Detection of Fallen Persons and Person Shadows from Drone Images

Taisei Ono, Haruka Egawa, Yuta Ono
Graduate School of Engineering, Kyushu Institute of Technology
Sensuicho1-1, Tobata, Kitakyushu, Fukuoka, 804-8550
E-Mail: ono.taisei774@mail.kyutech.jp, Japan

Seiji Ishikawa*, Joo Kooi Tan**
*Emeritus Professor, ***Faculty of Engineering, Kyushu Institute of Technology*
1-1 Sensui-cho, Tobata, Kitakyushu, Fukuoka, 804-8550, Japan
E-Mail: etheltan@cntl.kyutech.jp, Japan
https://www.kyutech.ac.jp/~etheltan/

Abstract

In recent years, the development of automatic search methods based on aerial images taken by drones has been attracting attention in order to prevent secondary disasters and to perform rescue operations quickly in the search for victims of natural disasters. Although various methods exist for automatic person detection for search, they are based on the assumption that the background area of a person captured by a drone camera is a uniform ground in which only those persons who require rescue exist without any shadows or trees. In this paper, we propose a method of automatic detection of both fallen persons and person shadows, or trees on the ground. The method is combination of Ri-HOG and Ri-LBP features to search for fallen persons. These features are robust to rotation. We then employ GrabCut and brightness values to detect shadows. The effectiveness of the proposed method was verified by experiment.

Keywords: Fallen person, Shadow of person, Drone images, Value histogram, GrabCut, Random forest.

1. Introduction

In the event of a disaster, the survival rate of victim declines rapidly with the passage of time, so rapid search activities are necessary. In addition, because of the risk of secondary disasters during disaster search activities, "quick" and "safe" search activities using drones have been attracting attention. As for person search, there is a person detection method using a heat source[1], but it is difficult to distinguish whether the person is standing or lying down. In addition, the fallen person detection method[2] and the fallen person head detection method[3], which focus on the shape of the person, incorrectly detect the shadow of an upright person whose shape is similar to that of a fallen person. Therefore, in this paper, we propose an automatic detection method for fallen persons and person shadows.

We define a shadow of a standing person outdoors as a person shadow, and propose a method of detecting person shadows from images captured by a camera mounted on a drone. Previous studies on shadow detection used RGB color information[4]. Since shadow colors generally have low brightness, we believe that the use of brightness is effective for shadow detection. Another study[5] used saturation as a feature for shadow detection, but saturation is easily affected by sunlight outdoors, so the color information changes significantly depending on light exposure. Other shadow detection methods include texture-based shadow detection[6,7], but these methods assume that texture exists in the shadows of all scenes in a video, and texture detection is difficult in aerial video taken with a drone due to resolution issues.

In this study, we propose a method of detecting fallen persons and person shadows using brightness as a feature. By discriminating between fallen persons and person shadows, it is possible to detect standing persons in addition to persons lying on the ground.

The proposed method uses the Rotation-invariant Histogram of Oriented Gradients (hereinafter referred to as Ri-HOG) feature[8] and the Rotation-invariant LBP (hereinafter referred to as Ri-LBP) feature[9] obtained from the circular cells to detect the shape of fallen persons and that of person shadows. After extracting the foreground by GrabCut[10] from the detected window, the brightness features are extracted to discriminate between a fallen person and a person shadow figure. Random Forest[11] is used as a discriminator for detection. There are two types of discriminators; a discriminator for person/shadow candidates that uses the images of a person falling down (a positive class) and the ground images (a negative class), and a discriminator for person/shadow candidates that uses the images of a person falling down after the GrabCut (a positive class) and the images of a person shadow after the GrabCut (a negative class). These detectors use brightness as a feature value.

2. Proposed Method

2.1. *Detection of fallen person and person shadow candidates*

In the proposed method, Ri-HOG[8] and Ri-LBP features are used for person and person shadow candidates detection. Ri-HOG features are HOG[12] features with rotation-invariance. Unlike HOG that uses rectangular cell, Ri-HOG uses cell arrangement that divides concentric circles. Ri-LBP[9] features are features that adds rotation invariance to Local Binary Pattern (LBP).

2.2. *Foreground region extraction*

GrabCut is used to extract foreground regions. In this section, we describe GraphCuts first and then GrabCut. Graph Cuts is a region segmentation method before applying GrabCut.

2.2.1. *Graph Cuts*

GraphCuts[13] is a method to efficiently find the combination of labels with the minimum cost, given a cost function that gives the minimum value when the labels are appropriately labeled. The cost function $E(L)$ used in Graph Cuts is shown below.

$$E(L) = R(L) + \lambda \cdot B(L) \tag{1}$$

$$R(L) = \sum_{p \in P} R_p(L_p) \tag{2}$$

$$B(L) = \sum_{\{p,q\} \in N} B_{\{p,q\}} \cdot \delta(L_p, L_q),$$
$$\delta(L_p, L_q) = \begin{cases} 1 & if L_p \neq L_q \\ 0 & otherwise \end{cases} \tag{3a}$$

$$B_{\{p,q\}} = \exp\left(-\frac{(I_p - I_q)^2}{2\sigma^2}\right) \cdot \frac{1}{\mathrm{dist}(p,q)} \tag{3b}$$

where P is the total set of pixels and p is its element. λ is a non-negative weighting of the relative importance of the domain term $R(L)$ and the boundary term $B(L)$.

The domain term $R(L)$ is a function that depends only on the labels given to the pixels, and is the sum of all pixels when a label is given to pixel p. It is computed from a region where the locations of the target and part of the background are previously taught. The boundary term $B(L)$ is a sum over all neighboring pixels regarding the continuity of neighboring pixels p and q, due to a priori knowledge of the circumstances under which the labels of neighboring pixels should be given and $B_{\{p,q\}}$ is a cost function, where pixels are adjacent to each other, $\mathrm{dist}(p,q)$ is the distance between pixels p and q. When pixels p and q have the same label, the cost is set to zero by $\delta(L_p, L_q) = 0$. σ is a value that accounts for image noise.

Let $\Pr(L|D)$ be the likelihood of data D given L, and let $\Pr(L)$ and $\Pr(D)$ be the prior probabilities of L and D. Then, the Bayesian equation gives the following relationship.

$$\Pr(L|D) = \frac{\Pr(L|D)\Pr(L)}{\Pr(D)} \tag{4}$$

Since $\Pr(D)$ is a fixed value, the problem is to find L that maximizes the right-hand side. The cost-minimization problem is then solved by taking the logarithm of the numerator of the right-hand side and multiplying it by -1 to obtain E(L). Compared with equation (1), the following is obtained.

$$\mathrm{R}(L) = -\ln P(L|D) - \ln P(L) \tag{5}$$

The histograms $\Pr(I|L)$ and $L=$ {"obj", "bkg"} are created from the specified target and the background regions, respectively. Using the probability density function, the domain terms are cost functions as follows;

$$R_p(\text{"obj"}) = -\ln \Pr(I_p / \text{"obj"}) \tag{6}$$
$$R_p(\text{"bgk"}) = -\ln \Pr(I_p / \text{"bkg"})$$

R_p("obj") is a function that doubles the probability density of pixel p when its label is the background and its pixel value, and the boundary term is then calculated by Eq(3b).

2.2.2. *GrabCut*

Based on the assumption that there is a difference between the pixel values of adjacent pixels of the target and the background in the input image, GrabCut[14] calculates the likelihood of the foregroundness and the backgroundness for each pixel in the specified rectangle based on color statistics using the GMM (Gaussian Mixture Model). Likelihood of the foregroundness and the backgroundness is calculated. Graph Cuts is then applied to the segmentation of the specified region. The GMM is relearned from the obtained results, and GrabCut is applied to it. This process is repeated to improve the segmentation accuracy.

2.3. *Extraction of V histogram features*

The Value histogram feature is a feature that excludes hue and saturation from the HSV (Hue, Saturation, Value: HSV) histogram. It is used to discriminate a fallen person and a person shadow.

2.3.1. *HSV Transformation*

First, the input RGB color image is converted from the RGB space to the HSV (Hue, Saturation, Value: HSV) space. The HSV transform is performed using the following equation. Here in case of $H < 0$, add 2π to H, and if $V\_max = 0$, $S = 0$ and $H = infinite$.

$$V_{max} = \max\{R, G, B\},\ V_{min} = \min\{R, G, B\} \qquad (7)$$

$$V = V_{max} \qquad (8)$$

$$S = (V_{max} - V_{min}) / V_{max} \qquad (9)$$

$$H = \begin{cases} \frac{G-B}{V_{max}-V_{min}} \frac{\pi}{3} & (\text{If } V_{max} = B) \\ \frac{B-R}{V_{max}-V_{min}} \frac{\pi}{3} + \frac{2\pi}{3} & (\text{If } V_{max} = R) \\ \frac{R-G}{V_{max}-V_{min}} \frac{\pi}{3} + \frac{4\pi}{3} & (\text{If } V_{max} = B) \end{cases} \qquad (10)$$

2.3.2. *Creating a value histogram*

The proposed method creates a two-dimensional histogram based on Value, excluding hue and saturation from the HSV histogram of a color image. The V histogram represents Value as a variable. The range of the Value is originally from 0 to 180, but it is changed to 45 levels from 0 to 44. and the intensity of the V histogram is normalized to 1 to speed up the computation.

2.4. *Detection of fallen persons and shadows*

The proposed method uses Random Forest as a discriminator, which is characterized by its ability to efficiently learn even high-dimensional features through random learning and its ability to suppress the influence of noise in the teacher signal by randomly selecting training data.

Fig. 1 Example images of the shadow of a person

Fig. 2 Examples of a training image for the shadow of a person detector

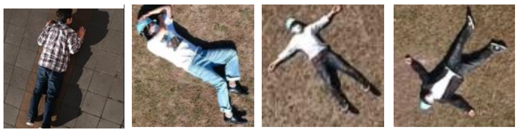

Fig. 3 Example images of a fallen person

Fig. 4 Examples of a training image for a fallen person detector

First, we detect a candidate of a person who fell down based on a discriminator trained using Ri-HOG and Ri-LBP features. Next, the foreground extraction is performed by applying GrabCut to the detection window of the fell-down person candidate. Finally, the foreground image is used to discriminate between a fallen person and shadow of a person based on a discriminator trained using the value V of the extracted foreground image.

3. Experiment

In the performed experiments, images of a fallen person are used as positive class images in the image database for the detection of a candidate fallen person, and the images other than persons are used as negative class images. INRIA Person Dataset[15] was also used as the negative class image. A discriminator is constructed from these training images. All training images are 61[pixels] in height and 61[pixels] in width.

The discriminator between a fallen person and a human shadow using Value was constructed using images such as Figure 2, which is a GrabCut image of a human shadow image shown in Figure 1, as the Positive

Fig. 5 Examples of detection results of a fallen person

class images in the image database, and images such as Figure 4, which is a GrabCut image of the a fallen person image shown in Figure 3, as the Negative class images. All training images are 500[pixels] in height and 500[pixels] in width.

3.1. *Experiment 1*

Four types of videos in which a person is captured by drone are used in the experiment. In all of these videos, there is only one fallen person to be detected.

The result of detection is evaluated with every detection window using Intersection over Union (IoU). The IoU threshold was experimentally set to 0.5. The experimental results showed that the detection rate was 0.80 for video 1, 0.71 for video 2, 0.78 for video 3, and 0.82 for video 4.

Examples of experimental results are shown in Figure 5. The red rectangle represents the detection result of a person who fell down. Figure 6 shows examples of the detection results of a fallen person after the GrabCut process.

3.2. *Experiment 2*

In this experiment, two types of videos in which a person shadow is captured by a drone are used. In all of these videos, there is only one person to be detected.

As in Experiment 1, detection is evaluated frame by frame usingIoU. The IoU threshold is experimentally set to 0.5. The detection rate was 0.83 for video 5 and 0.95 for video 6.

Examples of the experimental results are shown in Figure 6. The blue rectangle represents the result of shadow detection after the GrabCut process. Examples of the experimental results is shown in Figure 7, where the proposed method is applied to the image before the Grab Cut. The blue rectangle represents the result of shadow detection. Figure 8 shows examples of the shadow detection after the GrabCut process.

4. Conclusion

In this paper, we proposed a method of discriminating between fallen persons and person shadows from aerial video images. Experimental results show that the average detection rate of a fallen person is about 78% when there

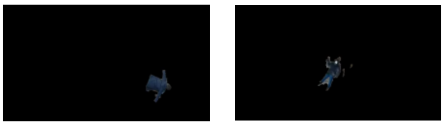

Fig. 6 Examples of detection results of a fallen person after the GrabCut process

Fig. 7 Examples of a detected person shadow

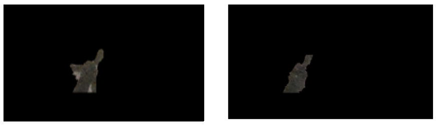

Fig. 8 Examples of a detected person shadow after the GrabCut process

is one fallen person in the frame, and the detection rate of a person shadow is about 89% when there is one person shadow in the frame. In the future, we aim at improving the detection rate in various cases, such as when the altitude is high.

References

1. A. Gaszczak, T. P. Breckona, J. Hana, "Real-time people and vehicle detection from UAV imagery", Proceedings of the SPIE International Conference on Robots and Computer Vision XXVIII: Algorithms and Techniques, Vol.7878, pp.78780B-78780B-13, 2011.
2. H. Egawa, S. Ishikawa, J. K. Tan, "Detection of a fallen person from UAV images using rotation invariant features", Proc. Biomedical Fuzzy Systems, pp.91-94, 2020. (In Japanese)
3. H. Egawa, S. Ishikawa, J. K. Tan, "Detection of a fallen person and estimation of the head position from UAV images", The 2021 International Conference on Artificial Life and Robotics (ICAROB2021), pp.279-282, 2021.
4. P. Kumar, K. Sengupta, A. Lee, "A comparative study of different color spaces for foreground and shadow detection for traffic monitoring system", Proceedings of the IEEE 5th International Conference on Intelligent Transportation Systems, pp.100-105, 2002.

5. V. Gomes, P. Barcellos, J. Scharcanski, "Stochastic shadow detection using a hypergraph partitioning approach", Pattern Recognition, Vol.63, pp.30-44, 2017.
6. A. Leone, C. Distante, F. Buccolieri, "A texture-based approach for shadow detection", Proc. IEEE Conference on Advanced Video and Signal Based Surveillance, pp.371-376, 2005.
7. A. Leone, C. Distante, "Shadow detection for moving objects based on texture analysis", Pattern Recognition, Vol.40, pp.1222-1233, 2007.
8. T. Fujimoto, Y. Kameda, I. Matsuda, S. Itoh, "Person detection from top view images using rotation invariant HOG features", Proc. The Institute of Image Information and Television Engineers, pp.17_10_1-17_10_2, 2014.
9. M. Pietikaine, T. Ojala, Z. Xu, "Rotation-invariant texture classification using feature distribution", Pattern Recognition, Vol.33, pp.43-52,2000.
10. D. Khattab, C. Theobalt, A. S. Hussein, M. F. Tolba, "Modified GrabCut for person face segmentation", Ain Shams Engineering Journal Volume 5, Issue 4, December 2014, pp.1083-1091.
11. L. Breiman, "Random forests", Machine Learning, Vol.45, No.1, pp.5-32, 2001.
12. N. Dalal, B. Triggs, "Histograms of oriented gradients for Person detection", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp.886-893,2005.
13. Y. Boykov, M-P. Jolly, "Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images", Proc. of IEEE International Conference on Computer Vision, I, pp.105-112, 2001.
14. C. Rother, V. Kolmogorov, A. Blake, "GrabCut": Interactive foreground extraction using iterated graph cuts",ACM Transacctions on Graphics, Vol, 23, No.3, pp.309-314, 2004.
15. INRIA Person Dataset, https://b.hatena.ne.jp/entry/pascal.inrialpes.fr/data/person

Authors Introduction

Mr. Taisei Ono

He received his B.E. degree in Department of Control Engineering in 2021 from the Faculty of Engineering, Kyushu Institute of Technology in Japan. He is acquiring the M.E. in the same University. His research interests are detection of fallen person, and shadow of person, image processing of drone images.

Ms. Haruka Egawa

She received B.E.and M.E. in Control Engineering from the Graduate School of Engineering, Kyushu Institute of Technology, Japan. Her research includes aerial images processing, human/fallen human detection and motion recognition.

Mr. Yuta Ono

He obtained the M.E. degrees in Department of Control Engineering from Kyushu Institute of Technology in 2017, Japan. He is presently a doctor course student in the same university. His research interests include feature extraction, machine learning, human detection and recognition.

Prof. Dr. Seiji. Ishikawa

Seiji Ishikawa graduated from The University of Tokyo and was awarded BE, ME and PhD there. He is now Emeritus Professor of Kyushu Institute of Technology. He was Visiting Researcher of The University of Sheffield, UK, and Visiting Professor of Utrecht University, NL. His research interests include visual sensing & 3-D shape/motion recovery. He was awarded The Best Paper Award in 2008, 2010, 2013 and 2015 from Biomedical & Fuzzy System Association, Japan. He is a member of IEEE.

Prof. Dr. Joo Kooi Tan

Prof. Tan is currently with the Department of Mechanical and Control Engineering, Kyushu Institute of Technology, as a Professor. Her current research interests include ego-motion analysis by MY VISION, three-dimensional shape/motion recovery, human detection, and its motion analysis from video. She was awarded SICE Kyushu Branch Young Author's Award in 1999, the AROB Young Author's Award in 2004, the Young Author's Award from IPSJ of Kyushu Branch in 2004, and the BMFSA Best Paper Award in 2008, 2010, 2013 and 2015. She is a member of IEEE, The Information Processing Society, The Institute of Electronics, Information and Communication Engineers, and The Biomedical Fuzzy Systems Association of Japan.