Bicyclist Recognition and Orientation Estimation from On-board Vision System

Yanlei Gu\(^1\)  \quad Shunsuke Kamijo\(^2\)

\(^1\) Institute of Industrial Science, The University of Tokyo
4-6-1 Komaba, Meguro-ku, Tokyo, 153-8505 Japan
(E-mail: guyanlei@i.k.u-tokyo.ac.jp, kamijo@i.is.u-tokyo.ac.jp).

Received on January 27, 2015
Presented at the JSAE Annual Congress on October 24, 2014

ABSTRACT: Bicyclists move at speed equivalent to a slowly moving vehicle, and sometimes share the road with vehicles in urban environments. Thus, bicyclists take more challenge for safe-driving compared to pedestrians. Therefore, accident avoidance system is expected to recognize the type of road user, and then can perform further behavior analysis for risk assessment based on the type of road user. Our contribution to this tendency consists of a method to distinguish bicyclists from pedestrians, and reliably estimate the bicyclists’ head orientation and body orientation from image sequences taken by an on-board camera. The output of the proposed method can be used for risk assessment.

KEY WORDS: Safety, Active Safety, Bicyclist, Recognition, Orientation Estimation, On-board Vision [C1]

1. Introduction

Accidents involving vulnerable road users such as pedestrians and bicyclist are one of the leading causes of death and injury around the world. In the other hand, there’s no doubt that reduction of these accidents is supposed to be considered in the development of autonomous driving. Bicyclists move at speed equivalent to a slowly moving vehicle, and sometimes share the road with vehicles in urban environments. Thus, bicyclists take more challenge for safe-driving compared to pedestrians. This paper focuses on this less discussed road user, bicyclists, and proposes a recognition and pose estimation method to ensure the safety of bicyclists.

Although all of road users should be paid attention to in driving, they have different moving patterns. Therefore, collision avoidance systems should provide different behavior analysis streams based on the type of road user. For example, bicyclists and pedestrians should be processed as different objects. In addition, bicyclists are often misclassified as pedestrians in the conventional pedestrian detection system\(^(1)\). Thus, first contribution of this paper is to propose a method to distinguish bicyclist from pedestrian. This recognition phase employs a cascade structure classifier. This classifier distinguishes bicyclist from pedestrian using multiple features and discriminative local area, in order to achieve a high recognition rate.

Besides the type of road user, the pose information is also significant in the perception and prediction for the road user’s behavior. In the domain of human pose, head orientation and body orientation are two of most important factors for traffic safety. The head orientation suggests the direction the bicyclist is paying attention to, and the body orientation strongly denotes the direction the bicyclist is traveling to. So the human pose can be represented as the head orientation and the body orientation in this paper. Thus, the second contribution of this paper is to realize the robust head orientation estimation and body orientation estimation for bicyclist. In this step, Semi-Supervised Learning (SSL) is applied in multi-class classification, which can reduce the difficulty of dataset preparation and improve classification rate. Moreover, temporal constraint and human physical model constraint are considered, which assist the orientation estimation to produce reasonable and stable result in video sequence.

The remaining part of this paper is organized as follows: Section 2 describes the related research works on recognition of pedestrian and bicyclist, and pose estimation. Section 3 presents the first contribution of this paper, a method to distinguish bicyclist from pedestrian. In section 4, the methodology of head orientation estimation and body orientation estimation is described. The experimental results will be demonstrated in section 5. This paper is ended up with a conclusion and a future work consideration in section 6.

2. Related Works

Bicyclist is a kind of pedestrian-like object. The comprehensive and intensive survey on the vision-based pedestrian detection could be found in related works\(^ (2-4)\). In hundreds of papers mentioned\(^ (2-4)\), the most representative one is the Histograms of Oriented Gradient (HOG) method\(^ (5)\). Although one decade has passed since the HOG feature was published, this method still affects the following research considerably. The Deformable Part-Based Model is a remarkable approach proposed for object detection recently, it also employs the HOG feature to describe objects\(^ (6)\). In addition, an extended version of HOG features, referred to as Co-occurrence Histograms of Oriented Gradient (CoHOG), was specially published for pedestrian detection at low resolution\(^ (7)\). The excellent performance of the HOG feature and
the CoHOG feature makes them attractive as the feature descriptor in object detection and recognition tasks.

Besides pedestrian detection, detection of other kinds of road users was also performed. Cho et al. detected bicyclist using both HOG method and Part-Based Model. Multiple view-based detectors: frontal, rear, and right/left side view, were used in detection. However, the difference between pedestrian and bicyclist was not addressed in their works. In addition, based on our previous work, almost bicyclists were detected as pedestrian in experiments, if we applied the pedestrian detector without considering the difference between pedestrian and bicyclist. So this paper will firstly focus on the distinguishing bicyclist from pedestrian.

Recently, researchers put their focus on the pose of pedestrian in the view of traffic safety. Schulz et al. proposed a framework for head localization and orientation estimation. The system consisted of eight separate classifiers associated with the different head pose classes. The head localization and pose estimation was achieved by comparing the output confidence values of all eight classifiers for all windows at all different scales. In order to get stable result, particle filters were employed for tracking in continuous frames. Gandhi et al. used HOG and Support Vector Machines (SVM) to recognize the body orientation of pedestrian. The body orientation of pedestrian was used for predicting the walking path. Compared to the above related works, this paper focuses on the pose estimation of the less discussed road user, bicyclists. Moreover, this paper considers not only head orientation and body orientation, but also the relationship of these two orientations in pose analysis.

3. Recognition of Bicyclist

In this paper, recognition of bicyclist is to distinguish bicyclist form pedestrian. In order to focus on the recognition bicyclist directly, this paper assumes that the exact region of the pedestrians and bicyclists is already obtained. This region could be gotten by any foreground segmentation method, such as stereo vision-based or motion-based. Even though the regions are analyzed by using conventional pedestrian detection method, such as sliding HOG-based pedestrian detector in the images, the bicyclists can not be distinguished because there is pedestrian-like object existing in the bicyclist images. The main reason is the training dataset does not include bicyclist images as negative samples in the conventional pedestrian detector. Suppose the complementary training dataset defines pedestrian and bicyclist as different categories, the combination of HOG and SVM is still the first basic option to realize this recognition task. Here, the image of the bicyclist includes both the area of bicyclist and the area of bicycle.

The appearance of the images of Fig. 1 demonstrates that the recognition task is challenging, because there is obvious difference in inner-class. Besides the inner-class difference, the inter-class similarity could be observed at the same time. In the bicyclist images, the area of bicycle wheel provides a cue, which could be possibly used for distinguish bicyclist from pedestrian. However, it is still a problem to distinguish the front view and the rear view bicyclist from pedestrian. The area of tyre is very narrow, so a more detailed descriptor is needed.

CoHOG feature is a powerful feature descriptor, proposed by Watanabe et al., which can express complex shapes of objects by using co-occurrence of gradient orientations with various positional offsets. Since the CoHOG has high dimensions and presents more detailed information, high detection performance was achieved in the pedestrian detection at low resolution. Therefore, we have reason to believe that CoHOG feature could be used in this research. But, calculation cost of CoHOG feature is expensive because the feature vector is high-dimensional. Therefore, the CoHOG feature will not be used always in our recognition task, it is employed as a verification step. The proposed recognition framework is shown in Fig. 2.

In this flowchart, there are two classifiers. First is “HOG-SVM classifier” connected with HOG feature, this classifier is trained by using full bicyclist vs. pedestrian dataset, which includes bicyclist images and pedestrian images captured from different views, such as the images in Fig. 1. All images are rescaled to 64 by 128 pixels, and then are transformed to HOG feature. This classifier can distinguish side view bicyclist from pedestrian. The second SVM classifier, “CoHOG-SVM classifier”, is located below the CoHOG feature extractor. The CoHOG feature extractor just considers a part of image, denoted by gray

---

**Fig. 1. Images used in the recognition of bicyclist**

We mask the face part in the figure because of the requirement of privacy protection.

**Fig. 2. Flowchart of the recognition of bicyclist**
Head and body orientation estimation has been developed for pedestrian in our previous work (15). Our proposed method will be applied on the new object, bicyclists. In addition, the more detailed direction discretization is achieved for the body orientation estimation in this paper, which can improve the estimation accuracy in the safety aspect. Moreover, we proposed to use the Semi-Supervised Learning in the training step instead of conventional supervised learning. Because this part is the next step of the recognition of bicyclist, the input is still the bicyclist image, such as the images in Fig. 1 (b).

4.1. Head localization and orientation estimation

The bicyclist head area is no always located at the top-center of images, as shown in Fig 1 (b). Therefore, in order to get the exact region of head for head orientation estimation, head localization should be conducted. In this paper, the head localization and head orientation estimation are processed at the same time. The head orientation is discretized into 8 directions with 45 degree interval. So the head localization and orientation estimation becomes a 9-class classification problem, and the training dataset includes 8 positive orientation classes and 1 negative non-head class. Because the resolution of head image is low, CoHOG feature is employed to extract more information from image instead of HOG feature used in our previous work (15).

4.2. Body orientation estimation

In the pedestrian body orientation estimation, 8-direction was used in our previous work (15). However, bicyclist usually has a higher speed compared to pedestrian, the more detailed direction discretization is needed. Therefore, 16-direction discretization is defined for bicyclist body in this paper. Here, the body image means the whole area of bicyclist image including head area, such as the image in Fig. 1 (b). The HOG feature is employed to describe the texture of bicyclist body. In addition, the training of body orientation classifier is conducted using Multi-class SVM as well. The estimated orientation of body image is corresponding to the orientation with highest probability in body orientation estimation.

4.3. Preparation of training data using Semi-Supervised Learning

In order to obtain head and body orientation classifier, the label of each training data is needed. However, it is difficult to decide the label of image near to the boundary between two neighboring classes, such as the head images with two possible labels in Fig. 4 (a). In this paper, we proposed to use Semi-Supervised Learning (SSL) method to label the hard images before the generation of classifier (17).

In SSL, the training data is divided into two categories: “strongly labeled data” and “weakly labeled data” (18). The strongly labeled data has unique label, which can be easily distinguished by human eye. And the “weakly labeled data” is the
images near to the boundary of two neighboring classes, which has two possible labels. The weakly labeled data does not have exact label at the beginning of SSL, so the weakly labeled data can not be utilized for training directly.

As shown in Fig 4 (b), the SSL firstly uses the strongly labeled data to train an initial classifier. And then, this generated classifier predicts labels for all “weakly labeled data” in test step. After the test step, SSL adds $N$ “weakly labeled data” images with highest likelihood and their predicted labels to “strongly labeled data”. The SSL repeats the above mentioned three steps until no new strongly labeled data appears. In the other word, the strongly labeled data determines an initial classifier, and the weakly labeled data greatly refines the classifier. In this paper, the SSL is applied for the preparation of the label of positive head image and the label of body image in training dataset.

4.4. Joint pose estimation

In order to get reasonable and stable estimation result from continuous frames, two constraints are considered in the joint pose estimation step: human physical model constraint and temporal constraint. Generally, the orientation difference between human head and body could not be over 90 degree. Moreover, since we focus on the video sequence, the human pose should not change dramatically in continuous frames. These two constraints are embedded in particle filter.

The flowchart of the proposed method is shown as Fig. 5. In order to get the initialization position and orientation of head, we use sliding window method in the first image ($t_0$). In the sliding window method, the probability of each estimated block image can be represented as $P(d_o \mid I, x_h, y_h, s_h)$, Here, $(x_h, y_h)$ is head position, $s_h$ means head size, $d_o$ denotes body orientation. $I$ is the bicyclist image. Moreover, the probability of body orientation can be described as $P(d_o \mid I, t)$. $d_o$ denotes body orientation. The body and head orientation constraint is applied to initialize the head position, the head scale, the head orientation, and body orientation. The estimation result is denoted by rectangle and arrows on left second image in Fig. 5. The arrow connecting to rectangle means head orientation and the arrow from the center of body means body orientation. After first frame, we employ particle filter to track the head position, the head scale and the head orientation, and the body orientation. The particles are distributed around the estimation result of last frame. The state of one particle $q$ is a vector including five elements: head position $(x_h, y_h)$, head size $s_h$, head orientation $d_h$ and body orientation $d_b$, which is described as equation (1).

$$q = [x_h, y_h, s_h, d_h, d_b]^T$$ (1)

The human physical model constraint is described as follows:

$$L_{model} = P(d_o \mid I, x_h, y_h, s_h) \times P(d_o \mid I, H, \theta_h, 0, k_{au})$$ (2)

where $p_\theta$ denotes von Mises distribution, also known as the circular normal distribution. $\theta_h$ and $\theta_b$ are the body orientation and the head orientation. The $k_{au}$ is a constant value that controls the von Mises distribution. Basically, the value of $L_{model}$ is inversely proportional to the difference between $\theta_h$ and $\theta_b$ And the value of $L_{model}$ is proportional to the probability of head orientation $P(d_h \mid H, \theta_h, 0)$ and the probability of body orientation $P(d_b | H)$.

The temporal constraint $L_{temporal}$ is a multiplication of two parts: $L_{temporal_head}$ and $L_{temporal_bdy}$ which are represented as equation (4) and (5). The format of $L_{temporal_bdy}$ and $L_{temporal_head}$ is similar to $L_{model}$, so the value of $L_{temporal_bdy}$ and $L_{temporal_head}$ is inversely proportional to the orientation difference in two frames.

$$L_{temporal}(q(t), I(t), q(t-1)) = L_{temporal_bdy}(q(t), I(t), q(t-1)) \times L_{temporal_head}(q(t), I(t), q(t-1))$$ (3)

$$L_{temporal_bdy}(q(t), I(t), q(t-1)) = P(d_o(t) \mid I(t)) \times p_d(\theta_h(t) - \theta_b(t-1), 0, k_b)$$ (4)

$$L_{temporal_head}(q(t), I(t), q(t-1)) = P(d_o(t) \mid I(t)) \times p_d(\theta_h(t) - \theta_h(t-1), 0, k_h)$$ (5)
\[
L_{\text{temporal\_head}}(q(t), I(t), q(t-1)) = P(d_{\text{data}}(I(t), x_c, y_c, x_d))
\times p(u(t) - u(t-1), 0, k_d)
\]

In the particle filter, the likelihood of one particle is represented as a multiplication of model constraint \(L_{\text{model}}\) and temporal constraint \(L_{\text{temporal}}\). The likelihood weighted particles generate joint orientation estimation result in each frame.

5. Experiments

In order to demonstrate the effectiveness of proposed methods, we did a series of experiments. The video sequences used in the experiments, were captured by the camera mounted on a car while driving around city under fine weather condition (sunny) in daytime. The camera was set to VGA resolution of 640 by 480 pixels. The pedestrian and bicyclist images used in experiments were cropped from these video sequences. This study considered pedestrians and bicyclists, whose height is in between 90 pixels to 200 pixels, as the interesting targets for test. The average height of the pedestrian and bicyclist images is 139 pixels.

5.1 Evaluation for recognition of bicyclist

The first experiment is related to the recognition of bicyclist. This experiment is corresponding to the proposed method in section 3. The number of images in dataset and their properties are shown in Table 1. The recognition rate in the test dataset is shown in Table 2.

The HOG-SVM classifier achieves to 94.9% recognition rate totally, but the correct recognition rate of bicyclist is relative low. In the experiment, 289 bicyclist images are misclassified, and most of misclassifications belong to the rear view and the front view, are shown in Fig. 6. After verification using CoHOG-SVM classifier, the classification rate is improved to 99.0% totally. 245 misclassifications (ground truth is bicyclist) are corrected after verification. This result proves the analysis in section 3. The HOOG feature could not distinguish front view and rear view bicyclist from pedestrian very well, because the front view and rear view bicyclist have similar patterns as the pedestrian in the HOG feature domain. The CoHOG-SVM classifier refines the classification result by using the more detailed feature and focusing on discriminative area.

Moreover, we prepared another kind of bicyclist images, which just includes the area of bicyclist. The number of images and property are totally same as Table 1. The effectiveness of the proposed method is certified again by using old pedestrian images, as shown in Fig. 1 (a) and Fig. 3 (a), and these new prepared bicyclist images, as shown in Fig. 7. The recognition rate in the test dataset is shown in Table 3. The test image does not include the area of bicycle as well. The experimental result demonstrated that the proposed method also can improve the recognition rate on the pure bicyclist images. It denotes that the posture of bicyclist is an effective cue for recognition as well.

5.2 Evaluation for Semi-Supervised Learning (SSL)

This subsection focuses on the preparation of training dataset and the evaluation of the proposed SSL method. The effectiveness of SSL is evaluated for head orientation estimation and body orientation estimation, respectively. In each evaluation, two kinds of training datasets are used. In the first kind of dataset, the labels of all images are decided manually. This dataset is used in the conventional supervised learning as a baseline for comparison. The images of the second kind of dataset are the same as the first kind of dataset, but the images are divided into two categories: “strongly labeled data” and “weakly labeled data”. The “strongly labeled data” is labeled manually, and the SSL will distribute label for “weakly labeled data”. Moreover, the test dataset is carefully prepared using manually labeling way.

The number of images in each dataset is shown in the Table 4. The recognition rate of the conventional supervised learning method and the proposed SSL method are compared, which is demonstrated in Table 5. In the calculation of recognition rate,
we consider an orientation estimation to be correct if the predicted orientation is identical to the true orientation or one of its direct neighboring orientation, which is adopted in the related research as well (10), so the resolution of head orientation estimation is 90 degree and the resolution of body orientation estimation is 45 degree.

Compared to the conventional supervised learning method, the classifier generated from SSL maintains or even improves the classification accuracy in head and body orientation estimation. The classification accuracy is the one factor to evaluate the effectiveness. There is another difference between the SSL and the conventional supervised learning method, which is the number of strongly labeled data used in training, as shown in Table 4. The experiment result demonstrated that the SSL used approximate half strongly labeled data, and achieved the similar or better performance compared to the conventional method. The proposed SSL can reduce the difficulty of dataset preparation without decreasing performance.

5.3. Evaluation of orientation estimation using video sequence

To demonstrate the performance of the proposed joint estimation method, 59 bicyclist image sequences are prepared, which includes about 2000 images. The experiment results are shown in Table 6. First experiment result is produced without considering any constraint. The second experiment uses temporal constraint, the third experiment is conducted using both temporal constraint and human physical model constraint. Three factors are considered in the evaluation. The head overlap ratio indicates the performance of head localization, is calculated as following equation:

\[
\text{Head overlap ratio} = \frac{(r_{\text{gt}} \cap r_{p})}{(r_{\text{gt}} \cup r_{p})}
\]

where \(r_{\text{gt}}\) is the area of the ground truth of head, and \(r_{p}\) denotes the predicted area of head. Besides the head overlap ratio, the classification rate of head orientation and body orientation are calculated as well. In calculation of the classification rate, we still consider an orientation estimation to be correct if the predicted orientation is identical to the true orientation or one of its direct neighboring orientation, which is adopted in the related work (10) and last subsection.

In Table 6, the second column denotes head overlap ratio, and the third and fourth columns indicate the classification rate of head orientation and body orientation are calculated as well. In calculation of the classification rate, we still consider an orientation estimation to be correct if the predicted orientation is identical to the true orientation or one of its direct neighboring orientation, which is adopted in the related work (10) and last subsection.

In Table 6, the second column denotes head overlap ratio, and the third and fourth columns indicate the classification rate of head orientation and body orientation are calculated as well. In calculation of the classification rate, we still consider an orientation estimation to be correct if the predicted orientation is identical to the true orientation or one of its direct neighboring orientation, which is adopted in the related work (10) and last subsection.

In Table 6, the second column denotes head overlap ratio, and the third and fourth columns indicate the classification rate of head orientation and body orientation are calculated as well. In calculation of the classification rate, we still consider an orientation estimation to be correct if the predicted orientation is identical to the true orientation or one of its direct neighboring orientation, which is adopted in the related work (10) and last subsection.

Compared to the conventional supervised learning method, the classifier generated from SSL maintains or even improves the classification accuracy in head and body orientation estimation. The classification accuracy is the one factor to evaluate the effectiveness. There is another difference between the SSL and the conventional supervised learning method, which is the number of strongly labeled data used in training, as shown in Table 4. The experiment result demonstrated that the SSL used approximate half strongly labeled data, and achieved the similar or better performance compared to the conventional method. The proposed SSL can reduce the difficulty of dataset preparation without decreasing performance.

Table 4. The number of images in dataset for SSL evaluation

| Training dataset for supervised learning (All data is strongly labeled) | Head | Body |
|--------------------------|------|------|
| Strongly labeled data    | 4017 | 4502 |
| Weakly labeled data      | 3234 | 2536 |
| Test dataset             | 3482 | 4644 |

Table 5. Recognition rate of conventional method and semi-supervised learning

|                      | Head | Body |
|----------------------|------|------|
| Conventional supervised learning | 90.9% | 98.2% |
| Proposed Semi-supervised learning    | 91.7% | 98.3% |

Fig. 8. Orientation estimation result on image sequence, the rectangle around of head means head localization result, and the arrow starts from rectangle denotes head orientation, and the arrow starts from center of image indicates body orientation.

We mask the face part in the figure because of the requirement of privacy protection, the pose estimation results were generated using original images without mask.
temporal constraint and human model constraint. The experiment results on last two frames demonstrate that the temporal constraint can improve the head localization accuracy, and the head orientation can be further corrected by human physical model constraint.

6. Conclusions and Future Works
This paper proposed a behavior analysis framework for bicyclist. The recognition of bicyclist and the pose estimation are successively realized in this framework. In the recognition phase, a cascade structure classifier is proposed to distinguish bicyclist form pedestrian, a high classification performance is achieved by deeply analyzing the discriminative area in the second cascade layer. This experiment result suggests that the usage of multiple features and the discriminative area can improve the recognition performance. In the pose estimation phase, a novel Semi-Supervised Learning is employed for data labeling. The experiment result demonstrates that Semi-Supervised Learning is beneficial to obtain accurate classifier in multi-class classification problem. In addition, the temporal constraint and human model constraint contribute to the stable and reasonable pose estimation result in image sequences. All of the experiment results prove that the proposed method is effective.

Even, the classification results of head orientation and body orientation are higher than 90%, but the localization overlap ratio of head is lower than 70%, this 30% error possibly causes the failure in orientation estimation. Therefore, one future work is to improve the head localization. In addition, the system performance at different conditions will be investigated. Moreover, other important information will be estimated from images in the future, such as velocity of road users, and other type of road users will be considered in this framework, such as motorcyclist.

Acknowledgements
The authors gratefully acknowledge the financial support of Semiconductor Technology Academic Research Center (STARC). This research is permitted by the Compliance Committee of the University. In this paper, the faces of pedestrians and cyclists were masked by following the privacy protection rules of the Society of Automotive Engineers of Japan.

Reference
(1) H. K. Kim, T. Takahashi, and S. Kamijo, “Pedestrian Categorization Using Heterogenous HOG Cascade and Motion Difference Silhouette,” Transactions of Society of Automotive Engineers of Japan, Vol.44, No.2, pp. 573-578 (2013).
(2) M. Enzweiler, and D. M. Gavrila, “Monocular Pedestrian Detection: Survey and Experiments,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 31, No. 12, pp. 2179-2195 (2009)
(3) D. Geronimo, A. M. Lopez, A. D. Sappa, and T. Graf, “Survey of Pedestrian Detection for Advanced Driver Assistance Systems,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 32, No. 7, pp. 1239-1258 (2010)
(4) P. Dollar, C. Wojek, B. Schiele, and P. Perona, “Pedestrian Detection: An Evaluation of the State of the Art”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 34, No. 4, pp. 743-761 (2012)
(5) N. Dalal, and B. Triggs, “Histograms of Oriented Gradients for Human Detection,” IEEE Conference on Computer Vision and Pattern Recognition, Vol. 1, pp. 886-893 (2005)
(6) P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, “Object Detection with Discriminatively Trained Part-Based Models,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 32, No. 9, pp. 1627-1645 (2010)
(7) T. Watanabe, S. Ito and K. Yokoi, “Co-occurrence Histograms of Oriented Gradients for Pedestrian Detection,” Advances in Image and Video Technology, Springer Berlin Heidelberg, LNCS 5414, pp. 37-47 (2009)
(8) H. Cho, P. E. Rybski, and W. Zhang, “Vision-based Bicyclist Detection and Tracking for Intelligent Vehicles,” 2010 IEEE Intelligent Vehicles Symposium (IV), pp. 454-461 (2010)
(9) H. Cho, P. E. Rybski, and W. Zhang, “Vision-based Bicycle Detection and Tracking using a Deformable Part Model and an EKF Algorithm,” 2010 13th IEEE Conference on Intelligent Transportation Systems (ITSC), pp. 1875-1880 (2010)
(10) A. Schulz, N. Damer, M. Fischer, and R. Stiefelhagen, “Combined Head Localization and Head Pose Estimation for Video-Based Advanced Driver Assistance Systems,” Pattern Recognition, Springer Berlin Heidelberg, LNCS 6835, pp. 51-60 (2011)
(11) A. Schulz, and R. Stiefelhagen, “Video-based Pedestrian Head Pose Estimation for Risk Assessment,” 2012 15th IEEE Conference on Intelligent Transportation Systems (ITSC), pp. 1771-1776 (2012)
(12) T. Gandhi, and M. M. Trivedi, “Image based Estimation of Pedestrian Orientation for Improving Path Prediction,” 2008 IEEE Intelligent Vehicles Symposium (IV), pp. 506-511 (2008)
(13) R. Labayrade, D. Aubert, and J. Tarel, “Real Time Obstacle Detection in Stereo vision on Non Flat Road Geometry through ‘Disparity’ Representation,” 2002 IEEE Intelligent Vehicles Symposium (IV), Vol.2, pp. 17-21 (2002)
(14) Y. Shibayama, H. K. Kim, K. Fujimura, and S. Kamijo, “On-board Pedestrian Detection by the Motion and the Cascaded Classifiers”, International Journal of Intelligent Transportation Systems Research, Vol. 9 No. 3, pp. 101-114 (2011).
(15) S. Yano, Y. Gu and S. Kamijo, “Estimation of Pedestrian Pose and Orientation Using on-Board Camera with Histograms of Oriented Gradients Features,” International Journal of Intelligent Transportation Systems Research, DOI 10.1007/s13177-014-0103-2, pp.1-10 (2014)
(16) C.-C. Chang and C.-J. Lin, “LIBSVM: a Library for Support Vector Machines,” ACM Transactions on Intelligent Systems and Technology, Vol. 2, No. 3, pp. 1-27 (2011)
(17) S. Yano, Y. Gu and S. Kamijo, “Pedestrian Orientation Estimation using On-board Monocular Camera with Semi-Supervised Learning,” JSAE Annual Congress Proceedings, No.49-14, pp.35-41 (2014)
(18) O. Chapelle, B. Scholkopf and A. Zien: “Semi-Supervised Learning”, MIT Press, (2006)