$, which is more robust to these kind of noises. GMM is very good in separating the foreground objects from the background but it can not compute the change in every pixel of the image. Usually crowded objects move in wide areas, and for flow segmentation problem, we need to detect change in every pixel. Therefore, instead of using the foreground mask generate by GMM, we generate another foreground mask $f_{hs(x,y,t)}$ by computing the dense optical flow, smoothed by gaussian and median filters. We use Horn and Schunk (HS), but any method for dense optical flow computation can be used. Since we compute flow vector at each pixel, so each pixel has the magnitude and direction values. we use magnitude information of the flow vector to generate the foreground mask, all the pixels which have higher magnitude than a predefined threshold will be classified as foreground. Direction information of flow vectors can be used in crowd flow segmentation, since we are segmenting the flows on the basis of orientations. Optimal foreground mask $f_{out(x,y,t}$ is obtained by logical product of $f_{hs(x,y,t)}$ and $f_{g(x,y,t)}$. Later on, we apply morphological processes like morphological opening and closing on the $f_{out(x,y,t}$. Morphological process smooths the section of contours, eliminates small holes and fills gaps in contours. Segmentation block segments the crowd flows into different clusters, $C'_{j(x,y,t)}$, by employing $K$-means clustering followed by blob absorption method. To estimate the number of people in each flow segment, we take logical product of each cluster $C'_{j(x,y,t)}$ and foreground mask $f_{out(x,y,t)}$ and count the number of people by blob analysis and blob size optimization methods.

**Crowd Flow Segmentation** After foreground extraction, the next step is to compute motion flow field. Motion flow field is a set of independent flow vectors and each flow vector is associated with its spatial location and its orientation. Since we compute motion flow field at every frame, therefore we termed it *Instantaneous flow field*, which captures temporal information of the motion patterns and can be used to learn motion patterns in the video. Consider a feature point $i$, its flow vector $F_i$ is represented by its location $X_i$ and its velocity $V_i$, i.e , $F_i = \{X_i, V_i\}$. Since we are computing the flow vector of each feature point that belongs to the foreground objects, therefore, instantaneous motion field is given by $\{F_1, F_2, ...F_n\}$. This motion flow field is a n x 4 matrix, where each row of matrix represents the feature point $i$ and column represents its corresponding spatial location and velocity. Each flow vector represents a motion in a specific direction, therefore, we can not infer any meaningful information about the dominant flows from motion flow field alone. For detecting dominant motion patterns, we need to compute similarity among flow vectors and cluster them into multiple groups. In order to cluster similar flow vectors, we employ $K - means$ clustering algorithm. This process of grouping flow vectors into distinct groups is called flow segmentation.

From our experiments, we observe that after $K - means$ clustering, small blobs appear, these small blobs represents small clusters and detected due to (1), if the object move slowly, the central vectors of pixels are not same as the vectors at the boundaries, and hence clustered into two different groups. (2) the optical flow vectors at the boundaries of two opposite is always ambiguous, and

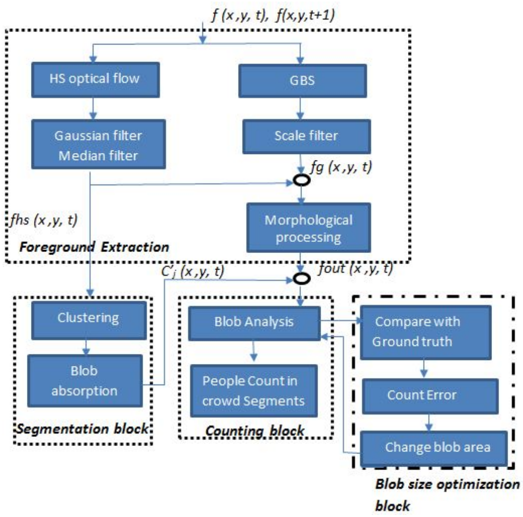


**Figure 6.** Crowd flow segmentation and counting

as a result small clusters appear at the boundaries of two opposite flows. In order to get rid off such kind of noisy clusters, we adopt a blob absorption approach which works on the principle of *Big fish eats small fishes*, where small blobs are absorbed either by a big cluster or by background. In our proposed blob absorption approach, background is also assumed as a big cluster, since in natural images, large amount of pixels correspond to background. A detailed description of blob absorption approach is given in [5]. Figure 7 shows a sample frame from a hajj video sequence, where the people are moving in two dominant directions. After employing K-means clustering (K=4,in this case), crowd is segmented into two main dominant flows with small clusters at the boundaries which are removed after employing blob absorption. After blob absorption, the crowd is segmented into dominant flows with the obvious boundaries. Now, the crowd is segmented into different segments. The next step is to count the number of people in each segment.

**Crowd Counting** In this section, we describe the methodology for counting the people in each segment. In low density crowds, where people are spread sparsely in the environment and each individual is clearly visible, we can use traditional methods of human detection and tracking to count the number of people. Therefore, it implies that in low density situations, counting people is a trivial job. whereas in high density situations, where the people in the environment are tight packed, highly occluded and due to less number of pixels per person, it is extremely challenging to detect the people using a human descriptor and hence it makes the counting problem even more difficult. Therefore, as a solution, we perform global analysis by employing blob analysis and blob size optimization

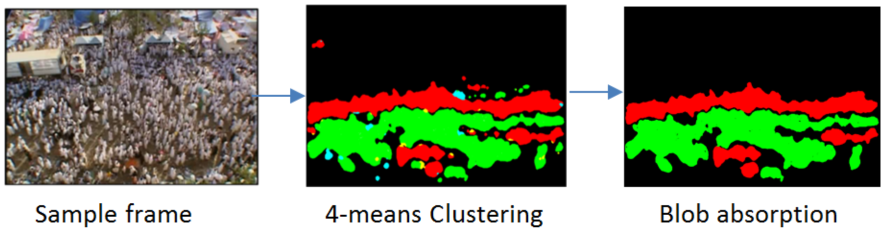


**Figure 7.** Results of 4-means clustering and blob absorption in a Hajj video frame

techniques on foreground image and estimate the number of people in high density crowds.

Blob analysis is a technique that computes statistics for blobs in an image. Blobs are the connected regions in the binary image and usually represent the moving objects in the scene. Since there are many blobs of different sizes representing different moving objects in an image, we need to find out optimum size of the blob that can serve as a threshold. The blobs with size/area above the specified threshold will not be considered (for instance, when counting pedestrians in road videos, these large blobs might be related to cars). For computing optimum size of the blob, we employ blob size optimization algorithm in [5] and [2]. In our experimental set up, in order to find the optimum blob size, we use four or five frames of video selected randomly. For each of the selected frame, we compute optimum size and final optimum blob size $A'$ is the mean of all four or five blob sizes. We use $A'$ for counting people in rest of video frames.

## 4 Crowd Behavior: Identifying Sources and Sinks

Crowded scenes are composed of large number of people. The people in the crowd exhibits different behaviors and understanding crowd behaviors without analyzing the actions of individuals (in crowds) are always advantageous to designers, planners and decision makers. Automatic detection of crowd behaviors have many applications, such as prediction of congestion which lead to unnecessary delays, detection of abnormal events which lead to crowd disasters. Crowd behavior modeling and understanding has important pre-processing task (i) extracting spatio-temporal motion information (e.g, trajectories), (ii) identification of source(entry) and sink(exit) points of trajectories, (iii) interaction of trajectories. We can not extract spatio-temporal motion information(trajectories) by employing crowd flow segmentation framework discussed in section 3, therefore in order to automate the process of crowd behavior understanding, we devise a new framework by adopting two novel algorithms, the first able to generate long, dense, reliable and accurate pedestrian trajectories and the second clustering them into dominant flows. The final global flows not only provide direct information about the characterization of flows but also provide a starting point for the further high level analysis of crowd behavior.

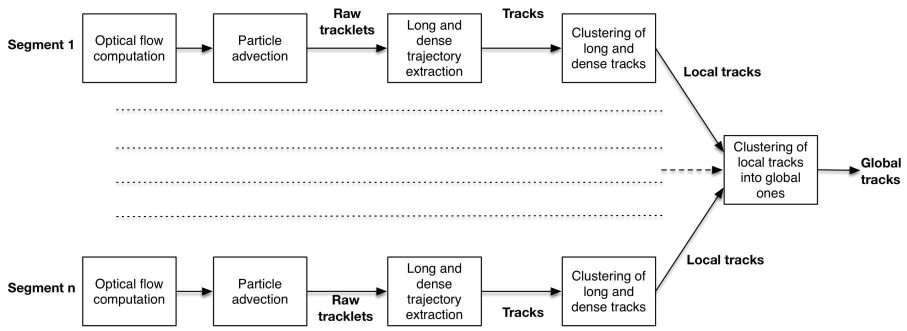


**Figure 8.** Source and sink identification framework

The approach starts by dividing the input video into multiple segments of equal length. The first frame of each video segment is overlaid by grid of particles initializing a dynamical system defined by optical flow as in [11]. Particle trajectories are extracted by integrating the dynamical system over time. We identify sources, sinks and characterize main flows by analyzing particle trajectories using unsupervised hierarchical clustering algorithm, where similarity among the trajectories is measured by Longest Common Sub-Sequence (LCSS) metric. The final global tracks are achieved by clustering local tracks through the same clustering algorithm. The above steps of the framework is illustrated in Figure 8 and described in the following section.

### 4.1 Achieving Reliable Trajectories

As mentioned above, the input to our framework is a sequence of video frames which is automatically divided into $n$ of segments, each of size $k$ frames. Since it is extremely hard to detect and track pedestrians in high density situations, therefore, we rely global analysis by employing optical flow.

**Particle Advection** In order to extract trajectories, we compute dense optical flow between two consecutive frames of every segment. We then initialize a continuous dynamical system by overlaying grid of particles on the initial optical flow field of the video segment and each position of the particle represent the source point. After initializing the grid of particles, the next step is to advect the particles in forward time over the optical flow field. As a result of particle advection, small duration trajectories called *tracklets* are obtained as shown in Figure 9.

**Achieving Final Flows Through Clustering** Tracklets achieved through the particle advection are short duration and therefore have a limited spatial extent. These tracklets fail to represent important characteristics of the overall motion. Moreover, they provide inadequate information that could help in identifying the

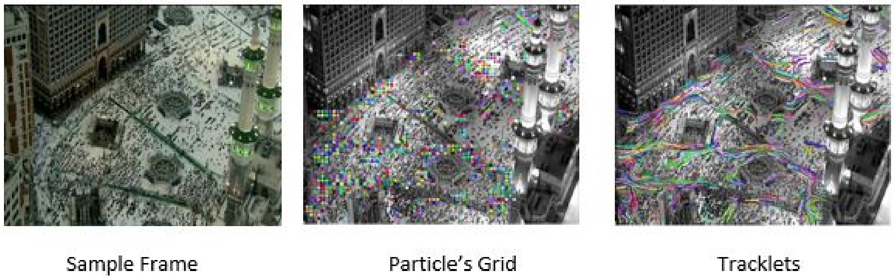


**Figure 9.** First image is the sample frame. Second image are the source points of particles while the third image are the particle trajectories obtained after the advection process

source and sink points of the dominant flows and hence inappropriate to directly applied for behavior understanding. In order to achieve better representation of motion information, we need to cluster these tracklets into longer trajectories. This becomes a combinatorial matching problem that we define and solve recursively.

Our proposed clustering algorithm is based on the assumption that the tracklets corresponding to single motion pattern are similar in orientation but the sources and sinks are spatially different. These tracklets start and end at different locations but the sink of one tracklet lie spatially close to the source of other tracklet. The algorithm in [6] exploits spatial closeness of source and sink locations and similarity among the tracklets by combining them into longer tracks.

After achieving long tracks, the next step is to cluster similar tracks into local tracks by employing hierarchical clustering algorithm using the following procedure.

1. We sort the tracks in descending order on the basis of their length. We compute the length of track as euclidean distance between its start and end point. Let $L_T$ is the sorted list of tracks.
2. We also set up a list of clusters $L_C$, initially containing one cluster associated to the first track $T_1$ (the longest one) in the list and it is considered as an initial cluster center.
3. We select bottom most track from the list, $T_s$, and compare it with the centers of all clusters present in $L_C$ using longest common sub-sequence metric. If similarity value is greater than a threshold, then tracklet $T_s$ is assigned to the current cluster, otherwise, we initialize a new cluster with the center $T_s$. We delete the tracklet $T_s$ from the list $L_T$ after assignment to a cluster.
4. If the cluster's size exceeds a positive value of $S$, then we update the cluster center by using $K^{th}$ order least square polynomial regression. We use $S = 30$ in our experiments.
5. We repeat the previous step until $L_T$ is not empty.

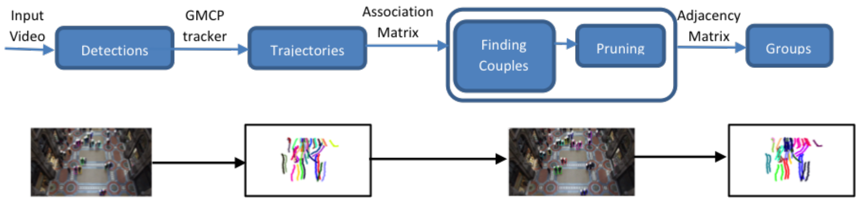


**Figure 10.** Proposed Methodology for Group Detection.

## 5 Group Detection in Crowds

In the previous frameworks, we tried to characterize crowd with a macroscopic perspective, providing information describing the behavior of segments or flows of pedestrians. However, in relatively low density situations it can be useful to characterize in a more fine grained way the analyzed situation. For instance, it has been where most of the people tend to move in groups. The members of the groups tend to maintain spatial and temporal correlations and therefore behavior of the crowd is usually influenced by these social relationships. Realizing the importance of group behavior and its influence on crowd, we propose an approach for automatic detection and tracking of groups in crowds.

The proposed approach starts by detecting individual pedestrians in video frame and then track the detected pedestrian through multiple frames using GMCP tracker by [13]. The trajectory of pedestrian is a set of tuples $(x, y, t)$, where $x$ and $y$ are the horizontal and vertical coordinates of the location at time $t$. We then define an *Association Matrix*, which captures the joint probability distribution of source and sink locations of all pedestrian trajectories. In order to capture the probability distributions of source and sink locations of trajectories, we assume two discrete random variables $\mathbf{X}$, representing "source" locations of the trajectories and $\mathbf{Y}$ representing "sink" locations.

An *Association Matrix* for $n$ trajectories is shown below.

$$P(X,Y) = \begin{Bmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1n} \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & p_{n3} & \cdots & p_{nn} \end{Bmatrix}$$

Each element of *Association Matrix* shows the probability distribution of source and sink location of a single pedestrian $k$ over all other $n$ pedestrians. Association matrix captures the walking behavior of a pedestrian relative to other pedestrians in the scene. A single pedestrian who is not a member of any group tends to stop or move freely in the environment. Moreover, he tends to keep his distance from other pedestrians. In the same way, members of the group tend to maintain small proximity within members and large with other individuals.

(a) Hotel Sequence (b) Gallery Sequence

**Figure 11.** Qualitative results of different video sequences

This type of behavior uniquely identify the group which can be captured by the association matrix.

A three step bottom-up hierarchical clustering approach is employed to discover couples. In the first step, we assign distinct cluster identifiers by treating each pedestrian as a separate cluster. In the second step, the algorithm adopts a greedy approach to find a best possible member for each pedestrian to form a couple. The algorithm groups two pedestrians in a group by measuring the difference between their probability distribution by using *Kullback-Leibler (KL) divergence*, also known as relative entropy, denoted by $D_{KL}(P_r||P_k)$. If $KL$ value is more than a predefined threshold, then pedestrians are termed as bad couples. After pruning of bad couples, groups are discovered using *Adjacency Matrix*, which captures the connectivity information among all pedestrians. Figure 11 shows qualitative results of the proposed framework.

## 6 Application of Computer Vision Toolset

The proposed computer vision tool set has two types applications, one to crowd manager and other to the modeler as shown in Figure 12. Due to to the complex dynamics of the crowd, crowd management is becoming a daunting job for the crowd managers and security staff. In such high density crowded situations, to ensure the safety of people, low cost surveillance cameras are installed at different locations that can cover the whole crowd. The analysts sitting in the surveillance room watching over multiple TV screens in order to detect some abnormal events. Such manual analysis of crowd is a tedious job and usually error prone. These problems necessitate the development of methods and tools that can automatically analyses the crowd and can give reliable estimate about the density and detect specific activities. The proposed tool set can provide information about the crowd size, distinct motion patterns, crowd behaviors and pedestrian groups detection.

During the last 15 years, many crowd and pedestrian simulation models have been proposed in literature [3]. Simulation models has been providing a support

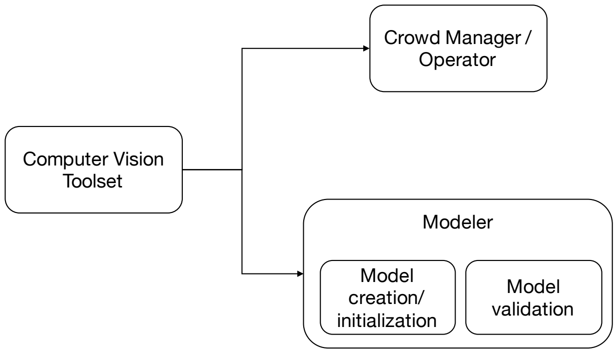


**Figure 12.** Application of proposed computer vision tool set

in decision-making since last many years. These simulation models mimic the real-world crowds but they must be calibrated and validated to assure the plausibility of their results. For example, in several crowd simulation models, it is desirable to model a real scenario, and in order to study how the physical environment would affect the flow of the people. In this case, it is important to achieve information about the motion of people, density, dominant directions, velocity. The manual extraction of this kind of information is very tedious, time consuming and usually prone to errors. This motivates the use of proposed computer vision tool set that can provide useful information that could be helpful in initial configuration of simulation models and also provide support in the validation phase.

In Figure 13, we presented an example where we capture information from the real time video, where the people are circulating around the Kaaba performing a religious ritual. In this example, we capture movements of the people by employing our method discussed in section 3, where a optical flow based dynamical system is initiated followed by the particle advection. As result of advection process, trajectories are obtained as shown in Figure 13(a). The achieved trajectories are clustered into a single dominant and coherent flow as shown in Figure 13 (b). This information is fed to the simulator in order to reproduce spiral movements of the people around a central object. The pedestrian motion is modeled by using cellular automata (CA). A more realistic behavior is obtained by incorporating floor field to the wall avoidance and lane formation as in [10]. In order to validate the circular movement of the people, we also initialize an optical flow base dynamical system followed by particle advection for the simulated video as shown in Figure 13(c). Trajectories achieved after applying particle advection process on a simulated video, are clustered into a single flow as shown in Figure 13(d). The circular flow detected in simulated video highlights the fact that the spiral movements of pedestrians are accurately modeled.

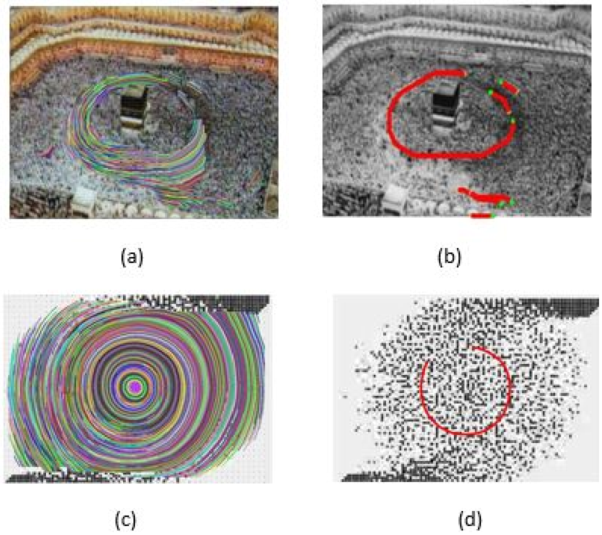


**Figure 13.** (a) trajectories extraced during the particle advection process on real time video, (b) final circular flow in real time video, (c) trajectories extracted through particle advection process in simulated video, (d) final circular flow in simulated video

## References


1. Ali, S., Shah, M.: A lagrangian particle dynamics approach for crowd flow segmentation and stability analysis. In: 2007 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2007), 18-23 June 2007, Minneapolis, Minnesota, USA. IEEE Computer Society (2007), http://dx.doi.org/10.1109/CVPR.2007.382977
2. Arif, M., Khan, S.D., Basalamah, S.: Counting of people in the extremely dense crowd using genetic algorithm and blobs counting. IAES International Journal of Artificial Intelligence 2(2), 51 (2013)
3. Crociani, L., Piazzoni, A., Vizzari, G., Bandini, S.: When reactive agents are not enough: Tactical level decisions in pedestrian simulation. Intelligenza Artificiale 9(2), 163–177 (2015), http://dx.doi.org/10.3233/IA-150086
4. Eibl, G., Brändle, N.: Evaluation of clustering methods for finding dominant optical flow fields in crowded scenes. In: 19th International Conference on Pattern Recognition (ICPR 2008), December 8-11, 2008, Tampa, Florida, USA. pp. 1–4. IEEE Computer Society (2008), http://dx.doi.org/10.1109/ICPR.2008.4761911
5. Khan, S.D., Vizzari, G., Bandini, S., Basalamah, S.: Detecting dominant motion flows and people counting in high density crowds. Journal of WSCG 22(1-2), 21–30 (2014)
6. Khan, S.D., Bandini, S., Basalamah, S.M., Vizzari, G.: Analyzing crowd behavior in naturalistic conditions: Identifying sources and sinks and characterizing main flows. Neurocomputing 177, 543–563 (2016), http://dx.doi.org/10.1016/j.neucom.2015.11.049

7. Li, W., Ruan, J.H., Zhao, H.A.: Crowd movement segmentation using velocity field histogram curve. In: 2012 International Conference on Wavelet Analysis and Pattern Recognition. pp. 191–195 (July 2012)
8. Moeslund, T.B., Granum, E.: A survey of computer vision-based human motion capture. Computer Vision and Image Understanding 81(3), 231–268 (2001), http://dx.doi.org/10.1006/cviu.2000.0897
9. Ozturk, O., Yamasaki, T., Aizawa, K.: Detecting dominant motion flows in unstructured/structured crowd scenes. In: 20th International Conference on Pattern Recognition, ICPR 2010, Istanbul, Turkey, 23-26 August 2010. pp. 3533–3536. IEEE Computer Society (2010), http://dx.doi.org/10.1109/ICPR.2010.862
10. Shimura, K., Khan, S.D., Bandini, S., Nishinari, K.: Simulation and evaluation of spiral movement of pedestrians: Towards the tawaf simulator. J. Cellular Automata 11(4), 275–284 (2016), http://www.oldcitypublishing.com/journals/jca-home/jca-issue-contents/jca-volume-11-number-4-2016/jca-11-4-p-275-284/
11. Solmaz, B., Moore, B.E., Shah, M.: Identifying behaviors in crowd scenes using stability analysis for dynamical systems. IEEE Trans. Pattern Anal. Mach. Intell. 34(10), 2064–2070 (2012), http://dx.doi.org/10.1109/TPAMI.2012.123
12. Srivastava, S., Ng, K.K., Delp, E.J.: Crowd flow estimation using multiple visual features for scenes with changing crowd densities. In: 8th IEEE International Conference on Advanced Video and Signal-Based Surveillance, AVSS 2011, Klagenfurt, Austria, August 30 - September 2, 2011. pp. 60–65. IEEE Computer Society (2011), http://dx.doi.org/10.1109/AVSS.2011.6027295
13. Zamir, A.R., Dehghan, A., Shah, M.: Gmcp-tracker: Global multi-object tracking using generalized minimum clique graphs. In: Fitzgibbon, A.W., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) Computer Vision - ECCV 2012 - 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part II. Lecture Notes in Computer Science, vol. 7573, pp. 343–356. Springer (2012), http://dx.doi.org/10.1007/978-3-642-33709-3_25