Image Tracking in Live Video Surveillance

Kokila.M [1], Nishanthi.H[2], Geethapriya.J[3]

[1,2,3] Department of Information Technology, Anand Institute of Higher Technology
Kazhipattur, Chennai, Tamilnadu, India

Abstract

Image Tracking Using Blob Analysis presents a novel framework for detecting flat and non-flat abandoned objects at a public place and determines which one remains stationary. The existing system doesn't detect the flat objects and also it is subjected to some false alarms. But in this prototype abandoned objects are detected by matching a reference and a target video sequence. The object is analyzed to ensure that it does not pose a threat to the security of the location. Four simple but effective ideas are proposed to achieve the objective: an inter-sequence geometric alignment, an intra-sequence geometric alignment, a local appearance, a temporal filtering and additionally we use blob detection algorithm to identify the flat, specific objects based on object attributes and interest points of the image.

Keywords:*Abandoned Flat and Non-flat objects, Reference video, Target video, Video matching, BLOB Analysis.*

1. Introduction

In recent years, visual surveillance by intelligent cameras has attracted increasing interest from homeland security, law, Enforcement and military agencies. The Detection of suspicious (dangerous) items is one of the most important applications. These items can be grouped into two main classes, dynamic suspicious behaviors (e.g., a person attempting to attack others) and static dangerous objects (e.g., luggage or bomb abandoned in public places). The scope of this paper falls into the latter category. Specifically, we investigate how to detect flat and nonflat static objects in a scene using a moving camera. Since these objects may have arbitrary shape, color or texture, state-of-the-art category specific (e.g., face/car/human) object detection technology, which usually learns one or more specific classifiers based upon a large set of similar training images, cannot be applied to our scenario. To deal with this detection problem,we propose a simple but effective framework based upon matching a reference and a target video sequence. The reference video is taken by a moving camera when there is no suspicious object in the scene, and the target video is taken by a second camera following a similar trajectory, and observing the same scene where suspicious objects may have been abandoned in the mean time. The objective is to find these suspicious objects. We will fulfil it by matching and comparing the target and reference sequences.

2. Method for detecting the abandoned object

Almost all current methods for static suspicious object detection are aimed at finding abandoned objects which is nonflat using a static camera in a public place, e.g., commercial center, metro station or airport hall. Spengler and Schiele propose a tracking/surveillance system to automatically detect abandoned objects and draw the operator's attention to such events [9]. It consists of two major parts: A Bayesian multiperson tracker that explains as much of the scene as possible, and a blob-based object detection system that identifies abandoned objects using the unexplained image parts. If a potentially abandoned object is detected, the operator is notified, and the system provides the operator the appropriate key frames for interpreting the incident. Porikli *et al.* propose to use two foreground and two background models [7] for abandoned object detection. First, the long- and short-term backgrounds are constructed separately. Thereafter, two foreground models are obtained based upon the two background models. The abandoned object can be detected by four hypotheses based upon the two foreground and two background models. Guler and Farrow propose to use a background-subtraction based tracker and mark the projection of the center of mass of each object on the ground[3].The tracker identifies object segmentations and qualifies them for possibility of a "drop-off" event. The stationary object detector running in parallel with the tracker quickly identifies potential stationary foreground objects by watching for pixel regions that consistently deviate from the background for a set duration of time. The stationary objects detected are correlated with drop-off events and the distance of the owner from the object is used to determine warnings and alerts for each camera view. The abandoned objects are correlated within multiple camera views using location information, and a time-weighted voting scheme between the camera views is used to issue the final alarms and eliminate the effects of view dependencies. Smith *et al.* propose to use a two-tiered approach [8]. The first step is to track objects in the scene using a trans-dimensional MCMC tracking model suited for generic blob tracking tasks. The tracker uses a single camera view, and it does not differentiate between people and luggage. The problem of determining whether a luggage item is left unattended is solved by analyzing the output of the tracking system in a detection process.

2.1 Overall System Architecture

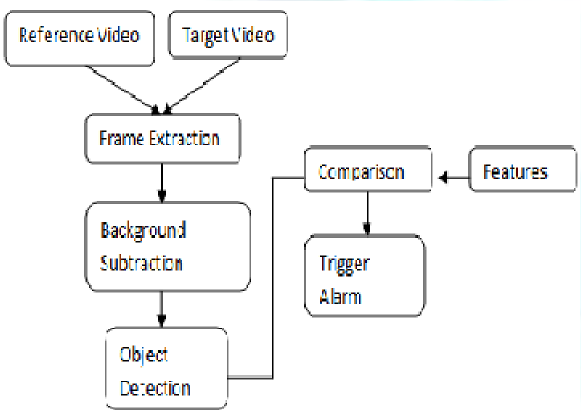

Fig.1 Overall Architecture

Given a reference video and a target video which are taken by a camera following similar trajectories, GPS is used to roughly align the two videos to reduce computational complexity (the GPS information is obtained every second, which corresponds to roughly every 10 m). Fig. 1.1 shows three corresponding frame pairs aligned by GPS, where the top and bottom rows are from R and T respectively (note the suspicious object in the target sequence).

Fig. 1.1 Example of GPS-aligned frame pairs. Top row: frames from reference video. Bottom row: frames from target video (an abandoned object appears on the right before the bushes).

Generally, it is hard to find the suspicious object if we only compare the GPS-aligned frames, which potentially have large viewpoint variation. This is because GPS alignment can only guarantee that the corresponding intersequence frame pair is taken approximately at the same geographical location, but cannot guarantee that the camera has the same view angle for R and T. In addition, due to speed variation between and the different position of the vehicle, alignment using only GPS information may lead to frame pairs separated by as much as 2 m in 3-D real world. Therefore, a fine geometric alignment is necessary. A feature-based alignment method is a better choice than an appearance-based one when the illumination conditions and is different. We propose to use 2-D homographies [2], [4] for fine alignment. The reasons are that homographies can align two images by registering their dominant planes, and that any nonflat objects (including suspicious and nonsuspicious ones) on the dominant plane are deformed while flat objects remain almost unchanged after alignment. We will show that the deformation caused by homography alignment plays a key role in our detection frame-work (especially when there is an large illumination variation). Therefore, we have two assumptions: the suspicious object is a nonflat 3-D object1 (Specifically, we are more interested in detecting the abandoned objects which has such a height as a suitcase or gift-boxes etc.), and when it is present in the target sequence, it lies on the ground instead of hanging in the sky, being buried underground or covered by other objects. Note that we do not make the assumption that the route (road) must be flat. In fact, the road can consist of a few flat segments.

2.2 Intersequence Geometric Alignment

The SIFT feature descriptor [6] is initially applied to the GPS-aligned frame pairs (we also tried the Harris corner detector, but the result is worse). To reduce the effect of SIFT features of high objects (e.g., trees) on the homography estimation, it is better to apply it only to the image area which corresponds to the ground plane. Therefore, the method proposed in [5] is used to estimate the horizon line passing through the vanishing point of the road. The horizon for straight road can be located at an accuracy of over 96%. For curved road, the vanishing point is detected

ISSN: 2320 - 8791
www.ijreat.org

as the one associated with the main straight part of the road. The performance of vanishing point detection is reported in [5] and the supplemental results for general road images can be found in the section of “General Road Detection from a Single Image” of our project. In addition, we emphasize that the homography estimation is insensitive to the accuracy of horizon detection: a detection error of 15 pixels higher or lower than the actual horizon has very little effect on the homography estimation. Only the SIFT features below the vanishing point are viewed as valid. Coarse correspondences between the valid SIFT features of and are constructed.

Specifically, we first compute a 128-dimensional SIFT descriptor for each key point of the reference and target frames (the extraction process just follows Lowe’s method). For each descriptor in the reference frame, we search its nearest neighbor in the target frame. Similarly, for each descriptor in the target frame, we search its nearest neighbor in the reference frame. If the two nearest neighbors are consistent, we view them as a match.

2.3 Intrasequence Geometric Alignment

The procedure for intrasequence geometric alignment is similar to that for intersequence alignment. The difference is that both the reference (the frame to be warped) and target frames are from the same video this time. Generally, the choice of depends upon the speed of the moving camera. If the camera moves fast, k should be set to a small number, and vice-versa. We take k=5 for our experiments, with the platform moving at an approximate speed of 30 km/h and the displacement of the camera between frames being about 10 m. Since the illumination variation between the intrasequence reference and target frames is usually small, the intrasequence alignment generally aligns the dominant planes very well (even when shadow appears in one and disappears in the other, as in the case shown in the top row of fig2)

Fig. 2 Suspicious object areas B (highlighted) based upon the intersequence alignment.

2.4 Temporal Filtering

We use temporal filtering on to get our final detection. Let k be the number of buffer frames used for temporal filtering. We assume that is the current frame and the remaining suspicious object areas in after intersequence and intrasequence. We also stack the homography transformations between any two neighboring frames of buffer.

3. Blob Detection Algorithm

A blob (binary large object) is an area of touching pixels with the same logical state. All pixels in an image that belong to a blob are in a foreground state. All other pixels are in a background state. In a binary image, pixels in the background have values equal to zero while every nonzero pixel is part of a binary object. You can use blob analysis to detect blobs in an image and make selected measurements of those blobs. Blob analysis consists of a series of processing operations and analysis functions that produce information about any 2D shape in an image. Use blob analysis when you are interested in finding blobs whose spatial characteristics satisfy certain criteria. In many applications where computation is time-consuming, you can use blob analysis to eliminate blobs that are of no interest based on their spatial characteristics.

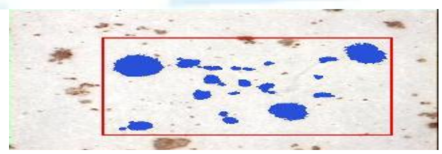

Fig 3 A View on Blob Method.

You can use blob analysis to find statistical information-such as the size of blobs or the number, location, and presence of blob regions. With this information, you can perform many machine vision inspection tasks, such as detecting flaws on silicon wafers, detecting soldering defects on electronic boards, or Web inspection applications such as finding structural defects on wood planks or detecting cracks on plastics sheets. You can also locate objects in motion control applications when there is significant variance in part shape or orientation.

In applications where there is a significant variance in the shape or orientation of an object, blob analysis is a powerful and flexible way to search for the object. You can use a combination of the measurements obtained through blob analysis to define a feature set that uniquely defines the shape of the object.

IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 1, Issue 1, March, 2013
ISSN: 2320 - 8791
www.ijreat.org

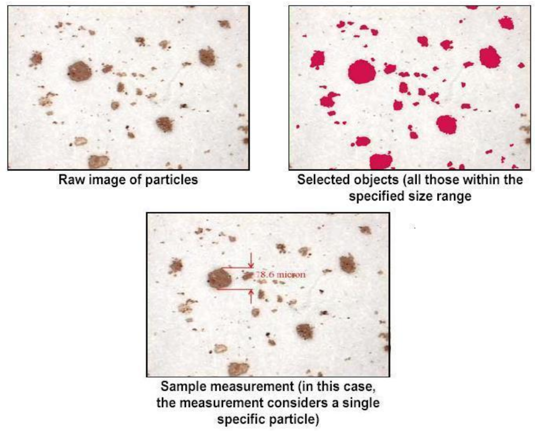

Fig. 4 Count Size with BLOB Analysis.

4. Conclusion

This paper proposes a novel framework for detecting flat and nonflat abandoned objects by a moving camera. Our algorithm finds these objects in the target video by matching it with a reference video that does not contain them. We use four main ideas: the intersequence and intrasequence geometric alignment, the local appearance comparison, and the temporal filtering based upon homography transformation.

We also use BLOB Analysis in addition to the above mentioned four techniques. Our framework is robust to large illumination variation, and can deal with false alarms caused by shadows, rain, and saturated regions on road. It has been validated on fifteen test videos.

5. References

[1] Hui Kong, Jean-Yves Audibert, Jean Ponce, Detecting Abandoned Objects with a Moving Camera
[2] D. Forsyth and J. Ponce, Computer Vision: A Modern Approach. Upper Saddle River, NJ: Prentice-Hall, 2002.
[3] S. Guler and M. K. Farrow, “Abandoned object detection in crowded places,” in Proc. PETS Workshop, 2006, pp. 99–106.
[4] R. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2003.
[5] H. Kong, J.-Y Audibert, and J. Ponce, “Vanishing Point Detection for Road Detection,” in Proc. IEEE Conf. Computer Vision, 2009, pp. 96–103.
[6] D. Lowe, “Distinctive Image Features from Scale-Invariant Key- points,” Int. J. Comput. Vis., vol. 60, no. 2, pp. 91–110, 2004.
[7] F. Porikli, Y. Ivanov, and T. Haga, “Robust abandoned object detection using dual foregrounds,” EURASIP J. Adv. Signal Process., vol. 2008,no. 1, pp. 1–11, 2008
[8] K. Smith, P. Quelhas, and D. Gatica-Perez, “Detecting abandoned lug- gage items in a public space,” in Proc. PETS Workshop, 2006, pp.75–82.
[9] M. Spengler and B. Schiele, “Automatic detection and tracking of abandoned objects,” in Proc. PETS Workshop, 2003, pp. 149–156.
[10] R. Szeliski, “Image alignment and stitching: A tutorial,” Foundation Trend Comput. Graph. Vis., vol. 2, no. 1, 2006.
[11] M. A. Fischler and R. C. Bolles, “Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography,” Commun. ACM, vol. 24, no. 6, pp. 381–395, 1981.