Unsupervised Person Detection through Occluding Foliage

A novel approach based on the variance between integral images

**Lisa Schneckenreiter**  
 JKU, Linz  
Artificial Intelligence  
K01507090@students.jku.at

**Daniel Deutsch**  
 JKU, Linz  
Artificial Intelligence  
[K01153228@students.jku.at](mailto:%4b%30%311%353%322%38@%73tu%64%65%6e%74%73%2ej%6b%75%2ea%74)

**Nikolas Boldis** JKU, Linz  
Artificial Intelligence  
[K11706481@students.jku.at](mailto:K11706481@students.jku.at)

**Uta Wagner**  
JKU, Linz  
Artificial Intelligence [K51904650@students.jku.at](mailto:%4b5%31%39%30%34650@%73t%75d%65%6et%73.j%6b%75%2ea%74)

1 Introduction

The detection of moving targets in a forest through occluding foliage is an essential task in wilderness search and rescue. Nathan et al. designed a drone-operated 1D camera array for parallel synthetic aperture aerial imaging and showed that anomaly detection greatly benefits from the integration of parallelly recorded imaged [1]. Here, we present a novel approach, applying computer vision methods to the pixel-wise variance between integral images from a series of recording timepoints.

2 Method

The training, validation and test datasets for this project contain 26, 11 and 13 samples, respectively, each consisting of 70 images from parallel recording by 10 cameras at 7 sequential timepoints.

In a first pre-processing step, the provided mask and homography matrices, are used to remove the timestamp labels and to warp all images per sample to the same perspective. The resulting images are then grouped by the timepoint of recording and averaged over all ten images recorded in parallel by the camera array. Subsequently, they are cropped to a size of 700x315 pixels to remove any areas not covered by all ten input images and the pixel-wise variance between the 7 integral images per sample is calculated. Assuming that a moving person would increase the variance in the red or the blue colour channel as opposed to the green channel, which would most likely represent movement of vegetation, we subtract the green colour channel (varG) of the variance image from the blue (varB) and red (varR) one, respectively, set all negative values to 0 and multiply by a scaling factor of 500 (equations (1) and (2)).

(1)

(2)

In the resulting two greyscale images the contrast is increased, and the brightness decreased by a custom-written function and subsequently, convolutions with a 5x5 dilation kernel, a 7x7 erosion kernel and another 5x5 dilation kernel are successively applied using OpenCV’s functions *cv2.dilate()* and *cv2.erode()* [2]. Finally, the images are binarized by setting a threshold at 1.

With the help of the OpenCV functions *cv2.findContours()* and *cv2.boundingRect()* [2], bounding boxes surrounding all bright spots in the binary images derived from the blue and the red colour channel, respectively, are created. After padding the rectangles by two pixels in every direction, overlapping boxes within one image are merged and the results increased by 14 pixels along the x-axis and 8 pixels along the y-axis. The biggest bounding box by area of the images derived from the blue and the red channel, respectively, is returned as a detection if bigger than 24 pixels in width and 16 pixels in height. In case no detections fulfilling these criteria can be found, the threshold for binarizing the processed images is lowered by 1 iteratively until at least one suitable bounding box is detected. As a final step, bounding boxes smaller than 38x30 pixels are padded to that size.

3 Results

On the given validation dataset an average precision of 0.574 was achieved.

4 Discussion

The described algorithm performs very well on the given dataset and solves the problem of detecting moving people in an occluded forest with high precision in a simple way without any need for machine learning.

However, one drawback is that the method is biased towards the detection of movement in the blue and red colour channel and therefore, will have difficulties detecting people wearing for example green clothing. This goes even further, as only up to one person will be found per colour channel, technically limiting the amount of detected people per sample to one wearing blue and one wearing red. To make this more general, one could choose all bounding boxes fulfilling a certain size criterion instead of only taking the biggest one. Moreover, the algorithm will always at least yield one detection, which is due to the iterative step in the thresholding of the binary image. Removing this iterative approach might lead to better generalization of the method, but probably decreases the performance on this specific dataset. We opted in favour for this trade-off because of the given challenge design.

REFERENCES

[1] Nathan RJAA, Kurmi I., Schedl D., Bimber O., 2021. Through-Foliage Tracking with Airborne Optical Sectioning. *arXiv*

[2] Bradski G., 2000. The OpenCV Library. *Dr. Dobb’s Journal of Software Tools*

Conference Name:ACM Woodstock conference

Conference Short Name:WOODSTOCK’18

Conference Location:El Paso, Texas USA

ISBN:978-1-4503-0000-0/18/06

Year:2018

Date:June

Copyright Year:2018

Copyright Statement:rightsretained

DOI:10.1145/1234567890

RRH: F. Surname et al.

Price:$15.00