Pattern Recognition Letters 44 (2014) 141–151

Contents lists available at ScienceDirect

Pattern Recognition Letters

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

A bio-inspired multi-camera system for dynamic crowd analysis ☆

Dimitrios Chrysostomou [c,*], Georgios Ch. Sirakoulis [b], Antonios Gasteratos [a]

[a] Laboratory of Robotics and Automation, Democritus University of Thrace, Dept. of Production Engineering and Management, Vas. Sophias 12, GR-671 00 Xanthi, Greece
[b] Laboratory of Electronics, Democritus University of Thrace, Dept. of Electrical and Computer Engineering, Panepistimioupoli, Kimmeria, GR-671 00 Xanthi, Greece
[c] Department of Mechanical and Manufacturing Engineering, Aalborg University, Fibigestraede 16, DK-9220 Aalborg East, Denmark

A R T I C L E I N F O

Article history:
Available online 8 December 2013

Keywords:
Crowd analysis
Bee colony
Multi-camera
Spider agents
Dynamic
Bio-inspired

A B S T R A C T

Analysis of crowd density has emerged nowadays as a hot topic issue related to the crowd safety and comfort and directly depended on the design and the operation of the crowded places under study. Usually multiple camera networks are employed to cover, monitor and improve the safety of people in large multifunctional crowded buildings. On the other hand, the art gallery problem is a computational geometry approach to a classical real-world visibility challenge. In a nutshell, it concerns the minimization of the free moving guards required to observe the entire gallery. In this paper we attempt to approach this problem from a novel perspective. To begin with, the number of guards are replaced by multiple cameras whose number should be minimized. At the same time, the observability of the camera network in the available space should be dynamically maximized, so as to observe the evolving density of the crowded areas adequately. In order to achieve this objective a twofold bio-inspired method is described and implemented, based on the emergent computation of swarms to come up with solutions in complex mathematical problems. More specifically, the observations on bumblebee colonies lead us firstly to the definition of artificial bumblebee agents used to determine the number of cameras needed to maximize the observability of a space given the safety specifications emerged from the crowd analysis. Secondly, the way the spiders wave their webs was used as a source of inspiration to determine the exact positions of the cameras in the given space by artificial spider agents. The feedback of the algorithm is then used to cover the areas with significant crowd density in a dynamic fashion. Experimental results show that the algorithm is capable of producing promising results where the areas with the maximum crowd density are continuously detected and covered in a dynamic way.

1. Introduction

As the worldwide urbanization keeps the population density levels high, crowd analysis applications are more timely than ever. In particular, a variety of applications such as crowd management, public space design, virtual environments, visual surveillance and intelligent environments are of great interest and are involved in the analysis of crowd dynamics and behavior. Furthermore, the rising popularity of multi-camera systems in conjunction with their low cost, rank them among the best solutions for a large number of problems related with the context of crowd analysis. On the other hand, different features in crowd analysis, such as crowd density, recognition, tracking, monitoring, etc. dictate the employment of specific computer vision techniques to tackle the problem of automatically extracting information to characterize different crowd events. Visual surveillance methods have been developed to estimate motion of objects and people in the scene, either in isolation or in groups (Hu et al., 2004). The variety and topology of sensors influence the process of scene capture while environmental conditions, such as natural and artificial illumination changes, often introduce great amount of noise. Moreover, the typology of the scene affects the required process to extract the most accurate information of a dynamic scene.

Among many efforts attempting crowd analysis by means of computer vision, Velastin et al. (2004) proposed a solution for occlusion handling in the context of crowd management, while Lin et al. (2001) calculated the crowd density and the optimal exit path using the visual feedback of the cameras in crowd evacuation situations. Especially for extensive public areas such as airports and stadiums, a multi-camera system is essential to deliver adequate data for segmentation, detection and tracking of multiple people in a cluttered environment (Mittal and Davis, 2003; Calderara et al., 2009). A thorough review of computer vision techniques and implementations for crowd analysis have been presented by Zhan et al. (2008) and Junior et al. (2010). It is clear that one of the most significant factors that determines the performance of a dynamic multi-camera arrangement is the occlusion avoidance of the observed space. Depending on the application, the occlusions

☆ This paper has been recommended for acceptance by Simone Calderara.

* Corresponding author. Tel.: +30 25410 79330; fax: +30 25410 79331.
E-mail addresses: dimi@m-tech.aau.dk (D. Chrysostomou), gsirak@ee.duth.gr (G.Ch. Sirakoulis), agaster@pme.duth.gr (A. Gasteratos).

0167-8655/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.patrec.2013.11.020

in a predefined space should be either null or restricted. Thus, an important issue in designing visual sensor arrangements is the appropriate placement of the cameras such that they achieve the minimum possible occlusions, i.e. the maximum visibility. The main focus of this paper is the maximization of the coverage with respect to overall crowd density, a dynamic camera topology and the network's capability to efficiently observe crowd dynamics.

In the view of the foregoing, the renowned Art Gallery Problem *(AGP)*, a well-studied visibility problem in computational geometry, gave thrust to studies of multi-camera topology in 3D monitoring space. Chvatal (1975) introduced AGP as the problem of guarding an art gallery with the minimum number of guards to observe the entire gallery. The layout of the art gallery is a close polygon and the covering points (vertices on the polygon) corresponds to the guards. The art gallery theorem, provides an upper bound on the minimal number of guards: $n/3$ guards are always sufficient and sometimes necessary to guard a simple polygon with m vertices. The original art gallery theorem is known to be a NP-hard problem (Culberson and Reckhow, 1994). Many variations of AGP have been studied in previous works that address a plethora of restrictions. More specifically, Urrutia (2004) introduces new directions of research based on watchman routes and floodlight illumination problems in his extensive survey. Recently, Bottino and Laurentini (2011) proposed an incremental algorithm for interior and edge covering which produces results with nearly optimal performance or close to the lower bound of the polygonal environment while Couto et al. (2011) proved that an exact solution is possible by discretizing the examined polygon in $O(n^3)$ iterations. As the majority of these formulations result in NP-hard problems, a profusion of optimization approaches have been studied and used in several different scenarios. A subset of them utilized only for perspective cameras and used for performance comparison with our method are a *Greedy* approach by Erdem and Sclaroff (2006), a *Branch* & *Bound* technique by David et al. (2007) and a *Genetic* method by Yao et al. (2010). Many recent works, examine the nature of this problem not only in the context of surveillance but of ambient intelligence too. Plenty applications of self-organization approaches have been presented lately as Bandini et al. (2010) proposed self-organization models for adaptive environments while Montagna et al. (2012) presented gradient-based patterns for anticipative adaptation. Besides, Dantu et al. (2012) examined the impact of deterministic and stochastic approaches in micro-aerial vehicle collectives and Bicocchi et al. (2012) worked towards self-organizing virtual macro sensors. However, in our problem there is a significant alteration. The main goal of the plethora of applications is to find the minimum number of guards that are able to move around the gallery and watch standard, let us characterize them, "static" areas of the gallery. The target of the proposed work, where the guards are replaced by cameras, is to minimize the number of the cameras, that have limited ability to alternate their position; while the coverage of the available space, time dependent, should be dynamically maximized as well, so as to observe the evolving density of the crowded areas adequately.

An optimization technique to determine the optimal number of cameras needed to cover sufficiently the examined environment, given the specific application of crowd analysis, is inspired specifically by the intelligent behavior of the bumblebees during their foraging process. Recently, researchers have experimentally discovered that bumblebees learn to fly along the shortest route between flowers even if the flowers are revealed in a different order (Lihoreau et al., 2010). Besides, in this paper a second bio-inspired solution based on the spiders' way of weaving their webs was selected, in order to outline the topology of the vision sensors. Craig (1987) indicates in his entomologist study that web-building spiders rely on the critical interplay between web structure and the biomechanical properties of their silks to successfully capture their prey. The energetic cost of web construction is considered here as analogous to the visual information that allows a complete coverage of the provided space. In a similar fashion, as spiders are inclined to minimize this cost (Peakall and Witt, 1976), the location of the cameras are chosen to maximize the provided information needed to capture most of the working environment and the vast majority of the crowded areas. Web-building spiders consider several attributes to choose the proper location to build their nests and minimize the energetic cost of it, such as the net metabolic increment in webspinning activity, the calorific equivalent of silk, and the web weight (Tanaka, 1989). Furthermore, Cranford et al. (2012) recently prove that many species tend to prefer remote places to set up their webs in order to avoid unwanted visitors that are able to ruin their structures. In an analogous manner, the location of our cameras should be in such a place where neither a human or any object may interfere with their proper operation.

In our approach, a number of parameters need to be set, such as: the number of the i available cameras, the number of the bumblebees ($Bees_{num}$) standing for the initial grid points for those i cameras based on the previous space exploration, ($Problem_{size}$) resulting from spiders' web topology by the number of ($Spiders_{num}$) interior angles corresponding to the space-polygon. In the next step, the proposed algorithm maximizes the coverage of the available space subject to the given number of cameras and optimizes the camera topology given the aforementioned locations. However, the observability of the pedestrians remains the primary criterion of our bio-inspired approach emerging the reallocation of cameras based on the spiders module. Moreover, experimental results confirm that the algorithm delivers promising results where the areas with the maximum crowd density are continuously observed and fully covered. According to the structure of this paper, a brief literature survey regarding latest trends in estimating crowd density can be found in Section 2, the functional capabilities of the camera network are explained in Section 3, while the formulation of the problem is thoroughly described in Section 4. The proposed twofold bio-inspired algorithm is based on the emergent computation of the bumblebee colonies along with the intelligent way used by spiders to wave their webs to result into a modified approach to converge to the optimum final solution as found in Section 5. Section 6 provides an outline of the algorithm's observability comparison between the proposed bio-inspired and other algorithms in static environments as well as a thorough presentation of the simulation results of the aforementioned twofold swarm algorithm when applied to static and/or dynamic indoor environments. Finally, conclusions and future work are drawn in the last Section 7.

2. Crowd density estimation

One of the most important problems in crowd analysis is the density estimation of the crowd under consideration. The correct evaluation of the examined crowd is vital in cases such as emergency evacuation of a building or the detection of a potential threat. There are three different clusters of models of density estimation, as reported in the literature, namely pixel-based, texture-based and object-based ones (Junior et al., 2010).

Pixel-based methods rely on the local features, such as individual pixel analysis obtained through background subtraction models or edge detection, to estimate the number of people in a scene. Since the origin of these features is primitive, this model is mostly used to estimate the density of the crowd rather than the precise counting of the people. Ali and Shah (2007) proposed the use of Lagrangian Particle Dynamics to achieve a sophisticated crowd motion understanding in the form of physically and dynamically expressive segments. In particular, by treating the moving mass as an irregular dynamic system, a flow field is generated,