# Detection of a Fallen Person and Estimation of the Head Position from UAV Images

**Haruka Egawa, Seiji Ishiakwa, Joo Kooi Tan**

*Department of Mechanical and Control Engineering, Kyushu Institute of Technology, Sensui-cho Tobata-ku*
*Kitakyushu, 804-8550, Japan*
*E-mail: {egawa-haruka , etheltan }@ss10.cntl.kyutech.ac.jp*
*http://lab.cntl.kyutech.ac.jp/~etheltan/*

**Abstract**

In order to search for victims in the event of a disaster, we propose a method of detecting a fallen person from UAV images and estimating its head position. Rotation-invariant HOG and Rotation-invariant LBP are used, so that they may detected a fallen person regardless of its body orientation. In addition, the position of the head of a fallen person is estimated using the peak of the gradient histogram. Experimental results show satisfactory performance of the proposed method.

*Keywords*: *Aerial photographs,* Fallen persons, Rotation-invariant HOG, Rotation-invariant LBP, *Random Forest, Gradient histogram.*

## 1. Introduction

Japan is one of the world's leading countries of earthquakes [1] and has been hit by disasters caused by earthquakes. In the event of such a disaster, it is often very difficult to search for the victims manually. Therefore, recently, attention has been paid on the development of an automatic search method for those victims from aerial images taken by drones.

There are various human detection methods for automatic search [2,3,4], but most of them are based on the premise that the target person's body orientation is unified and the subject is upright [2,3]. Exists. However, there is no uniformity in the orientation of the person's body taken with the drone. Therefore, in this paper, we propose an automatic detection method for a fallen person that does not depend on the body orientation of the detection target. In this method, after detecting a fallen person [5], his/her head position is estimated and detected. This enables quick support of the fallen person for communication, medical care, food, etc.

The proposed method uses Rotation-invariant Histogram of Oriented Gradients (Ri-HOG) features [6] and Rotation-invariant Local Binary Pattern (Ri-LBP) as features that are robust to rotation in order to represent a fallen person. ) Use the feature quantity [7]. In addition, Random Forest [8] employed as a classifier. In addition, the head position is estimated using the peak value of the gradient histogram in the circular region.

The proposed method is examined its performance by the experiment using bird-eye view videos.

## 2. Method

### 2.1. *Ri-HOG features*

The Ri-HOG feature is a feature of HOG [9] with rotational invariance. Unlike the original HOG, which uses rectangular cells, the cell arrangement obtained by dividing concentric circles is used in Ri-HOG. This section describes the Ri-HOG feature.

### 2.1.1. *Calculation of luminance gradient information*

The input color image is gray-scaled and the brightness gradient at each pixel is calculated. From the obtained luminance gradient, the luminance gradient intensity and the luminance gradient direction in the Cartesian coordinate representation are obtained.

Assuming that the origin $O$ is at the center of an input image, the difference between the declination $\varphi(x, y)$ of a pixel position $(x, y)$ provided by the polar coordinate representation and the obtained brightness gradient direction $\theta(x, y)$ of the pixel defines the brightness gradient direction $\theta'$ in the polar coordinate representation . $\theta'$ is defined by

$$\theta'(x, y) = \theta(x, y) - \varphi(x, y) \qquad (1)$$

### 2.1.2. *Creating a 2-D histogram*

In the proposed method, a concentric circle consisting of three large, medium and small circles is divided into 36 in the angular direction, and each small area is defined as a cell. The cell numbers of the small concentric circle are given as $1,4,7\cdots,106$ clockwise from the start point, starting from the area in the 0 ° direction. Similarly, the cell numbers of the middle circle are $2,5,8, \cdots, 107$, and the cell numbers of the largest circle are $3,6,9, \cdots, 108$.

The offset regions $W_1 (p), W_2 (p), \cdots, W_6 (p)$ are further set by dividing the semicircular region existing in the radial direction of the pixel of interest $p = (x, y)$ into six parts. Let $q$ be the pixel in the offset region, and, from pixels $p$ and $q$, a 2-D histogram is calculated to use the co-occurrence information of their luminance gradient directions $\theta'$. Six 2-D histograms are created with each cell as it has 6 offsets. The 2-D histogram is calculated using the following formula;

$$H_{ij}(o_1, o_2) = \sum_{p \in S_i} \sum_{q \in W_j(p)} (m(p) + m(q)) K(\theta'(p) \in o_1) K(\theta'(q) \in o_2) \qquad (2)$$

Here, $i$ is the cell number, $j$ is the offset number, $o_1$ is the luminance gradient direction of pixel $p$ after quantization, and $o_2$ is that of the offset pixel $q$ after quantization. $S_i$ is the cell of present concern, and $K()$ is the function whose value is 1, if the argument is true, and 0, otherwise.

### 2.1.3. *Normalization*

In order to reduce the influence of changes in local brightness and contrast, normalization is performed for each 2-D histogram obtained by Eq. (2).

Since the values of luminance gradient direction θ of the entire circular region in the Cartesian coordinate representation range from 0 ° to 360 °, it is quantized in 36 directions every 10 ° interval. A gradient histogram is created by adding the luminance gradient intensities in the circular region based on the luminance gradient direction in the quantized Cartesian coordinate representation. After that, the cells are rearranged in the clockwise order from the reference direction based on the angle at which the gradient histogram takes a peak, and the vector in which the 2-D histograms of each cell are connected is used as the Ri-HOG feature quantity. Here, the cell numbers of the smallest concentric circle are expressed as follows;

$$i_u = 1,4,7,\cdots,N \ (u = 1,2,3,\cdots 36) \qquad (3)$$

Using Eq.(3), the cells are connected as $i_1, i_1 + 1, i_1 + 2, i_2, i_2 + 1, i_2 + 2, i_3 \cdots, i_{36} + 2$ , if the peak of the gradient histogram is 0 °, and they are connected as $i_{36}, i_{36} + 1, i_{36} + 2, i_1, \cdots, i_{34} + 2, i_{35}, i_{35} + 1, i_{35} + 2$, if the peak of the gradient histogram is 350 °.

## 2.2. *Ri-LBP features*

The Ri-LBP feature is a feature obtained by adding rotational invariance to the Local Binary Pattern (LBP) [10]. This section describes the Ri-LBP feature.

### 2.2.1. *Local binarization*

The input image is gray-scaled, and, as shown in **Fig. 1**, the brightness values of the pixel of interest and its 8 adjacent pixels are compared. If the gray value of an adjacent pixel is larger than or equal to that of the center pixel, the adjacent pixel is given 1, otherwise given 0. By aligning the obtained 0 or 1 in the direction of the arrow shown in Fig. 1 (b), an 8-digit binary number is obtained, which represents the center pixel.

In Ri-LBP, the starting point is changed in turn to acquire eight 8-digit binary numbers, and the minimum value out of the 8 numbers represents the center pixel.

For example, in Fig. 1, the normal LBP value is $0100\ 0011_b = 67_d$ , but the Ri-LBP value is $0000\ 1101_b = 13_d$.
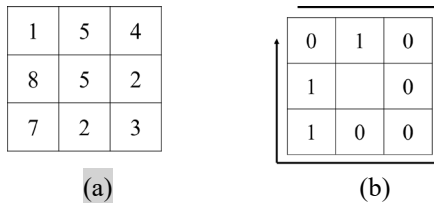
(a)                    (b)

Fig. 1. LBP

### 2.2.2. *Histogram creation*

The cell of the Ri-LBP feature is the same as the cell of Ri-HOG feature.

Since LBP is given as an 8-digit binary number, there are 256 values from 0 to 255, but in Ri-LBP, there is an 8-digit binary number that becomes the same value by changing the start point. For example, $1000\,0000_b = 128_d$ in LBP is expressed as $0000\,0001_b = 1_d$ when the start point is changed. Hence $1000\,0000_b$ and $0000\,0001_b$ are equivalent in Ri-LBP. After all, the total number of Ri-LBP values is 36.

Since there are cases where the number of pixels in one cell is less than 36, the proposed method divides 36 patterns into 9 patterns and creates a histogram with each cell. This histogram is normalized. The cells are rearranged based on the reference angle that gives the peak in the brightness gradient direction obtained in Section 2.1.3, and the histograms of all the cells are concatenated. This provides the Ri-LBP feature.

### 2.3. *Detection of a fallen person*

The proposed method uses Random Forest as a discriminator. The features of Random Forest include that it is possible to learn efficiently even with high-dimensional features by random learning, and that the influence of noise in the teacher signal can be suppressed by random selection of learning data.
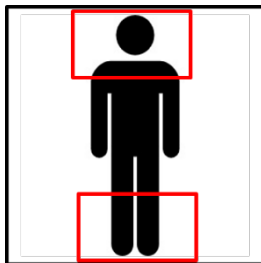


Fig. 2. Example of presumed head positions in the fallen person detection window.

### 3. Head position estimation

The head position of a fallen person is estimated using the angle at which the gradient histogram in the circular region obtained in Section 2.1.3 has a peak. When a fallen person exists in the detection window in the orientation shown in **Fig. 2**, the horizontal gradient becomes large and the peak angle of the gradient histogram is 0 ° or 180 °. Therefore, the areas shown by the red frame in Fig. 2 are the estimated head position.

### 4. Experiment

In the experiment, an image of a fallen person is used as an image of the Positive class of the image database. Images other than people are used for Negative class images. The INRIA Person Dataset [11] was also used as an image of the Negative class. A classifier is constructed using these training images. The size of the training image is all 61 pixels vertically and 61 pixels horizontally.

### 4.1. *Experiment 1*

The first experiment is the detection of a fallen person from an image. Four types of moving images taken with a multicopter are used in the experiment. In all of these videos, there is only one fallen person to be detected. In Experimental Video 1, there are some obstacles other than a fallen person. In the experimental video 2 and the experimental video 3, the amount of movement of the camera is large, and there are some frames in which there is no fallen person in the image. In addition, there is an upright person who is not fallen in the experimental video 4.

The detection is evaluated with each detection window using Intersection over Union (IoU). The threshold of the IoU is experimentally set to 0.5. As a result of the experiment, the recall was 0.990 for movie 1, 0.855 for movie 2, 0.565 for movie 3, and 0.875 for movie 4.

An example of the experimental results is shown in **Fig. 3**. The green rectangle is the manually set Ground Truth, whereas the red rectangle shows a detected fallen person

### 4.2. *Experiment 2*

In the second experiment, head of a fallen person is searched on the image of a detected fallen person. For the experiment, 137 images of the fallen person area cut out from the video taken by the multicopter are used.

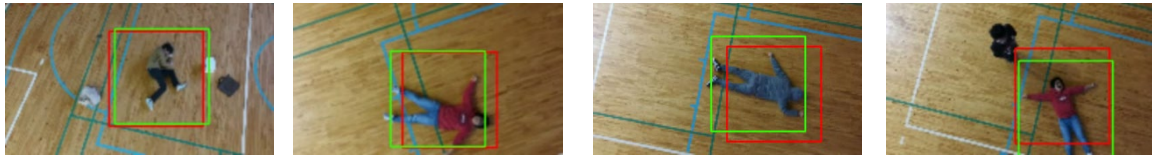*Haruka Egawa, Seiji Ishiakwa, Joo Kooi Tan*

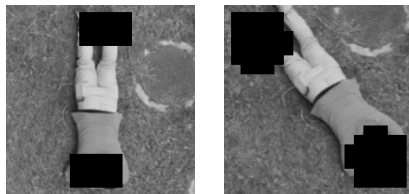Fig. 3. Example of detection result of Experiment 1 .



Fig. 4. Example of the estimated head position in
Experiment 2



Fig. 5. The result of head detection with a fallen person

In order to visualize the result, the estimated head position is painted black. The case where the estimated head position overlaps with the head of the actual lying person is regarded as true. As the result of the experiment, the correct answer rate was 73.7%. An example of the experimental results is shown in **Fig. 4**.

**Fig. 5** shows the results of head detection. It is indicated by a small light-blue square.

## 5. Conclusion

In this paper, we proposed a method of acquiring aerial images with a camera mounted on a multicopter and detecting a fallen person from the images.

In the experiment, the recall rate of the fallen person detection was about 82% when there was one fallen person in the image frame. The correct answer rate for head position estimation using the peak angle of the gradient histogram was 73.7%. In the future, we aim at further improvement of the detection rate by increasing the number of learning samples and detecting a fallen person with occlusion.

## References

1. 平成 25 年版防災白書 付属資料, 内閣府, 2013, p.1.
2. P. Viola, M. Jones: "Rapid object detection using a boosted cascade of simple features", *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 2001, pp.511-518.
3. R. Sato, S. Koshimura: "Development of a method for searching survivors in disaster-affected areas with UAV and image analysis", *J. JSCE*, Ser.B2, Coastal engineering Vol. 69, No. 2,2013, pp.I_1461 - I_1465.
4. A. Gaszczak，T. P. Breckona，J. Hana： "Real-time people and vehicle detection from UAV imagery", *Proc. SPIE Conference on International Robots and Computer Vision XXVIII:Algorithms and Techniques*, Vol.7878, 2011, pp.78780B-78780B-13.
5. H. Egawa, S. Isihkawa, J. K. Tan: "Detection of a fallen person from UAV images using rotation invariant features", *Proc. of Biomedical Fuzzy Systems Association*, 2020, pp.1-4(in Japanese).
6. T. Fujimoto, Y. Kameda, I. Matsuda, S. Itoh: "Human detection from top view images using rotation invariant HOG features", *Proc. of The Institute of Image Information and Television Engineers*, 2014, pp.17_10_1-17_10_2.
7. M. Pietikaine,T. Ojala,Z. Xu: "Rotation-invariant texture classification using feature distribution", *Pattern Recognition*, Vol.33, 2000, pp. 43-52.
8. L. Breiman: "Random forests", Machine Learning, Vol.45, No.1, 2001, pp.5-32.
9. N. Dalal, B. Triggs: "Histograms of oriented gradients for human detection", *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2005, pp.886-893.
10. T. Ojala, M. Pietikainen, D. Harwood: "A comparative study of texture measures with classification based on featured distributions", Pattern Recognition, Vol. 29, No. 1, 1996, pp. 51–59.
11. INRIA Person Dataset: （ http://pascal.inrialpes.fr/data/human/（ Accessed:10 Jan,2019）