Neurocomputing 242 (2017) 28–39

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Density independent hydrodynamics model for crowd coherency detection

Habib Ullah[a,*], Muhammad Uzair[a], Mohib Ullah[b], Asif Khan[c], Ayaz Ahmad[a], Wilayat Khan[a]

[a] Department of Electrical Engineering, COMSATS Institute of Information Technology,Wah Cantt, Pakistan
[b] Norwegian University of Science and Technology, Gjovik, Norway
[c] Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Topi, Pakistan

ARTICLE INFO

Article history:
Received 6 August 2016
Revised 22 December 2016
Accepted 3 February 2017
Available online 21 February 2017

Communicated by Jungong Han

Keywords:
Coherency detection
Crowed flow analysis
Smoothed particle hydrodynamics

ABSTRACT

We propose density independent hydrodynamics model (DIHM) which is a novel and automatic method for coherency detection in crowded scenes. One of the major advantages of the DIHM is its capability to handle changing density over time. Moreover, the DIHM avoids oversegmentation and thus achieves refined coherency detection. In the proposed DIHM, we first extract a motion flow field from the input video through particle initialization and dense optical flow. The particles of interest are then collected to retain only the most motile and informative particles. To represent each particle, we accumulate the contribution of each particle in a weighted form, based on a kernel function. Next, the smoothed particle hydrodynamics (SPH) is adopted to detect coherent regions. Finally, the detected coherent regions are refined to remove the effects of oversegmentation. We perform extensive experiments on three benchmark datasets and compare the results with 10 state-of-the-art coherency detection methods. Our results show that DIHM achieves superior coherency detection and outperforms the compared methods in both pixel level and coherent region level average particle error rates (PERs), average coherent number error (CNE) and F-score.

1. Introduction

Crowd flows represent movements of group of individuals that are pervasive in many real-world environments. Automatic coherency detection in crowded scenes is a challenging computer vision problem [1–3] and is useful in effectively decomposing scenes into meaningful parts. These parts can be exploited for automatic anomalous events detection and recognition.

The coherency and density of crowd are correlated. A uniform density over time represents a coherent flow of crowd. The coherency changes with the changing density flowing in multiple directions. Previous works [4–8] assumed that the level of movement remains uniform in a crowded scene. This means that the coherency of both low-density and high-density crowd flows remain consistent over time. However, this assumption may' not hold in many real-world scenarios where the density of people is changing over time. For example, high density of pedestrians on a pedestrian pathway can be observed during office hours whereas the same density reduces in later hours.

To address these challenges, we propose a density independent hydrodynamics model (DIHM) to locally model the movement of crowd without distinguishing pedestrians individually. Our proposed DIHM is based on smoothed particle hydrodynamics (SPH) [9], which is extensively used to solve fluid dynamics problems [10–13]. Our motivation for the SPH directly comes from the observation that the crowd flows in videos resemble fluid flows. In fact, SPH models both compressible and incompressible liquids, which implies that it is independent of changes in the volume. Considering these capabilities of the SPH, DIHM models coherent regions in the crowd and can cope with the density of people that varies over time.

We consider each moving object as part of the crowd and non-moving objects or groups of people as a background. Our proposed DIHM method is depicted in Fig. 1. Firstly, we extract a motion flow field from the video using the Farnebäck optical flow technique [14]. We then incorporate the enthalpy model [15] to remove the static particles that do not contribute to the detection of coherency. Subsequently, the orientation information of the particles is collected where each particle represents the location of a pixel. Secondly, the coherent regions in the scene are detected by employing the SPH model. Finally, to consolidate and to refine the coherency detection, an unsupervised, robust and

* Corresponding author.
E-mail address: habibkhandr@yahoo.ca (H. Ullah).

http://dx.doi.org/10.1016/j.neucom.2017.02.023

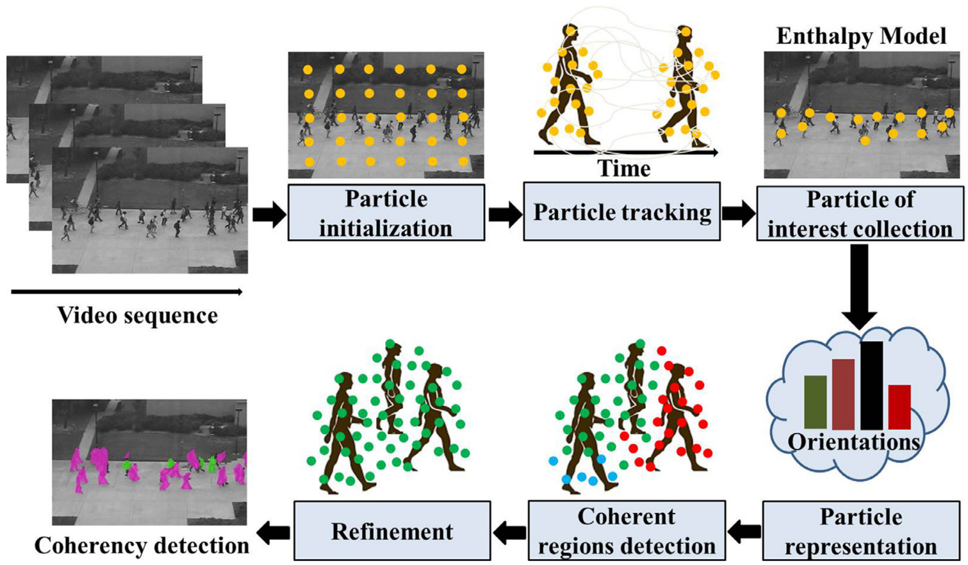

Fig. 1. Illustration of the proposed method. Each particle represents the position of a pixel. For the sake of visualization, only a limited set of particles are overlaid on the frame. Particle of interest collection removes static particles associated with non-motion regions. The retained particles are used to detect coherency in the next steps.

efficient multilayer spectral clustering (MLSC) [16] is exploited to group regions that are coherent both in appearance and motion. Thus, the overall framework renders more consistent coherency detection.

Our main contributions include the development of a novel density independent hydrodynamics model (DIHM) in conjunction with the enthalpy model for improved crowd coherency detection. One of the major attractions of the DIHM is its capability to handle changing density over time. To the best of our knowledge, we are the first to propose these models for coherency detection. Moreover, we extensively evaluate the proposed method on three benchmark datasets and compare our results with 10 state-of-the art methods. Our results show that the proposed method significantly outperforms all 10 state-of-the-art methods both qualitatively and quantitatively. Preliminary results of our proposed work on a few video sequences were presented in [17,18] where we detected only traffic accident [17] and dominant flows [18].

To evaluate the performance of the DHIM, we compare the results with 10 state-of-the-art coherency detection methods including the lagrangian particle dynamics (LPD) [4], the mixtures of dynamic textures (MDT) [5], the motion segmentation in crowds (MSC) [7], the spatio-temporal model (STM) [19], the local translation domain (LTD) [8], the detection of coherent motion (DCM) [20], the collective motion detection (CMD) [21], the segmentation based on dynamic system (SDS) [22], the trajectory clustering approach (TCA) [23], and the thermal diffusion process (TDP) [6]. Our results show that the DIHM achieves superior coherency detection. Moreover, our proposed DIHM outperforms the compared methods in both pixel level and coherent region level analysis in terms of average particle error rates (PERs), average coherent number error (CNE) and F-score.

The rest of the paper is organized as follows. In Section 2, an overview of related work is provided. The proposed method is presented in Section 3. Experimental results on the benchmark datasets are shown in Section 4 and the conclusion is presented in Section 5.

2. Related work

Since the methods for crowd coherency detection and anomaly detection are related to each other, we divide both of them into three broad categories based on the density of crowd. The methods targeting a maximum of two individuals are categorized as *individual level analysis*. Similarly, the methods targeting 15 and more than 15 individuals are grouped under the terms *low density flow analysis* and *high density flow analysis*, respectively. Table 1 summarizes the methods covered in this section according to their category, features and models used for representing coherency detection and anomaly detection, as well as the datasets on which these methods are evaluated.

In the individual level analysis category, Poling and Lerman [24] use nonlinear embedding of two-view point correspondences into a 9-dimensional space and identify the different motions by partitioning lower-dimensional subspaces. Narayana et al. [25] use optical flow orientations instead of the complete vectors and exploit the well-known property that under camera translation, optical flow orientations are independent of object depth. They introduce a probabilistic model that automatically estimates the number of observed independent motions and results in a labeling that is consistent with real-world motion in the scene. Shi et al. [26] exploit discrete cosine transform and a two-stage clustering strategy for tracked points to facilitate division of incomplete and corrupted trajectories against severe data missing and noises. Rahmati et al. [27] integrate prior knowledge in the form of weak labeling into motion segmentation. Using the example of Cerebral Palsy detection, motion patterns of infants are segmented into the different body parts by analyzing body movements. Qin et al. [28] combine the region saliency based on entropy rate superpixel with the affinity propagation clustering algorithm to get seeds in an unsupervised manner, and use random walks method to obtain the coherency results. Zhang et al. [29] use line integral convolution and information entropy to segment the binarization image where the size of the crowd is estimated by least squares fitting to count the number of individuals in a crowd. Wu et al. [30] propose a convex texture image segmentation model by extracting Gabor features and gray level co-occurrence matrix. Then, these features are fused together to effectively construct a discriminative feature space by concatenating with each other. Li et al. [31] deal with challenges in the motion segmentation problem, including perspective effects, missing data, and unknown number of motions. The 3-D motion segmentation is first formulated from two perspective views as a subspace clustering problem. It then combines the point correspondence information across multiple image frames via a collaborative clustering stage.