PhD Forum: Illumination-robust Foreground Detection for Multi-camera Occupancy Mapping

Nyan Bo Bo*, Sebastian Gruenwedel*, Peter Van Hese*, Jorge Oswaldo Niño-Castañeda*,
Dirk Van Haerenborgh†, Dimitri Van Cauwelaert*†, Peter Veelaert*† and Wilfried Philips*
*Image Processing and Interpretation, TELIN, Ghent University/IBBT, Belgium
†Vision Systems, University College Ghent, Belgium
Email: Nyan.BoBo@telin.ugent.be

***Abstract*—Foreground detection is an essential preprocessing step for many image processing applications such as object tracking, human action recognition, pose estimation and occupancy mapping. Many existing techniques only perform well under steady illumination. Some approaches have been introduced to detect foreground under varying or sudden changes in illumination but the problem remains challenging. In this paper, we introduce a new texture-based foreground detection method which is robust to illumination change. Our method detects foreground by finding the correlation between the current frame and a background model. A region with low correlation is detected as foreground. We compare the performance of our proposed technique with other techniques from literature (edge-based, ViBe and Gaussian mixture model) as a preprocessing step of the multi-camera occupancy mapping system. The evaluation demonstrates that our technique outperforms the other methods in term of object loss.**

I. Introduction

Nowadays, many image processing applications require reliable foreground detection as a part of their work flow, typically as preprocessing step. For instance, a multi-camera occupancy mapping described in [1] uses edge-based foreground detection [2] on each smart camera to estimate the location of people in the image plane and later fusion center makes the global estimate of the position of people in real world coordinates from estimations of all cameras. There are many techniques to detect foreground object in both indoor and outdoor environments. However, the performance of most techniques becomes unreliable when lighting varies abruptly.

Though several state-of-the-art techniques [3], [4], [5], which can adapt to lighting change have been introduced, the task of detecting foreground reliably remains challenging. The foreground detection method [2] based on analysis of image gradient is reported to perform much better than [3] and [5] in the presence of lighting changes but some parameters are light-sensitive.

Our method presented detects foreground by finding texture changes in the image of interest against trained background model. We use our method and three state-of-the-art techniques [3], [5], [2] as foreground detection for multi-camera occupancy mapping [1] and compare the performance in term of object loss. The results show that the occupancy map using our method achieves the lowest number of object losses.

II. Data

We use two video recordings that are captured by a network of six cameras (780×580 pixels at 20 FPS) in an 8.8 m by 9.2 m meeting room setup. Both recordings contain two persons walking around in the room while the light of the room is dimmed, switched on and off a couple of times to simulate the scenario of unsteady illumination. Each recording last a bit over one minute in which ground truth positions on the ground plane of each person are manually annotated in every second.

III. Methods

A. Foreground Subtraction Using Correlation

First, the background model is constructed by simply computing the average of the first few frames that do not contain any foreground object. We experimentally found that taking average of 50 fames is good enough to construct a background model. Once the background model is constructed, a sliding window Ω is swept one pixel per step over the frame of interest (FOI) and the background model. At each step, the correlation of pixels within the sliding window I and the corresponding pixels in the background model B is computed as follows:

$$\rho = \frac{\langle I, B\rangle_\Omega}{\|I\|_\Omega \|B\|_\Omega}. \tag{1}$$

If the correlation ρ is less than a threshold ρ_{thresh}, a pixel at the center position of the sliding window is classified as foreground and as background otherwise. Since our method uses correlation to detect changes in texture between the FOI and the background model, a less textured foreground region in front of a less texture background is often not detected. This creates holes in the foreground objects as shown in Figure 1(b). We overcome this issue by using the convex hull to produce the foreground silhouettes as shown in Figure 1(c).

B. Evaluation Method

We make a performance comparison of our method with the other methods [3], [5] and [2] by using each method as a foreground subtraction step of the multi-camera occupancy mapping described in [1]. In our comparison, we use the Bayesian estimator to fuse data instead of Dempster-Shafer reasoning. The parameters of our method are set as $sizeof\Omega = 10 \times 10$ and $\rho_{thresh} = 0.98$. These values are experimentally

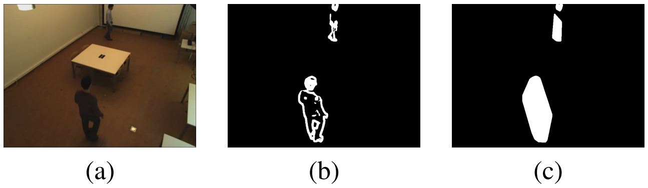

Fig. 1. An example of (a) input frame, (b) detected foreground using correlation, (c) output after applying convex hull.

TABLE I
NUMBER OF OBJECT LOSSES IN EACH VIDEO RECORDING.

Recording	GMM	ViBe	Edge based	Our method
1	33	42	9	2
2	42	43	10	1
Total	75	85	19	3

selected. Parameters of other methods are tuned to obtain optimal detection.

The system produces the locations of each person on the ground plane from the aforementioned video recordings. We use the total number of object losses per person in each video recording as a performance measure. Once the Euclidean distance between the estimated position of a person and the ground truth position is more than 70 cm, the object loss count is increased by one. A lower number of object losses indicates the better performance of the foreground detection technique.

IV. RESULTS AND DISCUSSION

The comparison results are shown in Table I. GMM, ViBe and the edge-based foreground detection method are presented in [3], [5] and [2] respectively. We can see that our method clearly outperforms the other three methods. We found that object loss in our method usually occurs when the light in the room is switched off due to the presence of excessive false positive detection.

Figure 2 illustrates that all four methods perform well under steady illumination. When the lighting of the room is reduced by about half , ViBe fails to detect both persons while the other three methods still detect them. However there are false positive detections at the feet in GMM and lower body parts are not detected by the edge-based method. Once the light is off, the detection of both ViBe and GMM become very unreliable. Although the edge-based and proposed method perform poorly, the full body of both persons are still detected by our method. We observe that a failure to detect full body (for example, the lower body parts are not detected) or excessive false positive often leads to object loss. This clearly shows the reason why our proposed method outperforms the other methods in occupancy mapping.

V. CONCLUSION

This paper presents a new illumination-robust foreground detection method. We compared of our method with other methods in literature and found that our method results in the lowest number of object losses under varying illumination.

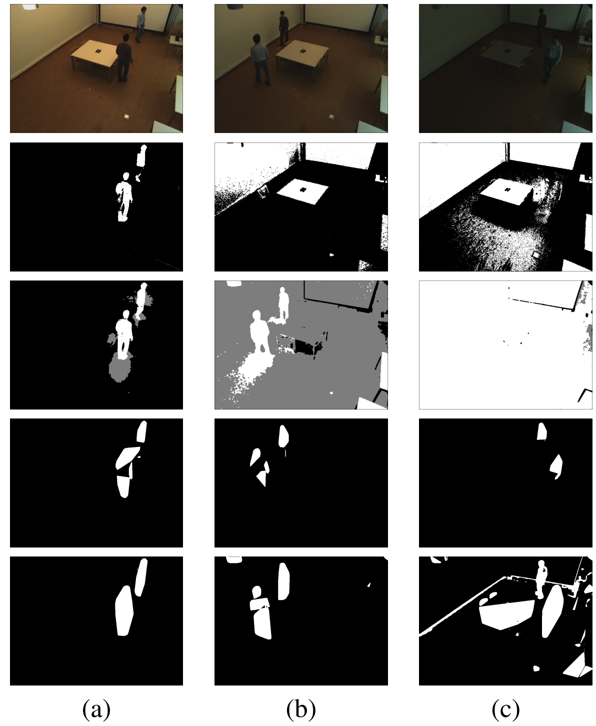

Fig. 2. Top to bottom: original image, foreground detection by ViBe, GMM (gray pixels are result of shadow detection and not qualify to be foreground), Edge-based and our proposed method when (a) there is normal lighting, (b) the lighting is reduced, (c) the lighting is off.

The method works well under various illumination condition without a need to tune parameters. Therefore, we conclude that the proposed method detects foreground more reliably than the other investigated methods. However, it does not update the background model, and thus relocation of static objects (for instance, tables and chairs) may cause false positives. As future work, we will integrate background model updating.

ACKNOWLEDGMENT

This work has been supported by the iCocoon project of the Flemish Interdisciplinary Institute for Broadband Technology (IBBT).

REFERENCES

[1] S. Gruenwedel, V. Jelača, J. Niño Castañeda, P. Van Hese, D. Van Cauwelaert, P. Veelaert, and W. Philips, "Decentralized tracking of humans using a camera network," in *Proceedings of SPIE, Intelligent Robots and Computer Vision XXIX: Algorithms and Techniques*, vol. 8301. SPIE, 2012.

[2] S. Gruenwedel, P. Van Hese, and W. Philips, "An edge-based approach for robust foreground detection," in *Proc. Advanced Concepts for Intelligent Vision Systems (ACIVS) 2011*, 2011, pp. 554–565.

[3] Z. Zivkovic, "Improved adaptive gaussian mixture model for background subtraction," in *Proceedings of the Pattern Recognition, 17th International Conference on (ICPR'04) Volume 2 - Volume 02*, 2004, pp. 28–31.

[4] J. Pilet, C. Strecha, and P. Fua, "Making background subtraction robust to sudden illumination changes," in *Computer Vision–ECCV 2008*, ser. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2008, vol. 5305, pp. 567–580.

[5] O. Barnich and M. Van Droogenbroeck, "Vibe: A universal background subtraction algorithm for video sequences," *IEEE Transactions on Image Processing*, vol. 20, no. 6, pp. 1709–1724, 2011.