Human Tracking in Cultural Places using Multi-Agent Systems and Face Recognition

Adel Hassan, Aktham Sawan, Amjad Rattrout, and Muath Sabha

Arab American University, Jenin, Palestine,
a.hassan10@student.aaup.edu,a.sawan@student.aaup.edu,amjad.rattrout@gmail.com
and msabha@gmail.com,
https://www.aaup.edu/

Abstract. Heritage places are considered among the most valuable places to any nation for maintaining their history. Computer Vision (CV) and Multi-Agent Systems (MAS) are used for preserving, studying, and analyzing cultural heritage places. This paper introduces a new technique that combines both CV and MAS to track visitors face in cultural places. The model consists of four layers of MAS architecture. The proposed system shows its ability to tackle the human face tracking problem and its flexibility to solve the problem with different tracking parameters. This paper also describes the ability of the agent-based system to deploy a computer vision system to execute different algorithms that fit in solving the human face recognition and tracking problem. The proposed system can be used in any similar place with real time agent-based human face-tracking system.

Keywords: Cultural Heritage, Multi-Agent systems, Computer Vision, Face Recognition, Human Tracking

1 Introduction

Computer vision is considered the most critical system that handles how can detect, recognize, and track objects in a complex environment such as heritage places. The tracking approach consists of two types; motion-based object tracking and recognition-based object tracking. Tracking consists of a set of factors that is influenced by a variety of factors, such as detectors, sensors and environment.

Multi-agent system suits this complex environment of tracking humans by its concept. Facial recognition for tracking detects, identifies, locates and tracks humans in the scene according to different parameters. These parameters, such as certain faces, human positions, camera position, environment, and surrounding effects, could be affected by the type of computer vision suitable algorithms used in the design. The paper suggests combining computer vision system technique with a multi-agent system to produce a piece of art that maintains any movements in the environment by a recognition-based approach.

The rest of this paper is organized as follows: Section 2 presents related works; Section 3 presents the problem and proposed system; Section 4 presents

the simulation of the proposed system and its results; section 5 is the conclusion and suggested future work.

2 Related Works

To build a model for extraction data from images and videos (series of images) using Computer vision, the system in general consists of two main processes, image processing and pattern recognition. These processes used to create new categories of information which understandable by the system [1].

Computer Vision tries to simulate human eyes in identifying objects and perceiving the environment, but this process involves a lot of challenges depending on the design and software used to do this recognition with limited resources and functionalities in the computer vision system. Pattern recognition is responsible for object identification from images that have been taken by external devices such as cameras and sensors [2]. Object detection and representation is considered as the most crucial process in finding moving entities from image sequences. Object tracking is the next necessary process to identify the coordination of moving objects in images or series of images (Video). Doing this needs robust object tracking algorithms to handle these processes. [3]. Sabha and Abu Daoud suggested an adaptive camera placement system for covering the complete area[11]. Their system could be applied on open cultural places for tracking humans.

There are many well known tracking algorithms and techniques such as Kalman, mean-shift, Camshift and LBP filters that achieve good results [2]. Xi divides object tracking subject into two categories, global or holistic visual representation local visual representation [7]. LBP (Local Binary Patterns) uses local-based feature representation especially in facial image analysis, and includes processes in object tracking methodologies such as face detection, face representation, facial analysis, classification and face recognition [5].

Object tracking methodologies could be put in two categories; recognition-based tracking and motion-based tracking. Recognition-based tracking is a modification of object recognition. The object is recognized in multiple series of images and its coordination of extraction. One of the drawbacks of this method is that the object that can be tracked only if the identification and recognizing process are happened, otherwise it cannot be tractable. Motion-based detection relies on the motion of object even if it has not been verified. This is done by comparing background and foreground images [9, 12]. Techniques that are used to handle the problem of target tracking dependent on the tracking parameters, these parameters are describes as follows:

(i) Trackers: Cameras, sensor or any sensing devices
(ii) Targets: Could be one or more objects and could be in stationary or move situation.
(iii) Environment: Outdoor / Heritage places (dynamic, unstructured environment) or Indoor (controlled, static, structured environment).

For facial tracking, LBP is considered the most robust, fastest and the most accurate method that could be used to recognize and track object in an image or video located in outdoor environment such as heritage places.

Heritage places are considered as complex environments that consist of a lot of objects. Tasks could be divided to do the job with faster and and more efficient way by assigning every task to one agent or multiple agents using multi-agent system MAS [4]. Agent is considered as an entity located in an environment and uses parameters to do a special action based on predefined goal for the entity. The agent could be reactive or proactive agent based on nature of their defined tasks. Multiple agents that can communicate with each others and perceiving the environment and give feedback with more rich information about the environment changes which formulates the MAS [10]. MAS characteristics used to build and design new model into computer vision area by making each agent to do a specific job such as capture images, identifying object and analysing a modality in these images. In addition to that, feeding back these information into such knowledge-base system to deal with extracting features at certain time doing more filtration and classification of object types and movements which cause changes between video frames in perceiving environment. [8] presents cooperative multi-target tracking system by introducing a set of active vision agents called AVAs, to track each object in the environment and reciprocate object information between agents. [13] presents how agents can corporate with each other to assign dedicated resources that fit with limitations in the environment using multi-sensor target tracking. [6] presents a different object tracking methods such as HAAR-like features [14], histograms of oriented gradients (HOG) [3], and local binary patterns (LBP) [15].

3 Proposed System

Proposed system by building a model that deals with three main categories; first is heritage place (Environment) since it consist of many elements and need to be organized in a unique forms to deal with them within the system, second is computer vision (CV) that will be the use to perceiving the environment, third is multi-agent system (MAS) that will be interface that is dealing with an environment based on computer vision to do the required action.

Architecture layer design using MAS techniques for proposed system that defines the policies, procedures, and functionalities between these elements shown in Figure 2.

Complexity of System: The complexity of System: The main characteristics that define the sophistication in proposed system shown in Figure 1 that arise as a result of a lot and huge elements that need to be dealing with, such as:

- Emergent behavior
- Self-organization
- Adaptation

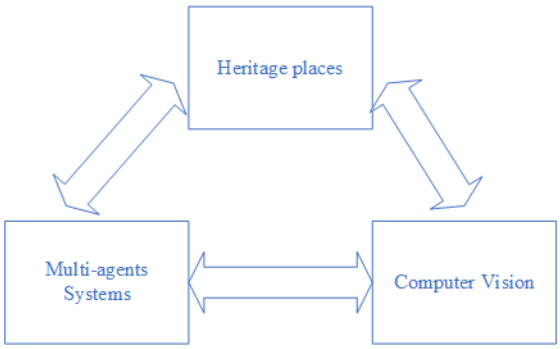

Fig. 1: Structure Model

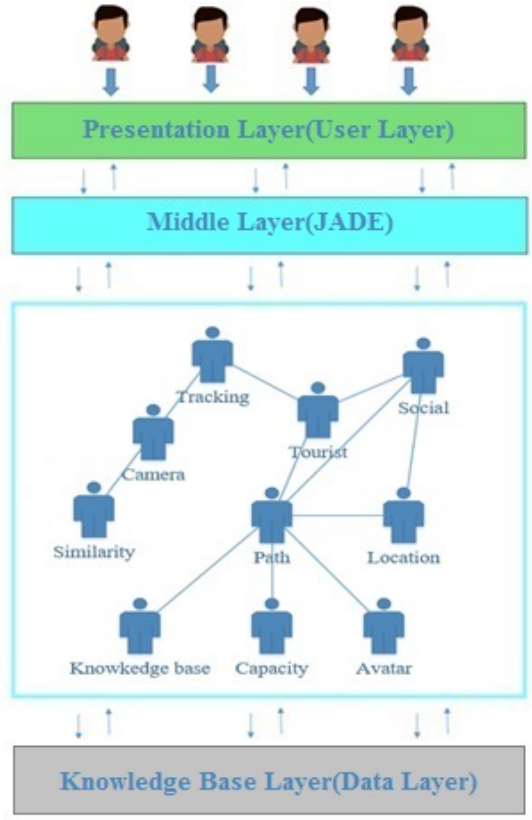

Fig. 2: Architectural Layer Mode

– Co-evolution

Computer vision system and multi-agent system working together in this paper using different agents to simplify complex tasks and operations that are in computer vision to perceiving, filtration, clarification, localization, detection, recognition, and tracking of human actions in a sophisticated and complex environment such as heritage places. Agents characteristics are used to communicate, collaborate, and cooperate with each other's to do the computer vision activities quickly and accurately. MAS architecture proposed in this paper is a layered approach that uses reactive and proactive agents to do different tasks to simplify the complexity of environment and Computer vision.

Computer vision using a facial recognition algorithm using LBP that uses different techniques such as template matching and support vector machines and linear programming to examine the facial expression. The template formed for each class face images, and then a nearest-neighbor classifier is used to match the input image with the closest template that a face image divide into small regions from which LBP histograms are extracted and concatenated into a single, spatially enhanced feature histogram.

Proposed architecture composed of four layers, Presentation Layer, Middle Layer, Operation Layer, and Database (Knowledge Base) Layer. These layers exchange data and information between each other's using different agents that are developed by JADE and communicate each other's reactively or proactively using Foundation for Intelligent Physical Agents (FIPA) Language to capture visitors and their reactions to improve the process of tracking the object target which is human face.

Different type of agents are created with specific role in each layer illustrate as follows:

(i) **The Presentation/User Layer**: This layer represents the access point for a user to the tracking system, and as more accuracy needed in the tracking process results, as more precise tracking parameters, the user should enter the system.
   – Tourist Agent: Agent representing tourists in the heritage area.
   – CAM Agent: Agent who is filming tourist agent.

(ii) **The Middle Layer**: In this layer, decision making takes place, and the success framework relies on the decisions made in this layer. The agent is responsible for deciding and choosing the best tracking methodology based on the user input tracking problem parameters.
   – Tracking Agent: Agent who captures and tracks tourist agents in the heritage area.
   – Path Agent: Agent who tracking paths of tourist agents.
   – Location agent: Agent who is tracking the location of a tourist agent.

(iii) **Operation Layer**: This layer includes all the agents that represent a phase in the tracking process, such as:
   – Similarity agent (uncertain): Agent who tries to find missing or unrecognized agents.

- Capacity Agent: Agent who monitors the capacity of POI.
- Social Agent: Agent, who is trying to spread a general message to more than one agent.

(iv) **Knowledge Layer**: knowledge base consists of a group of tracking methodologies, where each tracking methodology is a subset of following phases, Face detection, Face analysis, face recognition, and face tracking:
- Knowledge base Agent: Agent who has database and information about POI.
- Avatar Agent: Agent who explains POI information to tourist agent.

4 Simulation of Experiment

The simulation of the proposed system was built using JADE on eclipse with Open CV (Open Source Computer Vision Library) to analyze the video streaming. The primary goal of this paper is to create agents by using JADE and to build interactions between these agents using Agent Communication Language ACL Messaging. Experiments were performed on a single node laptop, four cores, Microsoft windows 10, Core i7 processor, 1.8GHz speed, 16 GB memory, and 500 GB HDD.

Interactions and activities between agents shown in Figure 3 describe how they cooperate and collaborate to get the required information about the visitor using face recognition approach in the heritage place. Activities sequence are as follows:

(i) Tourist Agent – Tracking Agent: Inform detection Confirm detection Begin Tracking.
(ii) Camera Agent – Similarity Agent: Get face image Send face image.
(iii) Avatar Agent – Knowledge Agent: Initiate type of avatar - Confirm type of avatar - Avatar begin presentation.
(iv) Capacity Agent – Social Agent: Get available location Send location information.
(v) Path Agent – Location Agent: Get location - Send location information - Confirm location of visitor.

Different data sets and algorithms are used to build the model. Data set containing 226 images and algorithms such as HARR, HOG, LBP are used. The proposed algorithm combines HUG and LBP algorithms together to produce better results compared to other algorithms as shown in Table 1.

Proposed system has been applied to video captured from entrance of heritage place and the results are shown in Figure 5 and Figure 6. It proof that images extracted from the video for face recognition methodology present perfect results that captures human faces and keeps tracking them when they are exploring the environment. Using the proposed system to be applied in a complex environment like heritage places can achieve better results for visitor tracking using the face recognition approach. The tracking agent and camera agent worked cooperatively to recognize human face when enters the environment and kept tracking her

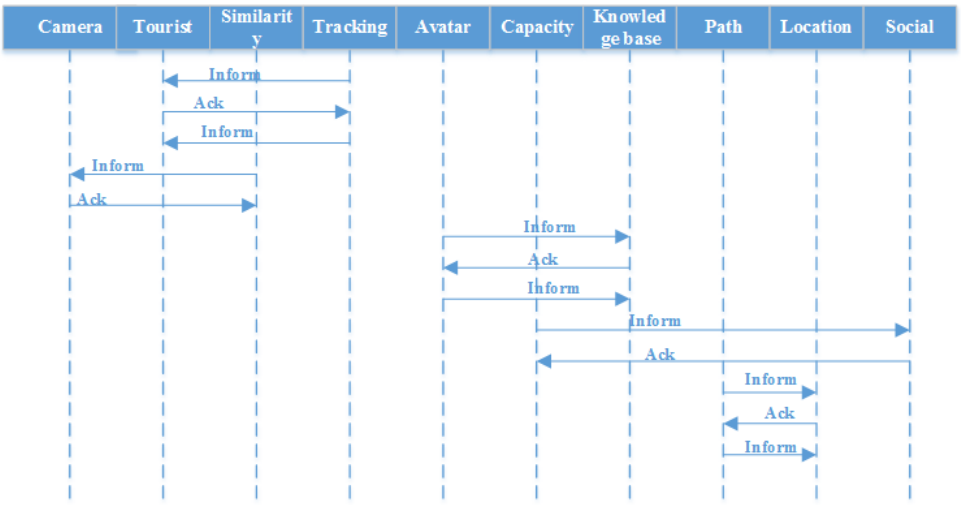

Fig. 3: Sequence Activity Diagram Between Agents.

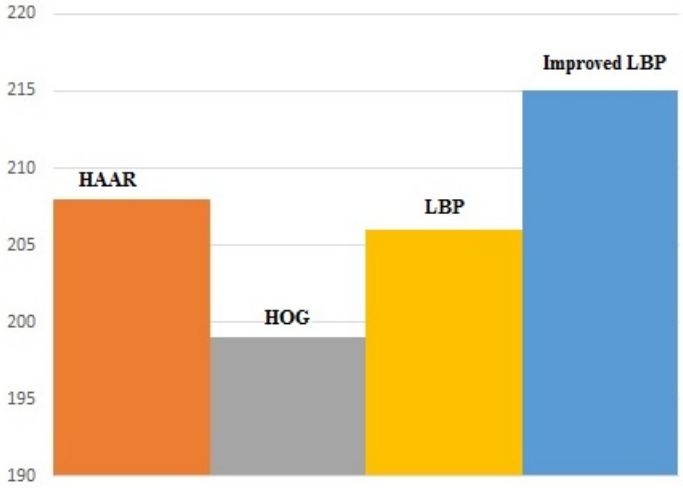

Fig. 4: Face Recognition result

till the human exit from the environment, then similarity agent enabled and terminated all other activated agents and sustain these activities in knowledge base by using knowledge agent.

Table 1: Face Detection Evaluation Comparison Results of 226 Images

Type	Total faces	Haar	HOG	LBP	Imporved
True positive	266	208	199	206	215
False positive	266	18	27	20	11
False Negative	266	55	63	41	33
Accuracy Rate	100%	92%	88%	91%	95%

Fig. 5: Face Recognition source

(i) The tracking agent initiates the activity and informs the tourist agent that a new visitor is detected.
(ii) The tourist agent confirms the activity and sends back to the tracking agent that is an agent detected and then the tracking agent initiates the request to camera image agent to capture the visitor's face.
(iii) A similarity agent then initiated to examine the detection activity to check the nature of the object, then send it to the camera agent to capture the object and return to the similarity agent that the object is human and detect the face based on algorithm face recognition.
(iv) The camera agent captures the face of the visitor and sends it to a similarity agent.
(v) Based on face detection and recognition, the Avatar agent will be introduced to the visitor about the area that available and suggests a tour based on capacity agent information that stored or retrieved by knowledge base agent.
(vi) After that, the path agent and location agent phase initiated to track the visitor in the environment and directed it until the tour is finished.

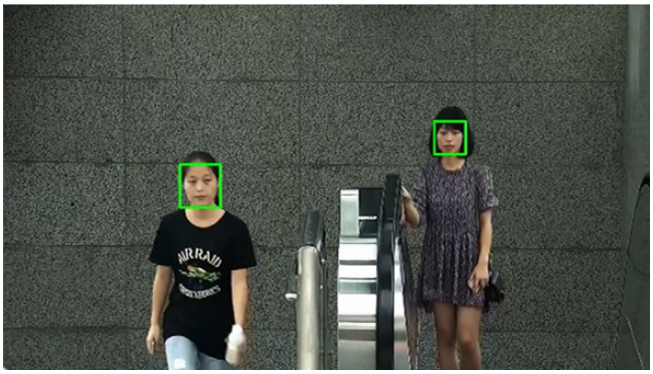

Fig. 6: Face Recognition result

(vii) In meanwhile, the social agent is activated between the visitor agents to explore different locations and suggest different tours based on information feedback form social agents.

5 Conclusion and Future Work

Heritage places are considered as complex environments that need special treatment which MAS can take care off, and could simplify tasks and manage and control different types of interactions.

In this research, the proposed system use multi-agent system and computer vision cooperatively to handle the tracking problem in such a complex environment of heritage places. As Computer Vision part, HUG and LBP algorithms have been used to improve the face recognition capabilities. The MAS used agents to do different tasks in the scene by adopting layered hybrid multi-agent system and benefited form the nature of agents that can work in cooperative, autonomy and self-organizing structures to manage the enormous changes and interactions in the environment because of the number of human in the environment and goals in heritage places. Besides, the proposed system managed to combine two algorithms to enhance the retrieval of the information and the tracking of a human faces in heritage places.

There are some problems in the existing human face recognition and tracking processes. Although the proposed system work efficiently when the face of a human is directed to the camera, but when human face is not in front of camera, or if the human is partially covered by an obstacle, the system is not able to recognize the human.

To overcome such these issues, we recommend to use addition algorithms to filter, classify and identify the human tracking and initiate new agents to tackle these difficulties to recognize human in heritage places.

References

1. Cosido, O., Iglesias, A., Galvez, A., Catuogno, R., Campi, M., Terán, L., Sainz, E.: Hybridization of convergent photogrammetry, computer vision, and artificial intelligence for digital documentation of cultural heritage-a case study: The magdalena palace. In: 2014 International Conference on Cyberworlds. pp. 369–376. IEEE (2014)
2. Coşkun, M., Ünal, S.: Implementation of tracking of a moving object based on camshift approach with a uav. Procedia Technology 22, 556–561 (12 2016)
3. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2005) 2 (06 2005)
4. Dorri, A., Kanhere, S., Jurdak, R.: Multi-agent systems: A survey. IEEE Access pp. 1–1 (04 2018)
5. Huang, d., Shan, C., Ardabilian, M., Chen, L.: Local binary patterns and its application to facial image analysis: A survey. IEEE Transactions on Systems, Man, and Cybernetics, Part C 41, 765–781 (11 2011)
6. Jiang, M., Deng, C., Yu, Y., Shan, J.: Multi-agent deep reinforcement learning for multi-object tracker. IEEE Access PP, 1–1 (02 2019)
7. Li, X., Hu, W., Shen, C., Zhang, Z., Dick, A., Hengel, A.: A survey of appearance models in visual object tracking. ACM Transactions on Intelligent Systems and Technology (TIST) 4 (03 2013)
8. Matsuyama, T., Ukita, N.: Real-time multitarget tracking by a cooperative distributed vision system. Proceedings of the IEEE 90, 1136 – 1150 (08 2002)
9. Murray, D., Basu, A.: Motion tracking with an active camera. Pattern Analysis and Machine Intelligence, IEEE Transactions on 16, 449–459 (06 1994)
10. Russell, Norvig, S.: Artificial intelligence: A modern approach. Prentice Hall, Englewood Cliffs, NJ (01 2010)
11. Sabha, M., Daoud, J.J.A.: Adaptive camera placement for open heritage sites. In: Proceedings of the International Conference on Future Networks and Distributed Systems. ICFNDS '17, Association for Computing Machinery, New York, NY, USA (2017), https://doi.org/10.1145/3102304.3109813
12. Sabha, M., Nasra, I.: Visitor tracking in open cultural places. In: Proceedings of the 4th Hyperheritage International Seminar. HIS.4, Jenin, Palestine (05 2017)
13. Soh, L.K., Tsatsoulis, C.: Reflective negotiating agents for real-time multisensor target tracking. IJCAI International Joint Conference on Artificial Intelligence (04 2003)
14. Viola, P., Jones, M.: Robust real-time face detection. International Journal of Computer Vision 57, 137–154 (05 2004)
15. Wang, X., Han, T.X., Yan, S.: An hog-lbp human detector with partial occlusion handling. In: 2009 IEEE 12th international conference on computer vision. pp. 32–39. IEEE (2009)