A Survey on Current Video Surveillance Systems

Sanket Walke[1] Ashwin Kulkarni[2] Omkar Langhe[3] Shubham Marne[4] Prof. Sandip Mahajan[5]
[1,2,3,4]Student [5]Assistant Professor
[1,2,3,4,5]Department of Computer Engineering
[1,2,3,4,5]Sinhgad Institute of Technology and Science, University of Pune, Pune, Maharashtra, India

Abstract— In today's world, surveillance plays an important role in securing property from various suspicious and intrusion activities. In such surveillance, technology plays an important role by making it easy for humans to gather information from surroundings at one place by using various electronic devices such as cameras, sensors and various monitoring devices. But search surveillance involves human interceptions to manage and control the system and makes the system dependable. Also to monitor and record the activity it is must to capture every frame and instance of movement. In The following implementation we use Surveillance Camera and Multi-sensor Data Fusion for improving detection, tracking and identification which also improves situation assessment and awareness. Also it provides spatial and temporal coverage, shorter response time and reduces communication and computing.

Key words: Human Detection, Tracking, Monitoring, Sensor, Multi-Sensor, Data Fusion

I. INTRODUCTION

Manually handled surveillance system addresses only what is being ordered to see and record, which sometimes loses the focus on the humans in the busy environments. The technical problems such as human detection, movement analysis, and pattern recognition involve very major fields of computer science are artificial intelligence, image processing, machine learning and much more. Video surveillance development is carried all over the world and has been a very trending topic of discussion.

As progressing in the 21st century we need systems which will do the most of tasks with any help of humans and make decisions on the basis analysis and patterns. There is certainly need of automatic systems which will reduce lot of human work. Deploying cameras and sensors is fine to the surveillance but to find the human resource to observe the result is much more expensive. Surveillance cameras are deployed in stores, banks, hospitals and parking lots and it is used only as forensic evidence thus losing its primary benefit as real-time monitoring system. The requirement of the surveillance system is 24x7 monitoring and to alert security officers for suspicious activity in progress or to a suspicious individual loitering in the parking lot, to prevent any crime from taking place.

Following literature review gives us an abstract idea of what type of research and development has taken in place in recent years and which technologies are used to make the research in action. Also to overcome the issues in the review we have proposed a system which we hope to overcome these identified problems and provide a better solution on video surveillance system.

II. LITERATURE REVIEW

A. Multiple Sensor Fusion and Classification for Moving Object Detection and Tracking

In this paper, the problem of intelligent vehicle perception have been reviewed. This paper focuses on the DATMO (Detection and Tracking of Moving Objects) component of the perception task. It has been proposed that the use of classification information as a key element of a composite object representation, where not only kinetic information but appearance information plays an important role in the detection, classification and tracking of moving object of interest.

By performing multi-sensor fusion at detection level, the impact of composite object description is analyzed. Three sensors have been used to define, develop, test and evaluate the fusion approach: LIDAR, radar and camera. Moreover, complete perception solution was evaluated using on-line and offline data from a real vehicle of the interactive European project.

Combining and Integrating class information at the detection level, allowed the fusion to improve the detection by considering an evidence distribution over the different class hypotheses of the detected objects. Because of this, reduction of the number of false detections and false classification at early levels of the DATMO component is improved.

1) Advantages

1) The tracking stage benefits from the reduction of mis-detection and from the more accurate classification information to accelerate the tracking process.
2) 3D based representations (e.g., voxel segments) can provide information about the geometry and class of the objects of interest around the ego-vehicle, and the common obstacle that generate false classifications (e.g., trees, bushes, poles).
3) An advantage of fusion approach at the detection level is that the description of the objects can be enhanced by adding knowledge from different sensor sources. For example, LIDAR data can give a good estimation of the distance to the object and its visible size. In addition, classification information, usually obtained from camera, allows to make assumptions about the detected objects. An early enrichment of objects' description could allow the reduction of the number of false detections and integrate classification as a key element of the perception output rather than only an add-on.

2) Disadvantages

1) In urban areas, vehicle detection and classification is still high, considering the increased number of moving obstacles and the confused environment.

B. Drone Assisted Disaster Management: Finding Victims via Infrared Cameras and LIDAR Sensor Fusion

In this paper, overall architecture for drone assisted disaster management and propose the suitable drone hardware with sensors for practical rescue operation has been discussed. Testing has been implemented on system in the lighted and limited lighting conditions, and sensors outputs virtualized by ROS provide the global and local maps of surroundings unknown environments. It has been observed and identified that laser scanner sensor and depth camera are insensitive to illumination change and thus can be fused together to offer meaningful information at natural disaster sites. Successful usage of these sensors enables rescuers to detect significant landmarks such as doors or boundary walls and find survivors from the disaster at the earliest time.

Among various types of robots, serial vehicles is called as drone. Drones are used to monitor and access disaster sites. Most important feature is compatibility and availability of searching survivors. Powerful drone requires multicore processors, low power NVIDIA GPU, 2GB RAM, Ethernet, real sense camera, accelerometer, position tracking. Drones are using LIDAR for sensor fusion. Sensor fusion is used in serial machines to gather the information and send it to the central system for decision making.

1) Advantages

1) Drones systems can be used in natural disaster environment to capture the various scenarios to take proper actions.
2) Drones are Cost effective, compact, and easy to operate in cluttered and confused environment, thus become affordable candidates for disaster assistance.
3) Drones can be operated in disaster locations like earthquakes, collapse buildings, etc., to explore the damaged location where human rescuers find challenging and difficult to explore.

2) Limitations

1) Drones are not autonomous.
2) Drones are not able to perform operations like localization and trajectory planning.
3) Also not able to implement path finding algorithm.

C. Vision-Based Target Detection and Localization via a Team of Cooperative UAV and UGVs

This paper tells us about target tracking and detection in real-time. In this paper, a new vision-based target detection and localization system is presented with the help of team of UAVs (Unmanned Aerial Vehicle) and UGVs (Unmanned Ground Vehicles). Since UAVs has lower resolution and broader detection range and UGVs have higher resolution and fidelity they are used together for better tracking and locating crowds.

A-Rule-of-thumb localization is method by a UGV is presented which gives us an estimated geographic locations of detected individuals. To detect individuals from UGV camera a histogram of oriented gradients (HOG) with a prominent performance at intermediate to higher resolution images is applied.

In order to detect the moving target (i.e. crowd) they have applied to different methods for UAVs and UGVs based on their physical features such as resolution and coverage. For this purpose they use (1) motion detection algorithm based on the optical flow for the UAV with slower background motion and (2) a human detection algorithm based on the HOG for UGVs with a higher resolution and portrait image of targets.

The First algorithm tells us about the extracting of several key points which have good features to be tracked in subsequent further slides. These key points are assigned by GoodFeaturesToTrack(GFTT) mothed which are invariable to rotation and spatial movements. GFTT sorts the pixels in autocorrelation matrix and chooses number of pixels as key points. In order to match the key points across successive frames based on the displacement they applied spatial optical flow concept and solved the tracking problem using pyramidal Lucas-Kanade (PLK) algorithm.

The Second algorithm tells us about the HOG based detection module for the UGVs. It contains some parameters such as window size and block stride. HOG can provide high accuracy human detection, though at a relatively high computational cost. After retrieving the HOG descriptor the algorithm applies a linear Support Vector Machine for human classification where it coefficients are extracted from a OpenCV trained classifier for detecting people. Once the target is detected, the algorithm sets a bounding box around each individual and a scale computation is applied on them to extract the image location of their feet. These coordinates are then passed into a function for computing target's real-world location.

1) TARGET Localization: Proposed Method [3]

The proposed Method is modularized in following modules:

1) Landmark Assignment by UAV

a) Vision-Based Landmark Detection:

Two Major Properties are required for the landmarks

1) To be detectable robustly in different illuminations and distances.
2) To be uniquely identifiable.

2) Moving Landmarks and Target Differentiation:

UGVs may move in UAVs detection range, and this may result in being detected as moving blobs by the UAV detection algorithm. In order to overcome this problem they introduced an assignment problem and solved it at every tracking interval for the best assignment of motion detected blobs to their corresponding color-detected landmarks. Hence, Moving landmarks can be discriminated from the moving targets.

1) Finding Real World Location by UAVs: [3]

Real-time localization algorithm we proposed in this paper implements the following generic steps:

a) Extracting landmark's real-world geographic locations (suing GPS).
b) Finding the image location of landmarks
c) Extracting the image location of targets (using the motion detection and solving the assignment problem).
d) Estimating the elements of transformation matrix between the two planes, using a system of equations.
e) Covering the image locations of the targets to their real-world geographic location, using the computed transformation-matrix.

3) Finding Real-World Location by UGV [3]

In this module, they have used triangle similarity for a pinhole camera model to get a good estimation of target's distance to the UGV camera.

4) Advantages

1) A very effective way of monitoring in suitable environment and accuracy is also maintained.
2) Very useful in tracking the geographical location of target
3) Individual or group.
4) No need of fixed deployment and ready to use anywhere anytime.

5) Disadvantages

1) The UGV's and UAV's are highly vulnerable and can be demolished in such operations.
2) The UGV's and UAV's are not optimum in rough environment and can produce false data.
3) Also the power factor is one major issue and cannot be used for long time.
4) Not an automatic system and UGV's and UAV's should be controlled manually.

III. RESEARCH METHODOLOGIES

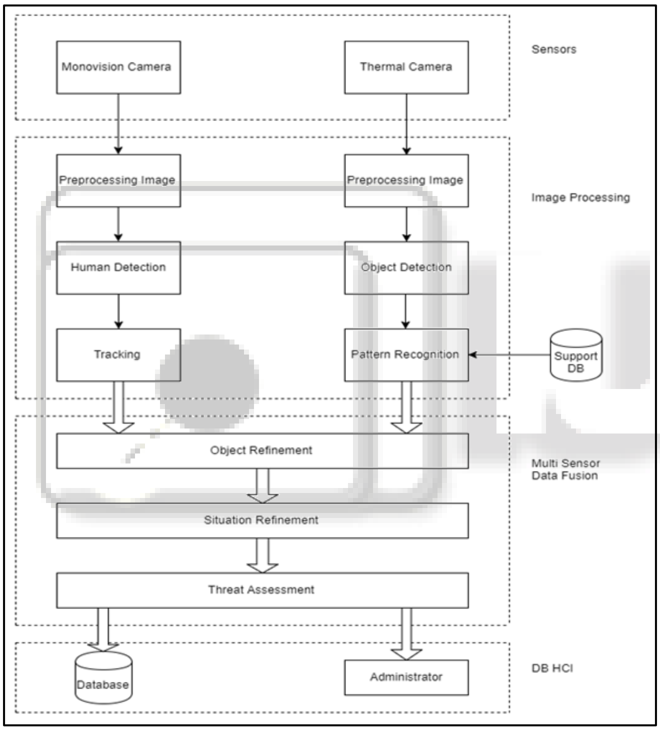

Fig. 1: Proposed System Architecture

Above system architecture is divided into 4 layers and each layer depicts the various functionalities carried out in the detection and monitoring and data-fusion process. This system will be a wireless system which will be much easier to install and can be used in much larger area. Let us understand how the how the process of above architecture is carried out.

A. Camera and Sensors Layer

Initially the camera and sensors work together to capture the ongoing activity. Once the activity is captured then it requires some intermediate to collect that data and pass it forward, this is taken care by interfacing device and passes the data to central system.

B. Image Processing Layer

Now the Image Processing block has got the data, the data is passed to it respective process. The normal frames and thermal frames are processed separately using threading. Normal frame is processed in 3 levels which are preprocessing image, human detection and tracking.

1) Preprocessing Image

In here the raw data collected by upper layer is converted into an understandable format on which further operations can be done.

2) Human Detection

In this block, for the detection of human algorithms like HOG will be applied and that information will be stored in the database.

3) Tracking

If a human is detected then this block will carry the tracking operation in which the movements and position of human will be monitored.

Same goes for thermal frames but here the processes are different such as preprocessing, object detection and pattern recognition.

4) Preprocessing Image

In here the raw data collected by upper layer is converted into an understandable format on which further operations can be done.

5) Object Detection

This is a very crucial task where we try detect or try to identify what kind of object that human is carrying. This will allow us to make certain decisions about that individual.

6) Pattern Recognition

Here pattern recognition requires a support database to refer some default provided patterns in order to recognize the object. After that both the data is passed to next level which is Data fusion layer.

C. Multi-Sensor Data Fusion Layer

In data fusion layer, there are also three levels to be followed which includes Object refinement, Situation refinement and threat assessment.

1) Object Refinement

This level combines sensor data to obtain the most reliable and accurate tracking and estimation of an entity's position, velocity, attributes, and identity. Although this level is not considered part of the high-level fusion, entity-tracking is analogous to tracking a phenomenon [6].

2) Situation Refinement

The Situation Refinement level develops a description of current relationships among entities and events in the context of their environment [6].

3) Threat Assessment

This level projects the current situation into the future to draw inferences about enemy threats, friend and foe vulnerabilities, and opportunities for operations [6].

D. Database and HCI Layer

The Human Computer Interaction (HCI) block provides an interface to allow a human to interact with the fusion system. The Database block provides management of data for fusion (sensor data, environmental information, models, estimations, etc.).

IV. Conclusion

So far, we have seen how the traditional systems works, what drawbacks did they had and how Proposed System Architecture offers to overcome the current situation. Also, we came across the functioning of proposed system architecture from which we came to know how each layer of the system contributes and works and how it actually performs the given task.

References

[1] Ricardo Omar,Chavez-Garcia and Olivier Aycard, “Multiple Sensor Fusion And Classification For Moving Object Detection And Tracking,” IEEE Transactions On Intelligent transportation systems, Vol. 17, No. 2, February 2016.
[2] Seoungjun Lee, Dongsoo Har and Dongsuk Kum, “Drone-Assisted Disaster Management: Finding Victims via Infrared Camera and Lidar Sensor Fusion,” IEEE 3rd Asia-Pacific World Congress on Computer Science and Engineering, 2016.
[3] Sara Minaeian, Jian Liu, and Young-Jun Son,” Vision-Based Target Detection and Localization via a Team of Cooperative UAV and UGVs,” IEEE transactions on systems, man, and cybernetics: systems, vol. 46, no. 7, July 2016.
[4] Q. Baig, “Multisensor data fusion for detection and tracking of moving objects fromadynamic autonomous vehicle,” Ph.D.dissertation, Lab.Inf. Grenoble, Univ. Grenoble, Grenoble, France, 2012.
[5] M. Paul, S. M. Haque, and S. Chakraborty, “Human detection in surveillance videos and its applications—A review,” EURASIP J. Adv. Signal Process., vol. 176, no. 1, pp. 1–16, 2013.
[6] Mr. Subrata Das, “Internet of Things - critical roles of data fusion, analytics, and intelligent agents”, published December11,2015[https://www.linkedin.com/pulse/internet-things-critical-roles-data-fusion-analytics-subrata-das Accessed:5November,2017]

