Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 125 (2018) 640–646

Procedia Computer Science

www.elsevier.com/locate/procedia

6th International Conference on Smart Computing and Communications, ICSCC 2017, 7-8 December 2017, Kurukshetra, India

A Weight Joint-Based Clustering (WJC) Method for Secure Monitoring System

Y. Jin[a], Abeer Alsadoon[a], P.W.C. Prasad[a]*, A. K. Singh[b], A. Elchouemi[c]

[a]*School of Computing and Mathematics, Charles Sturt University, Sydney, Australia*
[b]*Department of Computer Applications, National Institute of Technology, Haryana, India*
[c]*Walden University, USA*

Abstract

Currently, 3D depth sensors are used in monitoring systems to give accurate results in identifying objects. The existing monitoring systems which use 3D sensors are effective in recognizing object gestures, but do not consider an object's change in location. This paper proposes a new system that uses a 3D camera sensor and weight joint-based in clustering method to identify objects' change in location and to recognize object activities, using a 3D camera to capture the depth value of an object in order to measure its location with the aid of position data of the object's joints. Because the system uses depth value to identify the object's location, it does not require any extra processes for improving the accuracy of clustering where the object is available. The system provides 100% accuracy when recognizing objects' activities has low processing time and is cost effective.

Peer-review under responsibility of the scientific committee of the 6th International Conference on Smart Computing and Communications.

Keyword: Weight joint-based method; Clustering method; Activity recognition

1. Introduction

Monitoring systems are widely used in production management, security, medical care, etc. These systems are able to automatically recognize different objects' behavior. Several of them adopt artificial intelligence techniques to

*Corresponding author
E-mail address: CWithana@studygroup.com

1877-0509 © 2018 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the scientific committee of the 6th International Conference on Smart Computing and Communications
10.1016/j.procs.2017.12.082

analyze data in a way that can reduce processing time [1]. The artificial intelligence technique that analyses data similarity and classifies similar data as a cluster is called data clustering [2]. Also, among the monitoring systems currently in use, one can find 3D monitoring sensors such as Kinect for Windows which are able to detect objects in 3D. It is helping the monitoring system to recognize the object's exact location and its motion [3]. The current monitoring systems cannot detect or track the exact location of objects. To solve this problem, [1] have adopted a sound source sensor to measure an object's location. Kim et al. have provided a Weight Joint-Based Detecting method that uses a 3D sensor to identify the object's joints location [3]. Weight Joint-Based human behavior recognition algorithm is to provide a highly accurate model for recognition of objects' behaviors. This system relies on the depth values of the joint data to identify the object's location. Saad et al. have proposed a method that combines the Dynamic Time Warping Clustering method and the Hidden Markov model to provide a highly accurate system for recognition of objects' activities [2].

This paper is organized as follows: The following section 2 includes a more in-depth research on the current monitoring systems as well as an evaluation of the best existing solutions. In section 3, the paper presents the proposed system. Analysis and implementation of this system are given in section 4, followed by the conclusion of the proposed work.

2. 2. Literature Review

2.1 *Detecting and Tracking Objects*

Some of the existing monitoring systems provide methods for detecting and tracking objects. Region based tracking and contour based tracking methods are only some of the current examples [4]. They ensure the continuous tracking of an object in a detection area, by relying on a three coordinate target position detection method. Kim et al. have provided a Weight Joint-Based Detection method to identify objects' location [3]. They have adopted a 3D camera sensor to measure depth values of an object's joints in order to understand the object's location. An alternative detection method has been provided by Park et al. (2012). They have adopted the Symmetric Double Sided-Two Way Ranging (SDS-TWR) algorithm to measure the object's location. The method involves using four cameras and setting each of them up in a separate node of a rectangular area. Then it calculates the distance between the object's head and each camera in order to identify the object's location. Saad et al. have provided yet another method, using a Support Vector Machine (SVM) to classify whether a moving object is human or not. Firstly, SVM needs to train a device to understand the human blob in an image frame [2]. Zhang et al. provided a method that measures the object's location by using sound source localization technology; that means, it relies on the arrival time of a sound signal from different sound sensors to calculate the objects' location [1].

2.2 *Recognising the Object's Activity*

Kim et al. have provided a method using the sliding window filter method to recognize an object's behavior by comparing the object's behavior template with collected data [3]. Ouivirach et al. have provided a model using Dynamic Time Warping (DTW) and Hidden Markov models (HMM) to cluster an object's behavior pattern [5]. Kim et al. have also developed a method that uses the Scale Invariant Feature Transform (SIFT) dataset in order to recognize theft behavior [6]. This system compares the captured pixels of an image frame with the SIFT dataset to find the products that are contained in the image frame. Huan et al. have proposed a different model that considers the length of time an object is in the detection area in order to recognize any abnormal behavior [7]. This system checks whether the object is located in a sensitive area in an image frame. Another model collects the object's historical living data to recognize the object's abnormal motions usingthe historical data comparison method. It compares the object's motion with historical data. If it is significantly different, the system will recognize the object's motion as abnormal behavior [8, 9].

2.3 Current Best Solution

Kim et al. have provided a system based on a Weight Joint -Based (WJ) algorithm to recognize an object's static features [3]. This system relies on the depth value of the object's joint data to improve the accuracy of the object's location detection. It also provides a good method for the classification of the object's activity pattern, but it is not suitable for tracking changes in the object's location as it just considers differences in the position between the joints. It uses WJ method to detect the object's location. Fig. 1 is the flowchart of this system.

Ouivirach, on the other hand, have provided a clustering human behavior method (CB) which makes use of a 2D camera sensor based solution in order to recognize the object's activities [5]. This system uses the DTW method to identify changes in the object's location. However, because this system cannot detect the object's accurate location on its own, it requires an extra clustering process, the HMM. In order to build the HMM based clustering method, a lot of data is required. An alternative that does not require an extra clustering process involves using a 3D camera sensor to detect the object's location. This sensor enables the system to measure the object's exact location, and that is why it is the sensor proposed in this paper. Rahma et al. provided a solution which relies on region based tracking and a contour based tracking (R & C) method to continuously track an object's activity [4]. This system provides a good method to effectively and continuously track an object's location. The proposed system will use Opencv library in the process of tracking the object.

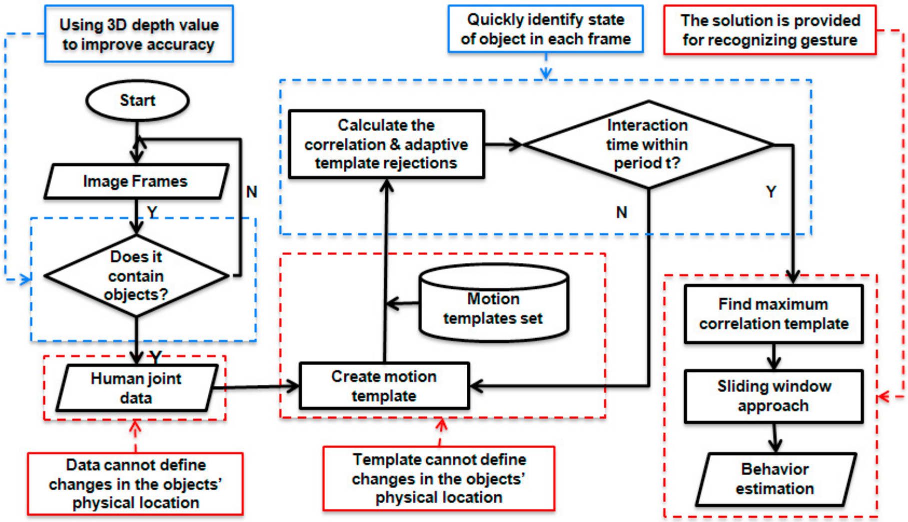

Fig. 1. Current best solution 1 (WJ)

3. Proposed Solution

The proposed system based on the Weight Joint -Based Clustering method (WJC) to recognize an object's activity is illustrated in Fig. 2. This system is composed of three parts, namely Extract object's features, Cluster object's activity pattern, and Recognise object's activity. In the first part (Extract object's features), the system collects the data on the object's joints to identify the object's location. In this part, the system uses a 3D camera sensor to extract 3D version values of the object's location data by using Weight Joint -Based Detection method. In the second part (Cluster object's activity pattern), the proposed system uses the clustering method to identify the object's location. The system then identifies the object's motion pattern based on the changes in the object's location. This is successfully completed through the use of the DTW clustering method. In the third and last part (Recognise object's activity), the system compares the actual object's motion pattern with a predefined motion pattern in order to recognise the object's activity.

3.1 *Detecting and Tracking Object*

This section is a hybrid of Kim et al.'s solution [3] and Ouivirach solution [5]. Its purpose is to improve the accuracy when detecting and tracking an object's location. The proposed system uses the Weight Joint -Based method to measure the object's location together with a 3D camera sensor. Firstly, the object is detected via a background subtraction method. If there is an object, the system will use the Weight Joint -Based Detection method to measure the positions of the object's 20 joints and identifythe object's exact position. The proposed system uses three values to identify a joint's position which include X: the width value of the joint in an image frame, Y: its height value in the image frame, and Z: the depth value in the image frame, i.e. the distance between the joint and the camera sensor. The system then uses region based and contour based tracking methods to continuously track the object's location.

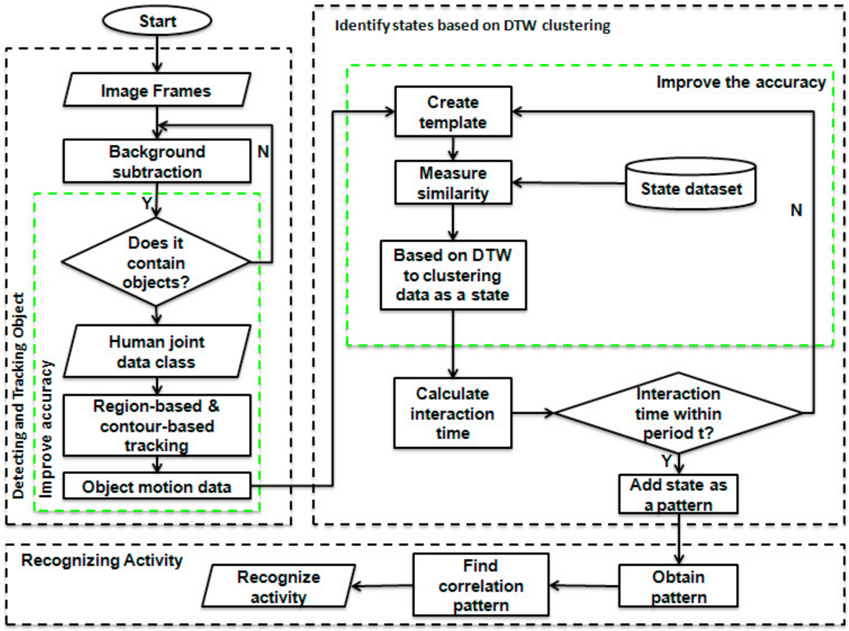

Fig. 2. Proposed WJC

3.2 *Clustering Object's Motion Pattern*

The system adopted by Ouivirach et al. (clustering human behavior method (CB)) [5], clusters data on an object' joints in order to identify changes in the object's location. The proposed system uses the Dynamic Time Warping clustering method to cluster an object's location. To identify a pattern in the object's change in location, the system requires first that the data on an object's joints be put into a cluster. Then, the system checks changes in the pattern of cluster numbers in order to recognise an object's activity pattern. In this case, 6 clusters have been created and each one represents a special part of the detection area. This system relies on the object's motion pattern to select the exact position of each cluster. The clusters are named 1, 2, 3, 4, 5 and 6, and their locations are available in Appendix 3. The proposed system compares the data on the object's joints provided by the 6 clusters in order to calculate the difference between the object's joints data and the 6 clusters. After that, the system will find the most similar clusters to represent the object's location in an image frame. By continuously clustering the data on the object's joints, the system will obtain the pattern of change in the object's location. This pattern data is initially unstructured and it just represents a sequence of the object's location in a continuum of image frames. For this system to work, this data needs to be optimised to a simplified and structured pattern. Based on how long the object stayed in each location, the system manages to optimise the unstructured activity pattern. The system confirms the object's length of stay in a location that is longer than a predefined period, henceforth called a threshold value. The proposed system counts the number of times a cluster number encounters an unstructured activity pattern and calls it number k. This represents the object's length of stay in a particular location. If k is larger than the threshold value, that is, if the duration in that location is longer than the predefined period, the proposed system will confirm the cluster number as an optimised activity pattern. If k is smaller than the threshold value, if the duration in the location is shorter than the predefined period, the system will reject the clusternumber and it will not go into the optimised activity pattern. This process enables the system to ignore errors which might occur for instance when the object is located at the border of two clusters. Algorithm 2 shows the process steps of clustering data on an object's location. The table in Appendix 1 shows the result of clustering the object's activity pattern.

3.3 Recognising Object Activity

In this part, the proposed system will compare the optimised activity pattern with the object's predefined activity patterns. If an optimised activity pattern matches an activity pattern A that already exists in the predefined activity patterns, the system will recognise the object's activity as activity A. If the system cannot match it to any activity pattern from the predefined ones, the system will recognise the optimised activity as unknown or undefined. Table 1 shows the result of recognising an object's activity.

Table 1. Result of recognizing an object's activity

	Recognised Pattern	Recognised Activity Name
1	3-2-1	Go Out
2	1-2-3	Come In
3	5-6	Come Down
4	3-2-1-4	Go to Bed
5	3-4	Go to Bed
6	4-1-2-3	Go to the Desk
7	4-3	Go to the Desk
8	3	Sit
9	Others	Undefined

If the pattern has 2 or more clusters, it does not classify as overstaying. This is because overstaying means that the object's location does not change; hence, the motion pattern would just contain 1 cluster. If an object's activity pattern contains just 1 cluster and this pattern repeats itself 4 times, the system will recognise the object’s activity as overstaying at the location. In this paper, the proposed system has predefined the number 4 for the length of time an object can stay in the cluster. Therefore, if the activity pattern repeats itself 4 times, that means the object is overstaying.

4. Results and Discussion

The proposed system has been implemented using a Python 2.7 project. It imports Opencv library to solve the computer vision process. The Opencv library method is mainly used to detect an object's joint data and to track objects. This project recorded each image frame as an image file at a resolution of 640 * 480 and 30 frames per second. A Kinect for Windows 3D depth sensor was used to capture image frames in depth. The data related to locations was collected as the first step of this work.The proposed system has 6 predefined activities,each related to one or more predefined activity patterns. The system relies on recognising an object's activity pattern in order to identify the object's activity. If its activity pattern does not match a predefined one, the system will recognise the object's activity as being undefined.

In this paper, the CB based system and the proposed system have been implemented. There are three possible experiments to compare these systems. The first and second experiments adopted the proposed WJC method to recognise the object's activity patterns. In order to optimise the object's activity pattern, the first experiment used 6 as the threshold value, whereas the second experiment adopted 10.The last experiment implemented a CB based system without the Hidden Markov Model (HMM) in order to improve the accuracy of the clustering process. It set the threshold value at 6. Evaluation of the system was based on its accuracy when recognising the object's behaviour and on the time it took to recognise the object's activity. Fig. 3 and 4 show the results of these three experiments.

As can be seen in Fig. 3, a comparison between the WJC based and the CB based systems shows that the WJC based system provides significantly higher accuracy when recognising the object's activity than the CB based system. When the threshold is 6, the WJC based system provides 100% accuracy when clustering the object's activity patterns. This result proves that the WJC based system is better capable of measuring the object's exact location than the CB based system. Also, because the CB based system cannot identify the exact location of an object, it cannot recognise the object's exact motion patterns either.

Fig. 3 also shows that in terms of recognition of "Sit" and "Come Down", the three experiments results are similar. They all provide 100% accuracy when recognising an object's activity. This is because the 2D values - X and Y - of cluster 3, cluster 5 and cluster 6 are easily separated from other clusters; therefore, depth value Z is not important in order to recognise the “Sit” and “Come down” activities.

By examining the recognition of the activities "Go to Bed", "Go to Desk" and "Undefined" in Fig 3, it can be seen that the WJC based experiments provide higher accuracy in terms of recognising activity patterns. The WJC based experiments are using 3D values to measure the object's location. Hence, the system can identify the exact location of the object and classify the object's correct motion pattern.

As for recognising the activities "Come In" and "Go Out", Fig. 3 shows that the WJC based experiments provide higher accuracy when recognising these activity patterns than the CB based experiment. The reason why the CB based experiment provides such poor accuracy is that it cannot identify cluster 2; therefore, the system is unable to recognise the object's exact activity patterns. Comparing the WJC based experiments, when the system uses 6 as the threshold value, it provides a higher accuracy than when that value is 10. The WJC based experiment, for which thethreshold value was set as 10, was unable to provide 100% accuracy when identifying the object's activity pattern. This is because the common duration of cluster 2 is between 8 and 15.

Fig. 4 shows that the processing times of the three experiments do not have significant differences. The CB based system is faster than the WJC based system, but the difference in processing time is often less than 0.005 seconds. Considering that an object's common activity pattern is around 10 seconds, such small difference is acceptable.

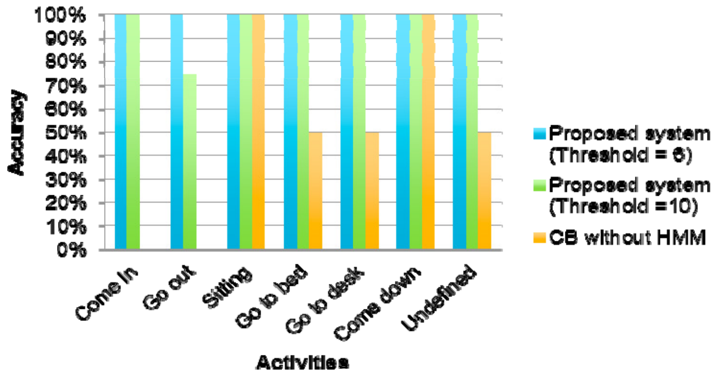

Fig. 3. Result of Accuracy

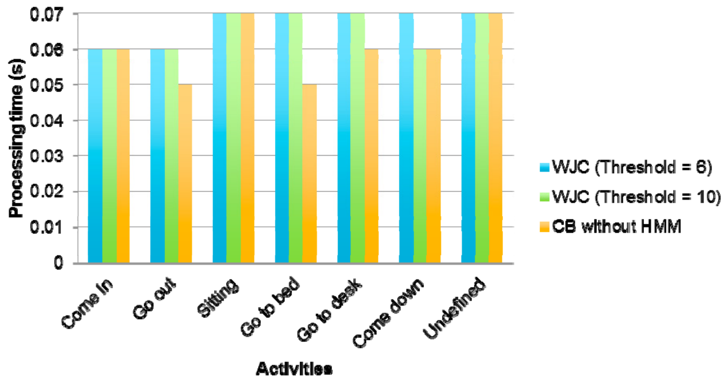

Fig. 4. Result of Processing Time

5. Conclusion

In this paper, a new model called Weight Joint -Based Clustering model has been provided in order to solve the accuracyand processing time problems of navigating the object. This model uses a 3D camera sensor to detect and track objects. The system uses not only depth value to weight objects' joints' data in order to identify the object's location, but also a clustering method to identify changes in the pattern of the its location. Finally, it relies on the object's motion pattern to recognise the object’s activities. The proposed system provides high accuracy at low cost when recognising an object’s activity. It provides a method to improve the utility of the current 3D monitoring system which ensures the system can recognize changes in the object’s physical location. The new model reduces the data requirements for clustering the object’s features, and as far as the processing part is concerned, it also provides satisfactory results in its implementation.

REFERENCES

1. Zhang, Zhao, and Guo, (2014) "Intelligent Video Surveillance System Based on Sound Source Localization Technology", In Applied Mechanics and Materials. 511 , 530-535.
2. Saad, and Hussain, (2010) "A Design and development of intelligent anomalous behavior and event detection system, 6th International Colloquium in Signal Processing and Its Applications (CSPA). 1-5.
3. Kim, Lee, Kim, Lee, Lee, Ju, and Myung ((2016) "Weighted joint-based human behavior recognition algorithm using only depth information for low-cost intelligent video-surveillance system", Expert Systems with Applications. 45, 131-141.
4. Rahma, and Jamil (2014) "The Proposed Design of the Monitoring System for Security Breaches of Buildings Based on Behavioral Tracking", International Journal of Computer Science and Network Security (IJCSNS). 14, 369-373 .
5. Ouivirach, and Dailey (2010) "Clustering human behaviors with dynamic time warping and hidden Markov models for a video surveillance system", In International Conference on Electrical Engineering/Electronics Computer Telecommunications and Information Technology. 884-888.
6. Kim, Yeom, Joo, and Park (2010) "Intelligent unmanned anti-theft system using network camera, International Journal of Control", Automation and Systems. 8, 967-974.
7. Huan, Tang, Wang, and Chen (2011), "Abnormal Motion Detection for Intelligent Video Surveillance", In Applied Mechanics and Materials. 58, 2290-2295.
8. Huang, Tian, Wu, and Zhou (2014), "A method of abnormal habits recognition in intelligent space", Engineering Applications of Artificial Intelligence. 29, 125-133.
9. Deng and Cheng (2012) "Research and Design of Intelligent Video Surveillance System", International Journal of Advancements in Computing Technology. 4(11), 378-388.