

High-precision bicycle detection on single side-view image based on the geometric relationship



Yen-Bor Lin*, Chung-Ping Young

Department of Computer Science and Information Engineering, National Cheng-Kung University, No.1, University Road, Tainan City 701, Taiwan, ROC

ARTICLE INFO

Keywords:

Bicycle
Computer vision
Ellipse detection
Triangle detection
Geometry
Algorithm

ABSTRACT

Improving the safety of transportation systems attracts lots of attention. Researchers introduced their methods to detect and analyze the vehicle and the pedestrian on the road to accomplish this goal. However, the bicycle is also a significant factor of the safety on a road. In this paper, a bicycle detector for side-view image is proposed based on the observation that a bicycle consists of two wheels in the form of ellipse shapes and a frame in the form of two triangles. Through the proposed triangle detection algorithm, the bicycle model and the geometric constraints on the relationship between the triangles and ellipses, the computation is fast according to the sample implementation and the evaluation of the reduced data amount. Besides, the training process is unnecessary and only single image is required for our algorithm. The experimental results are also given in this paper to show the practicability and the performance of the proposed bicycle model and bicycle detection algorithm.

1. Introduction

The vehicle and pedestrian detection attracted lots of attention to improve the safety of transportation systems. However, bicycles are also the significant factor of the safety on road. They share the road with vehicles and their speeds are comparable. According to the recent studies [1], the bicyclists are vulnerable and accounted for the high percentage of traffic fatalities. Therefore, it is important to invent a high accuracy bicycle detection algorithm.

Vision-based methods face the diversity of appearance and the variety of illumination condition. There are two categories of vision-based methods to detect a specified object. First, researchers proposed a set of feature vectors these were applied to train a classifier, and then used it to estimate whether the region of interest exists the object. Second, certain geometric properties were employed to detect the target. They discovered the shape, the color and the texture characteristics in common. In several years ago, Felzenszwalb et al. demonstrated a part-based model which outperformed the single template model by using a latent SVM with a variation of the histogram of oriented gradients (HOG) [2] features. As the computation technology steadily progresses, the part-based multi-class object detection [3,4] became popular in recent years. Although this kind of method is capable of detecting all classes of objects, the accuracy of the class-specific object detection is limited and is very computationally expensive. The shortcoming is that the detection accuracy depends on the

collected training data.

According to our observation, a bicycle consists almost of two wheels and two triangles in a side-view image. To the best of our knowledge, unexpectedly, this geometric relationship has not been used to detect a bicycle. Thus, a bicycle detection algorithm and a geometric model of bicycle for single side-view image are proposed and introduced in this article. The detected bicycle model on the single side view image is demonstrated in Fig. 1 as the running example in this paper. The advantage of this proposed method is that this provides insight into the processes involved in the human perceptual system instead of the considerable training process of the classifiers. The elliptic and triangular shapes are chosen because a triangle is still a triangle and an ellipse is almost still an ellipse after the perspective projection produced by the acquisition devices in the real world. The shape property of a bicycle is used to build the geometric model and the experimental results will be given to show the practicability and performance.

The remainder of this paper is organized as follows. Section 2 introduces the related works on the detection of the bicycles, the wheels and the triangles. The main contribution will be described in Section 3 including the proposed bicycle geometric model and the algorithm to detect the bicycle in a side view image. Experimental result will be given in Section 4 to show the practicability of the proposed model and the comparable performance using the implemented algorithm. At the end of the article, we conclude this work and state the future work.

* Corresponding author.

E-mail addresses: yen_bor@yahoo.com.tw (Y.-B. Lin), cypyong@mail.ncku.edu.tw (C.-P. Young).



Fig. 1. The detected bicycle model on the single side view image.

2. Related work

Rogers [5] proposed a bicycle counting system by extracting and tracking the blobs, then judged whether the moving blob is a bicycle by a template-based Hausdorff matching approach. It is required to pre-define a suitable set of templates for this kind of system to match different types of bicycles and it is computationally expensive.

Zhang et al. [6] proposed a wheel detector based on the improved codebook background segmentation. They chose the background model according to the light condition, that is the averaged gray value, added 3 new variables to the codewords, and then the background model was trained by the first 50 frames. The contour of the binarized image was translated to the Freeman chain code [7] for deriving the area and the perimeter of the estimating circle. Note that the circular Hough transform was used only to verify the detected circular contour. This is a general wheel detection method but the performance is sensitive to the light condition and does not resist the complex background.

Qiu et al. pointed out that the distance between points belongs to the same object does not change along time in the article [8]. They described a corner detection method based on the maximum of the Beaudet value on the horizontal edges instead of the whole image. The corners were treated as the feature set to track. The tracked corners were grouped according to their observation. The study proposed a two-layered judgment method. The moving object was estimated as a vehicle when the area of the grouped points was large. Otherwise, the second layer judgment would recognize the moving object as a bicycle when the average velocity of the grouped points was 5 times faster than others. The main errors include over-segmentation, under-segmentation and miss-segmentation. Based on the idea, Fujimoto [9] applied the optical flow algorithm to estimate the area of a moving object. The work also evaluated the width of the tire to estimate the tire angle and this study helped to estimate the traffic direction of a bicycle. Unfortunately, these detection after tracking based method do not account for the static objects.

Takahashi et al. [10] further employed the special motion pattern to judge whether the moving object is a bicycle. The moving object was firstly extracted by the proposed spatiotemporal 3D Gabor filters and the 3rd dimension is the going time. The velocity was also obtained. In order to discriminate the pedestrian from the two-wheel vehicle, a 2-class linear SVM with soft margins was used with the normalized shaped-based feature, HOG. In advance, the bicycle was detected by analyzing the difference of velocities and the existence of pedaling movement with the vertical moving direction relative to the object motion.

Lee et al. proposed the modified HOG block feature [11] and used the Adaboost method, instead of the SVM classifier, to detect the two-wheeler. Jung et al. also proposed an improved HOG named Multiple-Size Cell HOG (MSC-HOG) [12] and used the RealAdaBoost algorithm [13] uniting weak classifiers to detect the bicycle. Note that the bicycle

and the motorcycle were not separated. Kocamaz et al. proposed a multi-cue detection algorithm to detect the cyclists and count them as they crossed a virtual line segment [14]. They used the optical flow computed by the Horn–Schunck's method [15] to extract the foreground object in a video, the flow direction of the SURF [16] feature points, and labeled the trajectories for the cascaded classifier. Somasundaram et al. [17] compared the proposed Pyramidal HOG (PHOG) [18] feature with the SIFT [19] feature for the SVM classifier to detect the bicycle. The classification feature Multi-scale Block Local Binary Pattern (MBLBP) of the moving object was fed to the cascaded classifier trained by the AdaBoost algorithm to detect the bicycle in [20]. Yang et al. [21] proposed a two-stage multi-modal, which is for the three poses, bicyclist detection scheme on the motion based region of interest. They implemented the AdaBoost based algorithm to filter out the negative candidates according to the fast HOG feature and the Extreme Learning Machine (ELM) to detect the bicyclist. Shahrazi et al. [22] combined the HOG, Histogram of Shearlet Coefficients (HSC) and Multi-scale Local Binary Pattern (MLBP) features to improve the detection accuracy of bicycles. Tian and Lauer built different classifiers for eight bicycle orientations and revealed the relationship between the size and the position of the cyclist in the image based on the intrinsic and extrinsic parameters of the calibrated camera [23,24]. In their work, the extra camera calibration step is required. Furthermore, for an already existed image, the parameters and the setup configuration of the camera is unknown. Gu and Kamijo used the CoHOG [25] feature only on the lower part of the blob because of the expensive computational cost to verify the bicyclist detection, and proposed a semi-supervised learning (SSL) to estimate the orientation [26]. Laptev [27] boosted the different combinations of the histogram features for the objects in different classes in the BHog method. He also artificially enlarged the positive training set by adding noisy annotation to reduce the requirement of original training set. Wang et al. utilized the And–Or Graph [28] to make the rare object borrow the feature of the patches from other classes of objects for the training and detection by the dynamic structure optimization.

To overcome the high computational cost, the tracking after detection method became popular. The works in [29,30] were the examples to detect the bicycle with a deformable part-based model and to apply the extended Kalman filter for tracking. The features they used were the Principal Component Analysis (PCA) version of HOG feature and the integral HOG feature in the two works respectively. Chen et al. utilized the extra two near-infrared lighting and one capturing devices mounted on a driving vehicle to acquire the video in the nighttime, and they proposed the appearance-based and edge-based feature for the part construction and selection [31]. The used edge-based and appearance-based features were the Boundary Fragment Model (BFM) [32] and the HOG feature. However, the extra annotation of 13 keypoints on every training sample was required and this work focused on the distinction between bicycle and motorcycle. Yan et al. proposed the two groups of multi-Gaussian models to adaptively choose the threshold of the velocity and the size classifiers for each pixel because the speed and the size of the bicycle are different in different positions [33]. Messelodi et al. [34] focused on the distinction between bicycle and motorcycle. The wheel regions were analyzed in further. This would be a front or rear view image when the angle of moving direction with respect to the y -axis is smaller than a threshold. For the front and rear view images, their system judged the object as a bicycle if the thickness of its tires is smaller than a threshold. For the side-view images, the average illumination inside the detected wheels was compared to that of the background. If the average illumination was similar, the object was a bicycle. In their work, there were two SVM classifiers for the training of the thresholds for the view and object classifications individually. To reduce the computation cost, Nilsson et al. [35] analyzed the prior knowledge of the occurrence of the top center and down center parts of bicycles. They reduced the search space for the computation of the classifier cascade based on the split up

SNoW with local SMQT features.

The other choice to overcome the computational bottleneck is to speed up the computation by hardware techniques. For example, Moro et al. [36] implemented the pixel-wise feature on the GPU platform and used the mean-shift clustering method to perform the detection. Their main idea was similar to the work in [33].

Dukesherer and Smith [37] mentioned that a bicycle is composed of wheels, drive train and a frame. They proposed a three arc-based template to compute the forward directed Hausdorff distance and used this information to confirm that there is a bicycle detected by the circular Hough transform. The disadvantage is the limited usage only for a clean imagery and the template is not adaptive to the variety of natural scenes. Besides, the frame was not adopted to identify the bicycle in their algorithm although they mentioned it. To detect the wheels, Leung and Huang [38] introduced their work of great originality. They used a correlation algorithm to obtain corresponding edge points and estimated the disparity values by the triangulation formula for the estimation of the wheel planes. They transformed the ellipses on the wheel planes into the circles and then applied the Hough transform to derive the locations of wheels. Ardeshiri et al. [39] proposed the most similar idea to build a bicycle model that is a reasonable geometric relationship of the two ellipses those are the two wheels. This work assumed that there is only one bicycle in the image but this assumption is not applicable in real world. Furthermore, they still did not mention the geometric relationship of the frame and the two wheels.

The previous bicycle detection methods need a period of video sequence or a huge number of annotated images for the training and learning process. In runtime, they extract the moving object and then classify them. However, the nature made their computational cost much. The main disadvantages of the training based object detection methods are the requirement of collecting a large number of images annotated with the target object bounding boxes, the dataset bias and the costly rounds of hard negative learning. The further description of the disadvantages of the training based object detection methods could be found in [40]. In addition, the detected circles and ellipses do not mean the bicycle wheels exactly because there are still lots of other objects in this form of shape. Rosin and Ellis built the FABIUS system [41] based on their proposed shape-based templates to describe the appearance between the two wheels and the probabilistic matching called Subjective Bayesian Updating to detect the vehicles in an image. Our work was motivated by the FABIUS system to inspect the shape-based structure between the wheel pair. The main difference is that the matching step was improved with the proposed mathematically geometric model in our work. We introduce and propose a simple and robust bicycle detection algorithm based on the geometric insight for single side-view image. Without the demand of training process, the accuracy rate of this proposed algorithm is still acceptable and comparable as shown in Section 4.

Shape based method is a good choice for the detection of man-made objects. For example to detect the traffic sign [42,43], Loy and Barnes [42] presented the method using the symmetric nature of the triangular, square and octagonal road signs as well as the known number of sides. The fast radial symmetry transform [44] was extended for the shape detection. Unfortunately, they assumed that the traffic sign is parallel to the image plane that is not applicable in the case of the bicycle detection task. Ahan et al. used the color clues and carried out the geometric property analysis, e.g., the ratio between the square of the perimeter and the area and the number of sides, to detect the road signs and applied the distortion invariant fringe-adjusted joint transform correlation (FJTC) based similarity measure to recognize the road signs [43]. Our work was also motivated by [43] to analyze the geometric relationship of the bicycle in a side view image. The two mainly used shape are the triangle which is the component of a bicycle frame and the ellipse which is the form of the two wheel after the perspective projection from the real camera model. The related triangle

and wheel detection algorithms are also introduced in the following paragraphs.

The wheel detection algorithms [6,38,45] proposed by researchers in the past were based on the assumption that the elliptic object detected in the road image is a wheel such that the wheel detection problem was reduced to the ellipse detection problem. The bicycle wheel detection method in [46] extracted the dark-gray-color model patches and validated the wheel candidates by the three constrains including the constant color region inside the wheel, the range of size and the ratio of the width to the height. In the recent years, there are popular algorithms for the examples as published in [47–51] to detect the ellipse. The randomized Hough transform (RHT) [47,48] improved the computational efficiency of the original Hough transform based ellipse detection. Zhang and Liu decomposed the parameter space of the Hough transform to achieve computational efficiency [49]. Libuda et al. proposed a bottom-up method to group the extracted elements into an ellipse stage by stage based on their relationship [50]. In [51], the method used the edge curvature and the convexity in relation to other edge contours as clues for identifying edge contours that can be grouped together. In the work, a two dimensional Hough transform with a relationship score was still performed in an intermediate step. The existing algorithms to detect the ellipse performed well. However, the assumption is not always true because there are lots of objects in the form of an ellipse or a circle on the road.

The triangle detection algorithms could be found in the literatures [52–54]. An intuitive method to detect the triangle is to segment the image into regions according to the known color information [55] of the target object and to check the amount of different gradient angle values on the boundary of the regions. This kind of implementation could be found in [54]. However, the color information and the pattern on the bicycle were unknown and various such that the method could not be applied in this work. Besides, He and Ma [52] used Hough transform to extract the straight line segments and utilized the intersection of angular bisector to find the center of the inscribed circle of the triangle. Then, the coordinates of three vertices of the triangle were determined. With those information the triangle could be detected. They applied their proposed algorithm on the traffic sign detection to show the robustness to the rotation invariance and the size scalability. Lv and Luo proposed the triangle detection algorithm using the HSI color model and the Radon transform which detects the straight lines [56]. In their implementation, the sum of the three intersection angels was considered to consist a triangle. Liu and Wang [53] pointed out the property of arbitrary triangle that the distance of the incenter to the edges of the triangle is equidistant and equal to the radius of the inscribed circle. Based on the property, the proposed energy distribution map and the gradient orientation information were adopted to detect the triangle.

The triangle detection and ellipse detection were usually treated respectively because of the different applications. However, for the proposed bicycle detection algorithm, both of them are required to be detected efficiently. In order to make the detection of these two kinds of shapes rapid, we proposed an efficient triangle detection algorithm as described in Section 3.4 according to the ellipse detection framework proposed in [57] to accomplish this mission to extract the possible triangle pairs for the bicycle frame and the ellipse pairs for the wheel pairs simultaneously.

3. Methodology

Fig. 2 shows the software blocks to detect a bicycle overall. The purple blocks are the proposed work flows in this paper. In our proposal for this sample implementation, the width and the height of the search region to detect the frame of the bicycle are determined by the distance between the two detected ellipses and the length of the major axis of the detected ellipses respectively. The relationship between the two triangles is constrained as shown in Fig. 3 and will

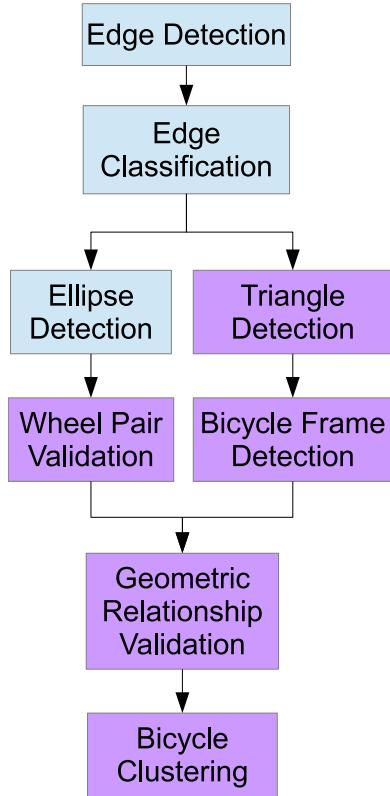


Fig. 2. The flow chart of the overview to detect a bicycle.

be introduced in the following sections. Then, the relationship among the detected frame which is composed of the two triangles and the two detected wheels these are the two ellipses is proposed to make sure that the two detected ellipses are the pair of wheels of a bicycle. In the following section, the running example as shown in Fig. 4 is given for readers to visually understand the proposed algorithm.

3.1. The wheel detection – the ellipse detection

The ellipse detection algorithm in [57] is used to detect the wheel candidates in the image. In order to get the whole view of our proposed algorithm, the ellipse detection algorithm is briefly introduced here. The algorithm classifies arcs into four types of arcs and evaluates the possibility to consist an ellipse by the combinations of three arcs in different types. The arcs are classified by the slope and the convexity. Straight afterwards, all possible triangles in an adaptively defined

search region are extracted and evaluated to compose the frame of the target bicycle according to the geometric relationship and the constraints described in the following sections.

3.2. The wheel pair

There are lot of objects in the form of ellipse shape in the world. To filter out unreasonable wheel candidates, the distance between two wheels of a bicycle is constrained as being described based on the size of the ellipses in Section 3.6. This constraint roughly improves the estimation of the likelihood of the wheel detection. An ellipse pair will be confirmed as a wheel pair according to the geometric relationship with respect to the bicycle frame in advance.

3.3. The search region

In order to reduce the computation, we just consider the straight lines in the proposed search region. It is inside the region bounded to the dotted gray lines as shown in Fig. 3. Fig. 4(c) shows the ellipse candidates. The lower boundary is the connected line between the two centers of the considering ellipse pair. The side boundaries are the extension of the two major axes of the considering two ellipses and the lengths are 1.5 times from the centers upward. The upper boundary is the straightly connected line of the two upper endpoints of the two side boundaries.

In further, another search region was proposed to consider the straight and curve contours for the front component of the bicycle frame. For example, the straight line was considered to represent the down tube and the curve or straight contour was considered to represent the top tube. This search region is bounded by the darker gray lines as shown in Fig. 3. The lower boundary is still the connected line between the two centers of the considering ellipse pair. The side boundaries are the extension of the seat tube and the head tube. The upper boundary is the parallel line of the connected line between the two centers of the ellipse pair.

3.4. The triangle detection

We need a triangle detection algorithm in order to extract the bicycle frame efficiently. Hough transform is a well-known method to detect the generally regular shapes including the circle, the ellipse and the triangle. A Hough Transform based method votes on the parameter space by every edge pixel and the parameter suits having the maximum voted values would be the detection result. However, the performance in computational efficiency and memory consumption will be the bottleneck of whole system. Therefore, we proposed a triangle detec-

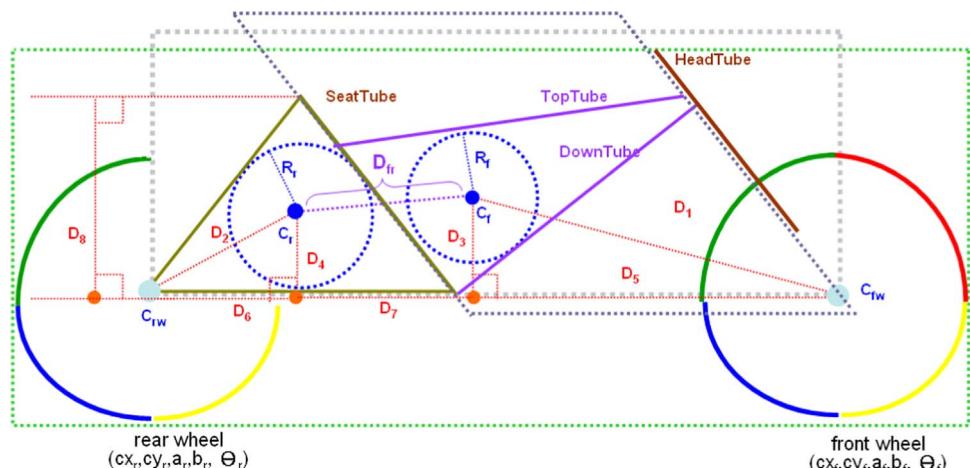


Fig. 3. The bicycle model defining the geometric relationship between the frame and the wheels.

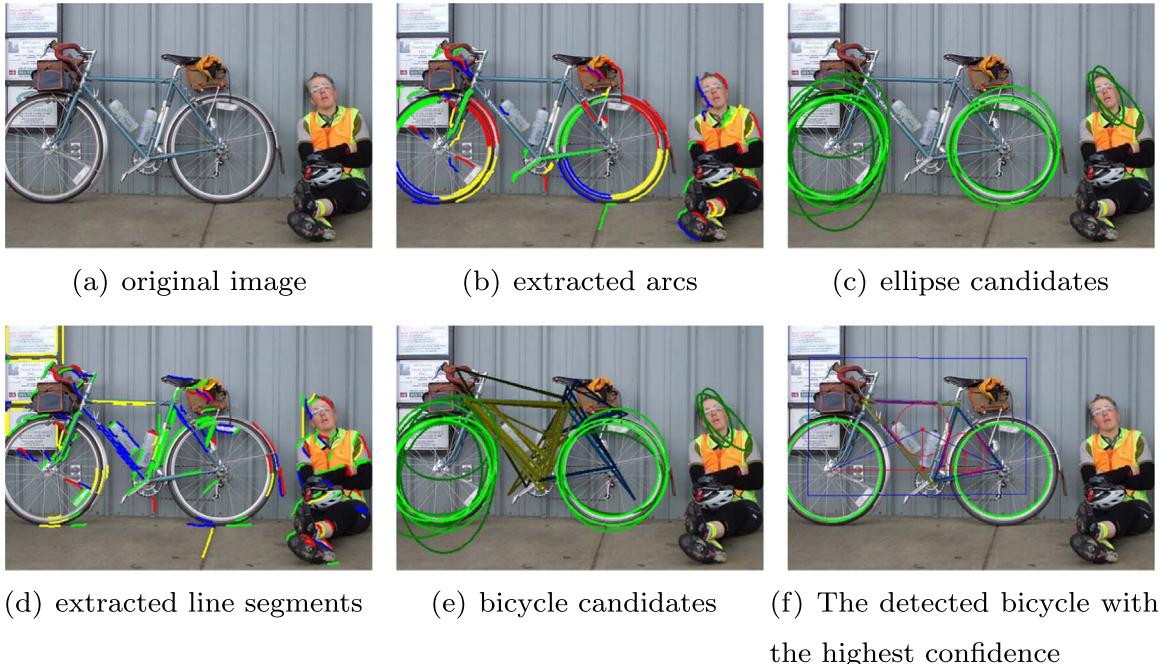


Fig. 4. A running example for the proposed bicycle detector.

tion algorithm according to the ellipse detection framework proposed in [57] to obtain the triangle candidates and used the triangularity metric [58] to extract the triangles to accomplish this mission for the bicycle frame detection.

3.4.1. The extraction of straight lines

First of all, the straight lines are the basic components to consist a triangle. A straight line is still a straight line after the perspective projection produced by the acquisition devices in the real world. However, the straight segments were dropped during the phase to extract arcs as shown in Fig. 4(b) in the original ellipse detection process [57] according to the threshold of the length for the shorter side of the contour bounding box. Obviously, the straight segments are required in our bicycle detection algorithm. We pick up them in the search region described in Section 3.3. In the original process, the segments would be divided into two types those are positive and negative diagonal segments. We add two new types of segments, including horizontal and vertical segments, and there are totally four types of straight segments extracted as shown in Fig. 4(d). They are classified according to their slopes. For the horizontal and vertical segments, a small tolerant range of the calculated slopes is defined to avoid the digital quantization error. As shown in Fig. 5, the segments with the calculated slope between the tolerant ranges defined by $0 \pm \theta_e$ are treated as horizontal line segments. In the same way, the segments are treated as vertical line segments by $90^\circ \pm \theta_e$. In addition to the horizontal and vertical segments, the other two types are the straight segments having positive slope and negative slope respectively. The extracted straight segments are collected to compose the triangle candidates.

3.4.2. The classification of triangles

In order to make the triangle pairing process more efficient, we detected the triangles by the combination of the types of straight lines. The triangles are categorized into four types named 01NPP, 02PNN, 03NPN and 04PPN as shown in Fig. 6. In this notation, P represents for the straight line segments with positive slope. Oppositely, N represents for the straight line segments with negative slope as shown in Fig. 5. The prefix number is the serial number of the type of triangle and the three letters in postfix are the types of the three consisting straight line

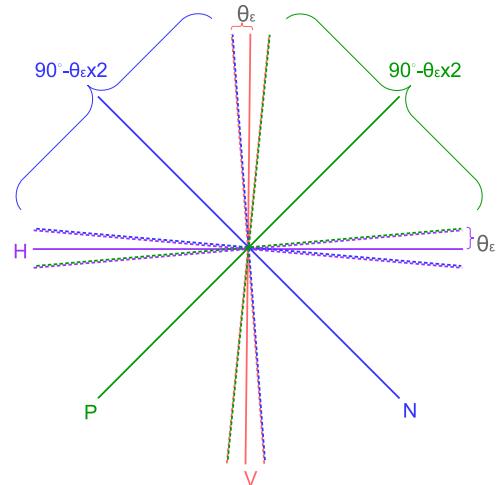


Fig. 5. The tolerant range of the calculated slopes to detect the horizontal and vertical line segments.

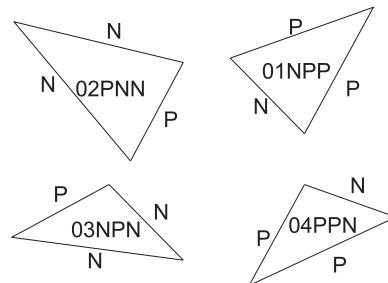
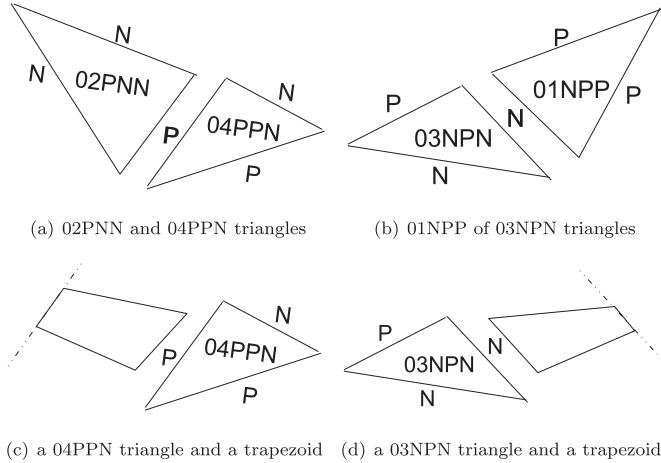


Fig. 6. The classification and the notation of triangles.

segments. The edges of a triangle are coded in counter-clockwise direction and the first one is the candidate of the shared edge for the triangle pairing phase which will be introduced in Section 3.5. This classification step reduces the data of all combinations of the triple straight lines to compose triangles.

In order to remove the duplicate detection of triangles, the detected

**Fig. 7.** The shape pairs to compose the bicycle frame.

triangles are clustered in further. The clustering conditions are based on the distance between the centers of their inscribed circles and the difference in the length of their radii which represent their sizes. The triangles would be merged into a group when their difference ratio is 10% shorter than the other one and the distance is short.

3.5. The bicycle frame detection

Recall that a triangle is still a triangle and a straight line is still a straight line after perspective projection, we proposed a simple method to extract the bicycle frame candidates. This bicycle frame detection method is based on the proposed triangle detection as described in Section 3.4. There are two types of shape pairs to compose a bicycle frame as shown in Fig. 7.

First, there are two conditions to make a pair of triangles. The triangle with the type of 01NPP matches the triangle with the type of 03NPN and the triangle with the type of 02PNN matches the triangle with the type of 04PPN. Second, to make a pair of a triangle and a trapezoid, as shown in Fig. 7(c) and (d) the trapezoid with positive orientation matches the triangle with the type of 03NPN and that with negative orientation matches the triangle with the type of 04NPN. This pairing step reduces the data amount of the shape pairs to compose a bicycle frame. In addition to the diamond frame based bicycle, the Dutch bicycle is another popular style of bicycle. The main feature of the Dutch bicycle is the step-through frame as shown in Fig. 9(e). Therefore, the bicycle tube is not always presented by a straight contour and there is sometimes presented by a curve contour. In order to make the proposed bicycle detection algorithm account robust to the variety of the bicycle frames, the 01NPP and 02PNN triangles could also be consisted by the curve contours, and the longer sides of a considering trapezoid could also be curve contours. This reacts the fact that the shape of top tube and the down tube may be curve or straight. For the curve contours, we make them simulate the straight contours in the way of connecting the endpoints of the contour straightly. Therefore, in the following sections, the curve and straight contours will be illustrated in a uniform way to present the geometric relationship.

3.6. The bicycle model

As mentioned above, a bicycle consists of the two wheels in the form of ellipse shapes and a frame in the form of two triangles usually. Many triangle candidates appear as shown in Fig. 4(e), but there should be exactly one frame, which means two triangles, for a bicycle. The proposed model is illustrated as shown in Fig. 3. We introduce the notation on this illustration as follows. The centers of the inscribed circles of rear and front triangles are notated by C_r and C_f . The radii of

the two inscribed circles are notated as R_r and R_f respectively. The distance between the two centers of detected ellipses is divided by the projected points of C_r and C_f on the straightly connected line between the ellipse centers into three segments as notated from D_5 to D_7 , and the projection distances are notated as D_3 and D_4 . Besides, D_2 denotes the distance between the center of the rear inscribed circle and the center of the rear ellipse. Similarly, D_1 denotes the distance between the center of the front inscribed circle and the center of the front ellipse. In addition, D_8 denotes the projection distance from the top of the rear triangle to the line which connects the two centers in a pair of detected ellipses. Finally, D_{fr} denotes the distance between C_r and C_f . In summary, the distances labeled from D_1 to D_7 , the distance D_{fr} , the radius lengths of the two inscribed circles, these labeled as R_r and R_f , and the minor and major axes, labeled as a_r , a_f , b_r and b_f , of the two ellipses are the parameters to describe a bicycle and these are the model parameters.

In general, both the values of R_r and R_f are smaller than those of a_r , a_f , b_r and b_f . The value of D_5 is almost equal to half of the distance between the centers of the two ellipses because this is preferred to make the center of gravity be located in the lower middle from the viewpoint of the bicycle mechanism design for the balance and the stability of the forces. We formulate these in the following equations. It is worth to note that different combinations of the model parameter values may represent different styles of the bicycle.

$$R_r < \min\{a_r, b_r\} \quad \text{and} \quad R_f < \min\{a_f, b_f\} \quad (1)$$

$$D_5 \simeq \frac{D_5 + D_6 + D_7}{2} \simeq D_6 + D_7 \quad (2)$$

We give the geometric constraints on this bicycle model to eliminate unreasonable bicycle candidates. First, the rear wheel must overlap the rear triangle of the frame as constrained by (3). Second, the sizes of the two triangles of the frame is constrained in a suitable scale with respect to the detected ellipse according to (4). Similarly, the sizes of the two ellipses of a wheel pair are similar in size. Third, the horizontal position of the frame is between the ellipses along the direction of the connection of the two centers, and the vertical position of the frame is above the connection line. The relatively horizontal position has already been constrained by (2) logically. However, the relatively vertical position is now constrained by (5). Finally, the two triangles of the same frame share a common or a parallel pair of straight line. Thus, as shown in Fig. 4(c), the human head in the form of ellipse shape is excluded because there is no pair of ellipse including this head conforms to the proposed geometric bicycle model. Finally, the bicycle with the highest confidence is given as shown in Fig. 4(f):

$$D_2 > \max\{a_r, b_r\} \quad (3)$$

$$\text{ratio}_{LB} < \frac{R_r}{\max\{a_r, b_r\}} < \frac{R_f}{\max\{a_r, b_r\}} < \text{ratio}_{UB} \quad (4)$$

$$D_3 < \max\{D_8, 1.5 \times \max\{a_f, b_f\}\} \quad \text{and} \quad D_4 < 1.5 \times \max\{a_r, b_r\} \quad (5)$$

According to the observation, a bicycle is generally three times as long as the radius of the consisted wheels, and the major axis of the ellipse of a wheel is in the vertical direction. The first constraint named Γ_1 is on the ellipse pairing to present the wheel pair based on the lengths of the semi-major axes of the two ellipses which are almost equal to each other as given in (6) and the distance between the two ellipse centers which is several times as long as their minor axis as given in (7). The second constraint named Γ_2 is that any vertex of the detected triangle locates in the search region which is introduced in Section 3.3:

$$\frac{|\max\{a_r, b_r\} - \max\{a_f, b_f\}|}{\max\{a_r, b_r, a_f, b_f\}} < Th_{\Gamma_{1s}} \quad (6)$$

$$\frac{D_5 + D_6 + D_7}{\min \{a_r, b_r, a_f, b_f\}} < Th_{I_{1d}} \quad (7)$$

The third constraint named Γ_3 is that the pair of triangles share a common edge or a parallel pair of edges. The constraint Γ_4 represents that the position of the inscribed circle of the rear triangle is above the connected line between the two centers of the considering ellipse pair. This constraint cooperates with the equation in (8) and (9) to constrain the relatively vertical position of the front triangle:

$$L_{fw,rw}(C_f.x, C_f.y) \times L_{fw,rw}(C_r.x, C_r.y) > 0 \quad (8)$$

$$L_{fw,rw} : \frac{y - Cy_r}{x - Cx_r} = \frac{Cy_f - Cy_r}{Cx_f - Cx_r} \quad (9)$$

where $L_{fw,rw}$ denotes the connected line of the centers of the ellipse pair. We name the aforementioned D_5 constraint as Γ_5 for easier explanation. According to the observation, the rear triangle of the bicycle frame generally intersects the rear wheel. This constraint named Γ_6 and is given as

$$D_6 < \min \{a_r, b_r\} \quad (10)$$

The bicycle frame takes the force of the bicyclist's pedaling at its lower position. In order to make sure the lower vertices of the triangle pair close to each other, the constraint, Γ_7 , on the position of the down tube is given and is formulated as

$$\frac{\text{ProjDistance}(\text{LowerVertexOfRearTrl}, \text{DownTube})}{2R_r} < Th_{I_7} \quad (11)$$

In physical, the sizes of the triangles in a pair to compose a bicycle frame are almost approximate to each other. The constraint Γ_8 is stated to model this fact and is formulated as

$$\frac{|R_f - R_r|}{\max \{R_f, R_r\}} < Th_{I_8} \quad (12)$$

The proposed 8 constraints are summarized in Table 1 and all the parameters of the constraints of the bicycle model will be evaluated in Section 4.2.1.

4. Experimental result

In order to demonstrate the practicability and the performance of the proposed bicycle detector, the proposed algorithm has been implemented in C++ and the program was compiled on Microsoft Visual Studio 2013. The OpenCV v3.0 library [59] was utilized to adopt the built-in data structures, to apply the Sobel filter, to obtain the gradient values, and to implement the adaptive Canny edge detector. The experiments were conducted on Intel Core i7 870 2.93 GHz processors with 8 GB RAM under windows 7.

The touring-bike and mountain-bike subsets in Caltech256 [60], the 12 subsets related or similar to the shape of bicycle in ImageNet [61] organized according to the WordNet [62] hierarchy, PASCAL Visual Object Classes (VOC2012) [63], CBCL Street Scenes [64] and cityscapes [65] datasets were chosen to evaluate the performance because these datasets provide the ground truth and they are popularly

Table 1

The proposed constraints on the bicycle model.

Constraint	Description in short
Γ_1	The ellipse pairing based on the sizes and the distance
Γ_2	The search region
Γ_3	The common edge of the two triangles
Γ_4	The triangles are the upper part of a bicycle
Γ_5	The relatively horizontal position of the front triangle
Γ_6	The rear triangle intersects the rear wheel
Γ_7	The position of the down tube
Γ_8	The size of the triangle pair

used in the computer-vision domain in recent years. The performance in terms of precision, recall and the F -score which is formulated by the following equation:

$$F_{\text{score}} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (13)$$

where the precision is the ratio of the number of true positive instances these are truly the bicycles, and the recall is the ratio of the number of true positive instances to the total number of bicycles in the dataset. The condition to recognize the detected bicycle as a true positive is the ratio of overlapped area between the corresponding bounding boxes of the detected bicycle with respect to the ground truth annotation. The ratio of overlapped area is calculated as Ratio_{ov} in (14), and the threshold is set to be 70% because the bounding box of our detection result excludes the handlebars, the seat post and the saddle:

$$\text{Ratio}_{ov} = 1 - \frac{A_{\text{det}} \oplus A_{\text{gt}}}{A_{\text{det}} \vee A_{\text{gt}}} \quad (14)$$

where A_{det} denotes the area of the bounding box of the detected bicycle and A_{gt} denotes the area of the bounding box of the ground truth provided from the dataset.

In order to compare the detection performance, the ratio of F_{score} to execution time is proposed too for the consideration in the two viewpoints simultaneously. We denoted this ratio as $FTRatio$ and formulated it by the following equation:

$$FTRatio = \frac{F_{\text{score}}}{\text{ExecutionTime}} \quad (15)$$

The larger value of $FTRatio$ presents the better performance of the bicycle detector. This will be used during the evaluation and comparison with the state of art works in Section 4.2.5.

There are 30 607 categorized images including 192 bicycle images in Caltech256 database which is more pure and suitable for the basic verification of the bicycle detectors with respect to the various classes of natural and man-made objects. VOC2012 contains the released datasets from 2007 to 2012 and the number of evaluated image files is 22 263. ImageNet is similar to VOC2012 but it contains larger number of images than VOC2012. Since the total number of images in ImageNet is more than 1 461 406, we only evaluated the 12 selected subsets related or similar to the shape of bicycle in ImageNet. They are bicycle, tandem, exercise bike, motorcycle, mountain bike, ordinary bicycle, safety bicycle, scrambler, bicycle chain, bicycle seat and bicycle wheel subsets. The selected 12 subsets contain 16 617 images totally. VOC2012 and ImageNet datasets did not focus their collection on the street or the road images such that they might miss certain cases. Therefore, CBCL Street Scenes and cityscapes datasets were also evaluated to show the detection performance in case of the street image. They contain 3547 and 2975 images respectively. In summary, these 76 009 images were evaluated. The experimental results will be introduced in the following sections.

The proposed method focused on the side-view image, and therefore, the front-view and rear-view bicycle images were excluded in the further evaluation and discussion as described in Section 4.2.5. Some samples of the excluded images from the chosen datasets are shown in Fig. 8. After the described images were excluded, there were 30 597 images from Caltech256, 3502 images from CBCL Street Scenes, 2839 images from cityscapes, 15 058 images from ImageNet and 21 993 images from VOC2012 to show the effectiveness and the efficiency of the proposed bicycle detection algorithm.

4.1. The qualitative evaluation

Some visual results are given in Fig. 9 to show the qualitative performance visually. First, Fig. 9(a) and (b) demonstrated the basically expected results in the case that the bicycle are in the left and right direction respectively in the two exactly side-view image

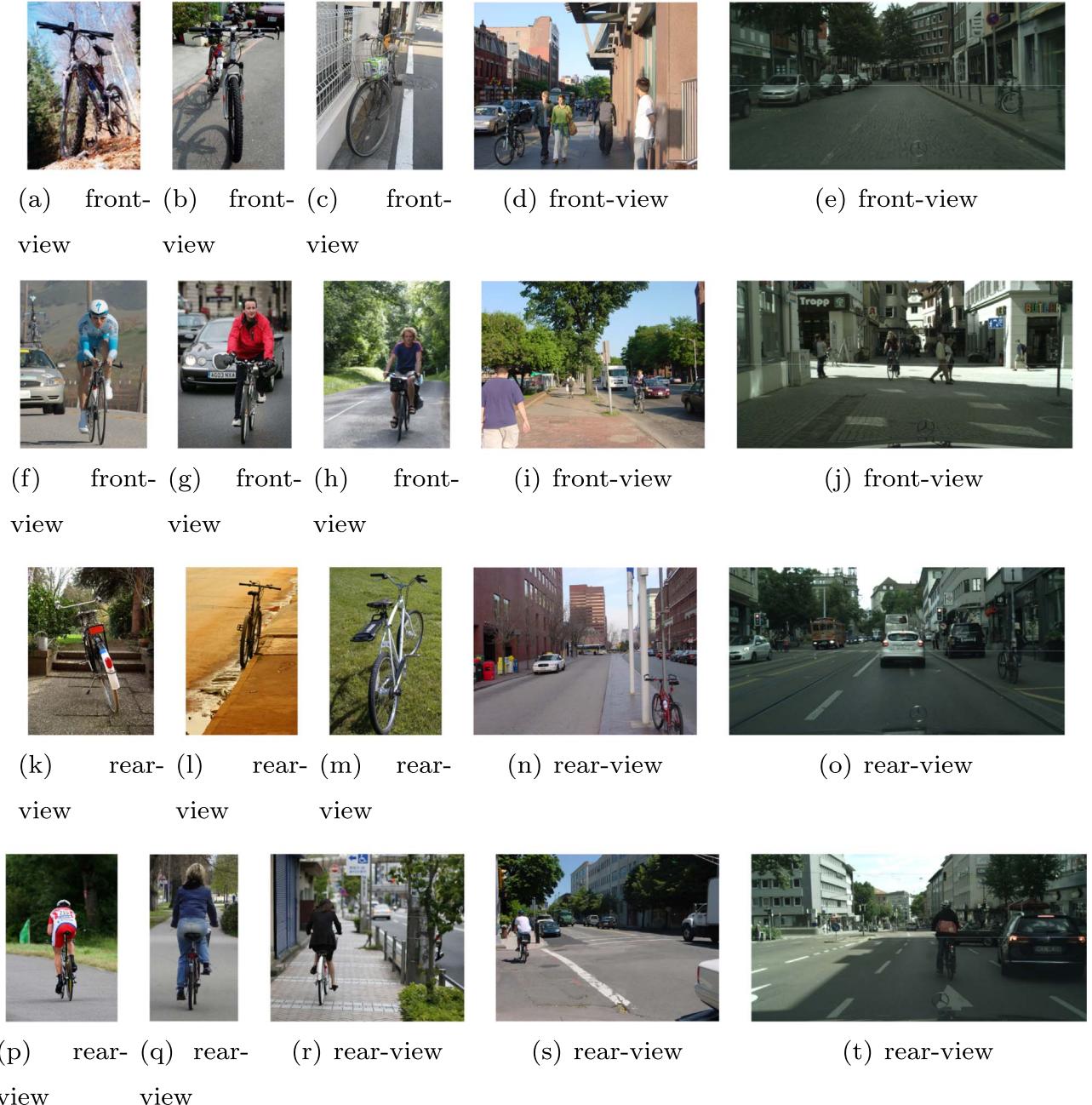


Fig. 8. Some samples of the excluded images from Caltech256, VOC2012, ImageNet, CBCL Street Scenes and Cityscapes.

whose background is clean. Second, Fig. 9(g)–(j) showed that the bicycles in the horizontal and vertical patterns of the background were detected correctly. In further, the bicycles in the mesh pattern of the background were also detected successfully as shown in Fig. 9(n). As we can see, the straight line patterns in the forms of horizontal, vertical and even the mesh did not affect the performance of the proposed method to detect the bicycle because the line segments which do not conform to the proposed bicycle model were eliminated based on the geometric relationship between the triangle candidates and the detected ellipse pairs. Another type of the distortion of the background is the complex patterns, for example lots of various objects and the leaves as shown in Fig. 9(k)–(m), and the bicycles were also successfully detected in this scenario. Besides, the bicycles in a direction of certain angle were also detected as shown in Fig. 9(o)–(s) because the perspective projection preserves the triangle and the circular wheel stays in the form of ellipse. Moreover, the other popular bicycles with special bicycle frames as shown in Fig. 9(f) and the Dutch bicycles as

shown in Fig. 9(c)–(e) could be detected. The informal bicycle frame for the proposed bicycle model was considered and included. In addition, the mixture of challenge situations for example in Fig. 9(r) and (s), the bicycles are also detected correctly. Finally, in the case of the occluded bicycles as shown in Fig. 9(t)–(y), they were also obtained when the three of four arcs to compose each wheel and the bicycle frame could be extracted. Fig. 9(y) also shows the detection result on the child bicycle especially. The visual results conclude that the proposed algorithm is robust to the angle of the bicycle, the complex patterns in the background and the variety of bicycle mechanism. The visual results showed the robustness of the proposed bicycle detector in general situations.

4.2. The quantitative evaluation

The quantization results from our proposed detector over the five datasets are illustrated in Figs. 15–17. Their F_{score} are 0.89, 0.57,

**Fig. 9.** Some visual results to show the qualitative results.

Table 2

The quantitative evaluation of the proposed constraints on the bicycle model.

Constraint	The extraction constraints on the bicycle model							
	Γ_1	Γ_2	Γ_3	Γ_4	Γ_5	Γ_6	Γ_7	Γ_8
Triangle amount %	523 090	400 589	14 056	6496	5356	4756	2726	2686
	100	76.58	2.69	1.24	1.02	0.91	0.52	0.51

0.47, 0.47 and 0.56 in average for the Caltech256, CBCL Street Scenes, cityscapes datasets, the selected 12 subsets in ImageNet and VOC2012 datasets respectively derived by the measured precision and recall. First of all, the evaluation on the bicycle model parameters will be given. Next, the performance of this proposed bicycle detector is related to the accuracy of the used ellipse detector because we define the search region for the frame of bicycle candidates according to the location and the size of the detected ellipses. Therefore, the effect of the accuracy of the ellipse detection will be evaluated in Section 4.2.2. Then, the bicycle frame detection algorithm and the proposed constraints on the bicycle model will be evaluated to show the reduced computational consumption in Sections 4.2.3 and 4.2.4, and the numeric results is summarized in Table 2. The table will be explained in the following section. At the end, the overall performance and the comparison with the state-of-art works, including BHog [27], AndOrG [28], YOLO [66] and DPM [3], will be given to show the effectiveness and the comparable performance of the proposed bicycle detector for side view images (Fig. 10).

4.2.1. The evaluation of the bicycle model parameters

The four model parameters named $Th_{\Gamma_{1s}}$, $Th_{\Gamma_{1d}}$, Th_{Γ_7} and Th_{Γ_8} as described in Section 3.6 influence the detection performance of the proposed algorithm in terms of the F_{score} . The influence of the parameter $Th_{\Gamma_{1s}}$ is investigated in Fig. 11(a). The size constraint on the ellipse pairing was defined as the ratio of the difference in length between the two longer axes to the longest axis in a pair of ellipses. In Fig. 11(a), the best results was obtained when the value of $Th_{\Gamma_{1s}}$ was around 0.5 and the performance becomes worse with the larger difference in size. For another parameter of the constraint on the

ellipse pairing, Fig. 11(b) shows the influences of the distance between the two centers in a pair of ellipses. The detection performance raised slightly when the value of $Th_{\Gamma_{1d}}$ was increasing to enlarge the search region and then fell because of the increasing occurrence of false positives when the value of $Th_{\Gamma_{1d}}$ was larger than 2. These two measured results mean the difference in size between the two ellipses in a pair has to be smaller and the distance has to be within a reasonable range for a physically feasible bicycle in real.

As illustrated in Fig. 11(c), we show the impact of Th_{Γ_7} which was defined as the ratio of the project distance from the lower vertex of the rear triangle onto the down tube to the length of the rear triangle's diameter. This is related to the position of the lower vertices in a pair of triangles and is utilized to constrain the position of the down tube. As this parameter was loosen with the various values from 0.1 to 0.5, the detection performance grew significantly. Oppositely, the performance dropped dramatically when the value of Th_{Γ_7} was larger than 0.55. In Fig. 11(d), the more tolerance of the difference in length between the two radii of the incircles in a pair of triangles was allowed, the detection performance got better. However, there were two peaks in this measured result because the different styles of the bicycles contributed to the two grouped values of the ratio of the two radii of the two incircles in a pair of triangles. Although the larger of the Th_{Γ_8} values for the proposed algorithm gained the more possibility to find different styles of bicycles, detection performance still dropped when the value was too large to obtain too much false positives.

Therefore, the parameters, $Th_{\Gamma_{1s}}$, $Th_{\Gamma_{1d}}$, Th_{Γ_7} and Th_{Γ_8} , of constraints on the proposed bicycle model were set as 0.55, 1.95, 0.55 and 0.5 respectively.

4.2.2. The effect of the accuracy of the ellipse detection

It is worth to note that the evolution of performance from Figs. 12 and 13 demonstrates that the recall of this bicycle detection rises and the precision stays high if we loosen the reliability threshold of the ellipse detection algorithm to obtain more ellipses in the image. The improvement of the value of recall ratio is obtained without the loss of precision because more ellipse pairs detected means more occurrence of the mentioned geometric relationship and more possibility to be verified as a bicycle. By the way, this result also presents the proposed algorithm performs similar to the different types of bicycles. This

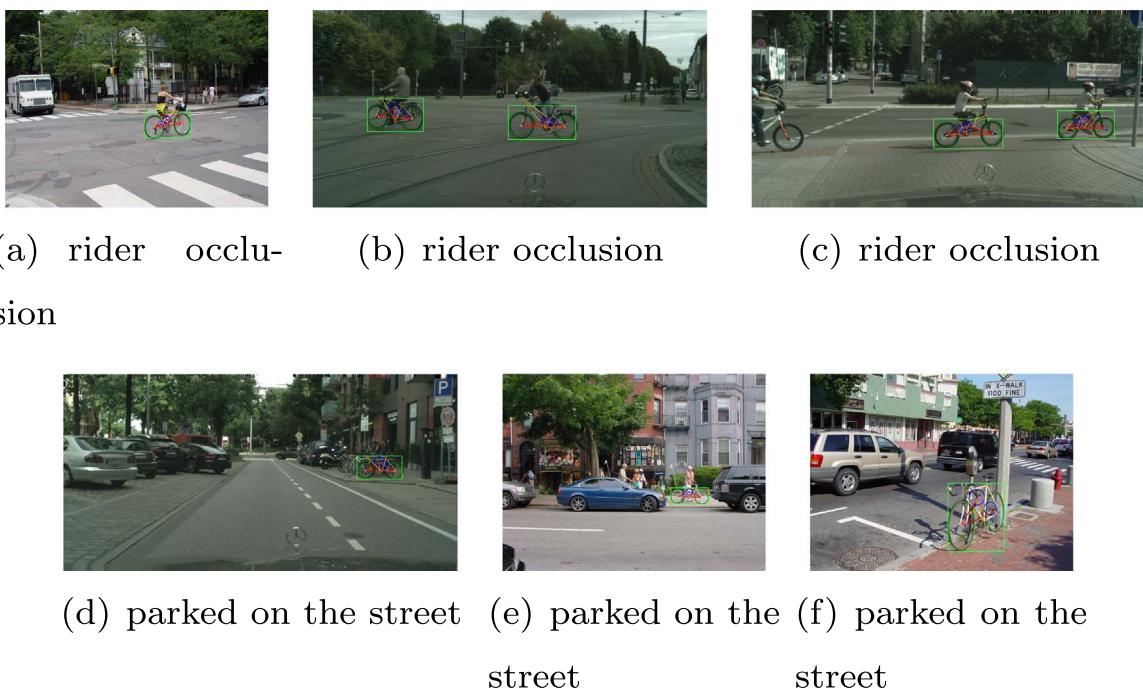
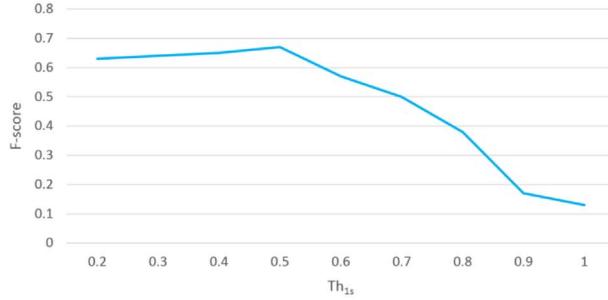
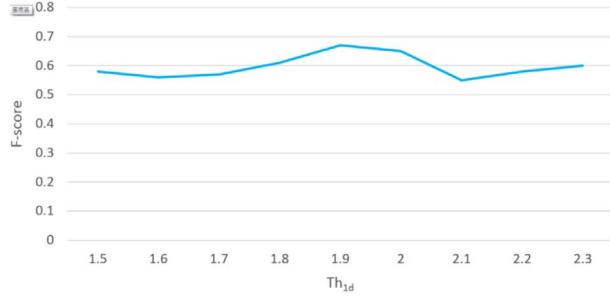


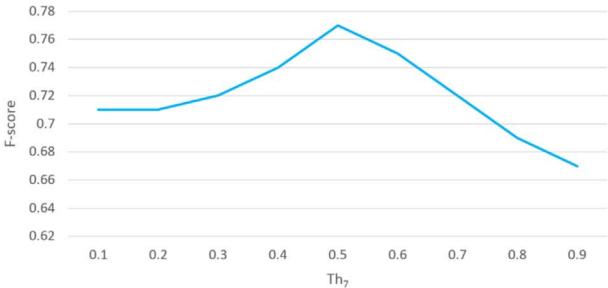
Fig. 10. Some visual results to show the qualitative results on the street images.



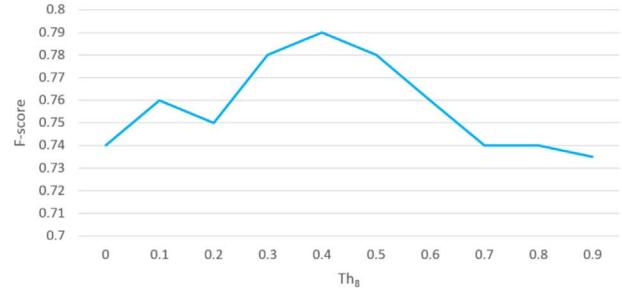
(a) the size of the ellipse pair



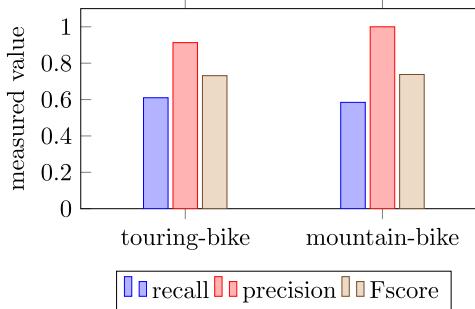
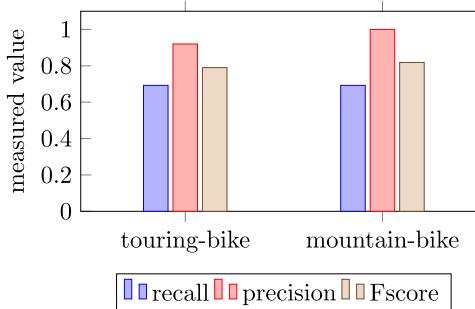
(b) the distance between the centers in a pair of ellipses



(c) the position of the lower vertices of the triangle pair



(d) the size of the triangle pair

Fig. 11. The evaluation of the bicycle model parameters.**Fig. 12.** The quantitative results in terms of recall, precision and F_{score} with the reliability threshold 0.6 for ellipse detection.**Fig. 13.** The quantitative results in terms of recall, precision and F_{score} with the reliability threshold 0.4 for ellipse detection.

experimental result showed the effectiveness of the proposed geometric relationship method to detect the bicycles in the side view images.

4.2.3. The evaluation of the bicycle frame detection

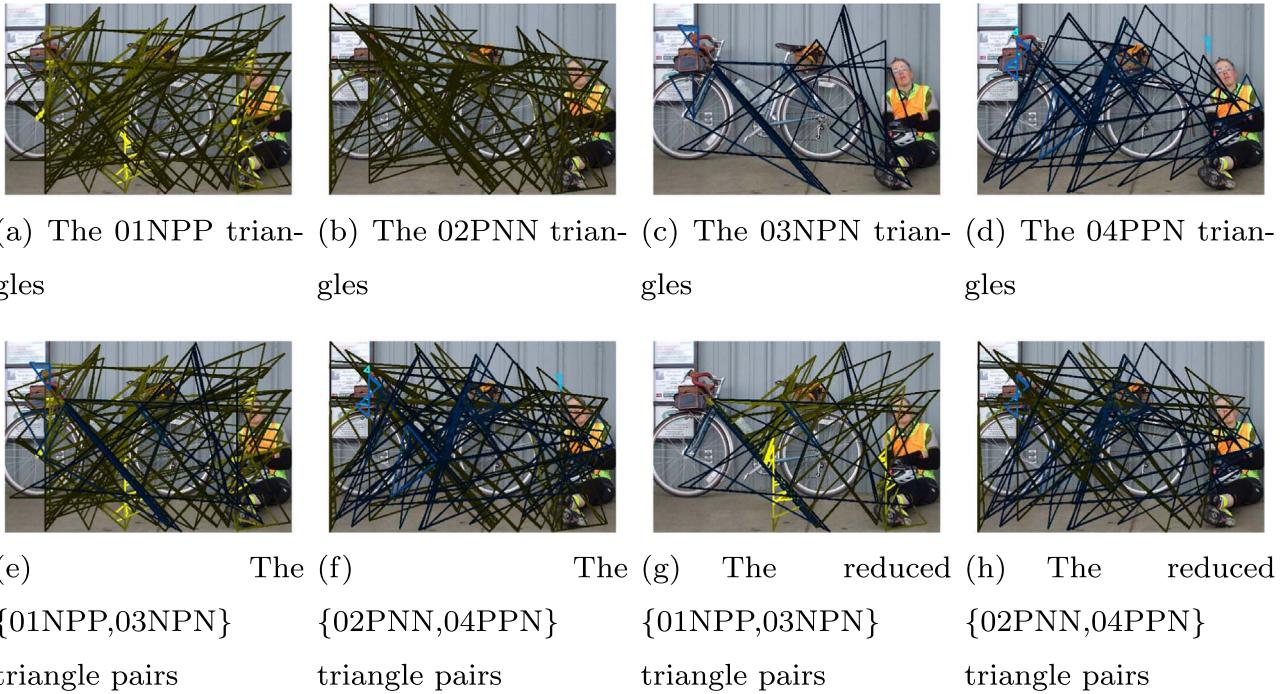
To show the computational efficiency on the proposed triangle

detection algorithm, the data reduction is evaluated. For the running example in Fig. 4(a), there are originally 85 type-P straight line segments and 68 type-N straight line segments. The basic requirement is the existence of 3 straight line segments and one of them is of different type. Therefore, there would be at most $C_2^{85}C_1^{68} + C_1^{85}C_2^{68} = 436\,390$ sets of triple straight line segments to consist the triangles. However, there were only 23 413 triangles considered in our proposed algorithm and the triangles are shown in Fig. 14(a)–(d) for the four types of the triangles respectively. Then, the 4011 triangle pairs are produced according to Section 3.5 as shown in Fig. 14(e) and (f) for the {01NPP, 03NPN} triangle pairs and the {02PNN, 04PPN} triangle pairs respectively. Next, the triangle pairs that have a shared edge are extracted in Fig. 14(g) and (h). Finally, the triangle clustering is performed to extract the final bicycle frame candidates. Note that the search region constraint as described in Section 3.3 is not applied to evaluate the amount of triangles reported here.

As shown in Fig. 14, this evaluation is quantitatively summarized in Table 3 in terms of the number of the considered triangles. As we can see, the amount of data to process is reduced dramatically.

4.2.4. The effectiveness of the proposed constraints on the bicycle model

To take the insight into the performance of the proposed constraints introduced in Section 3.6, the reduced data amount was evaluated on the running example as shown in Fig. 4. The number of triangles left for the next constraint is shown separately in Table 2. Again, the results show that the amount of triangles is gradually reduced. There were still a little bit many triangles after applying constraint Γ_8 . This is because reported triangles had not been clustered here.

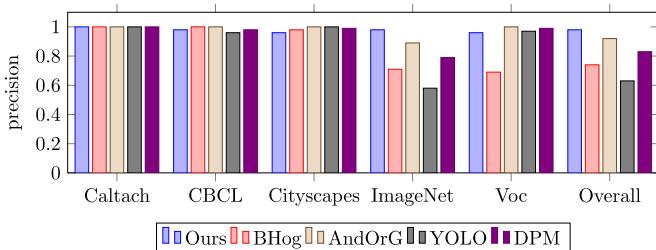
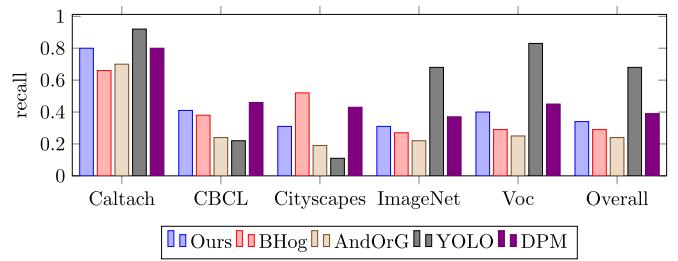
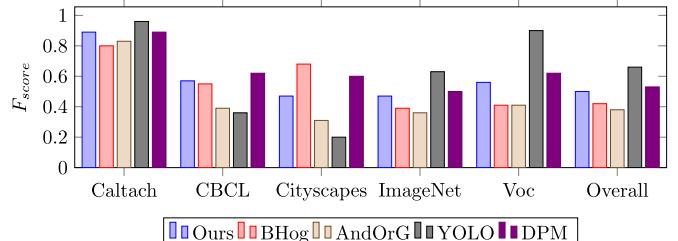
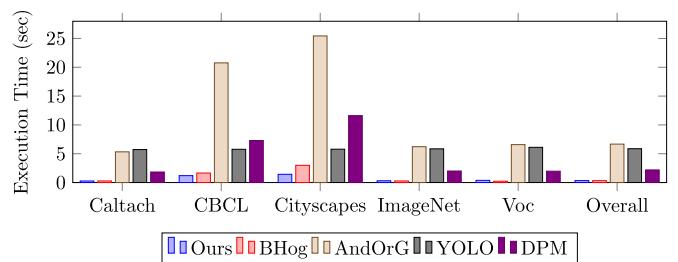
**Fig. 14.** The evaluation of the bicycle frame detection.**Table 3**
The quantitative evaluation of the bicycle frame detection.

Constraint	The extraction constraints on the bicycle frame			
	None	Triangle classification	Shared an edge	Clustering
Triangle Amount	436 390	23 413	4011	208
%	100	5.37	0.92	0.05

4.2.5. The overall performance and the comparison with the state-of-art works

Figs. 15, 16 and 17 illustrated the performance of our proposed bicycle detection algorithm in terms of precision, recall and F_{score} , respectively, over the Caltech256, CBCL Street Scenes, cityscapes datasets, the selected 12 subsets in ImageNet and VOC2012. The measured precision and recall are 0.98 and 0.34 from the proposed bicycle detection algorithm in this article over the five datasets and the F_{score} is 0.50 derived by the definition in (13).

As shown in Figs. 15–17, the overall precision of the detection results from the proposed method outperformed that from the comparing methods on the Caltech, CBCL Street Scenes and cityscapes datasets. The overall recall of the detection results from the proposed method was comparable to that from the comparing methods except to YOLO, and so was the overall F_{score} . However, the execution time of YOLO was considerably longer than that of our proposed method as

**Fig. 15.** The performance comparison over the chosen datasets in terms precision.**Fig. 16.** The performance comparison over the chosen datasets in terms recall.**Fig. 17.** The performance comparison over the chosen datasets in terms F_{score} .**Fig. 18.** The performance comparison over the chosen datasets in terms execution time.

shown in Fig. 18. BHog is the fastest method but its detection accuracy is a little lower than ours. However, the speed of our detector is comparable to that of BHog. To compare the detection accuracy and the execution time simultaneously, the $FTRatio$ metric as described at

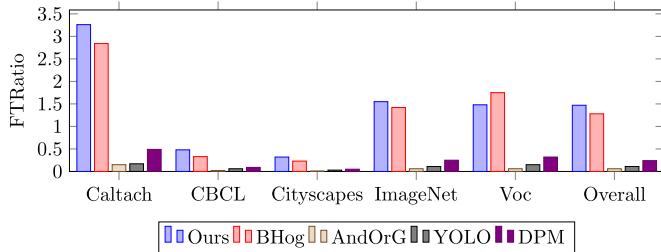


Fig. 19. The performance comparison over the chosen datasets in terms the ratio of F_{score} to execution time.

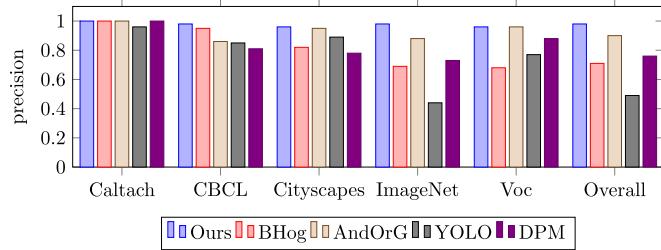


Fig. 20. The performance comparison over the chosen datasets excluding the front and rear view bicycle image in terms of precision.

the beginning of Section 4 was evaluated and illustrated in Fig. 19. As a result, the performance of our proposed bicycle detector was shown to be the best one in terms of $FTRatio$.

Furthermore, the precision, recall and execution time were evaluated once again on the updated datasets which eliminated the front view and rear view bicycle images (Figs. 20–23). This rearrangement is according to the prerequisite of the proposed bicycle detector in this paper. The measured precision and recall are 0.98 and 0.43 from the proposed bicycle detection algorithm in this article over the five datasets. The proposed method consumed 0.33 s per image in average.

The F_{score} is 0.60 and the $FTRatio$ is 1.83. These quantitative results indicated that the performance of our proposed bicycle detector was effective and efficient.

Except to the poor edge map due to the illumination condition and the extremely low resolution, there were still several reasons of the false detection in the images including false positives and false negatives. The improvement of the edge detection is out of the scope of this paper. The details of the false detections will be discussed in Section 4.4. The experimental results show that the proposed bicycle model and the detection algorithm are robust so long as the frame is composed of two triangle-like shapes even if the style of bicycle dramatically changes (Fig. 24).

4.3. The case study: the robustness to the bicycle pose

To evaluate the effectiveness and the robustness, the experimental result on the bicycle images in the different orientations was given in this section. We captured the photos for the different bicycle poses and

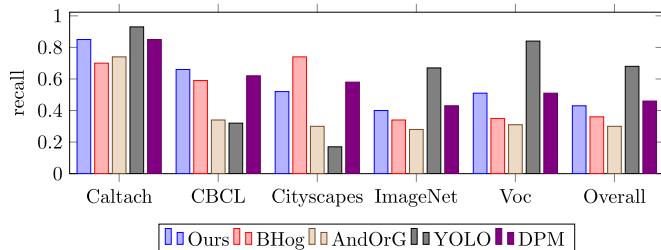


Fig. 21. The performance comparison over the chosen datasets excluding front and rear view bicycle image in terms of recall.

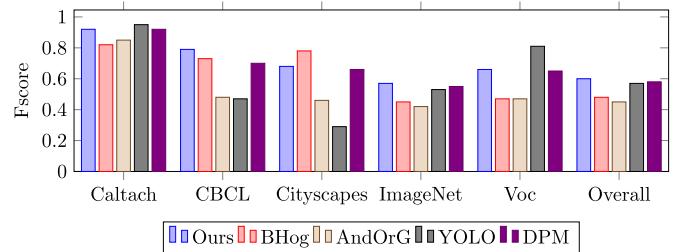


Fig. 22. The performance comparison over the chosen datasets excluding the front and rear view bicycle image in terms of F_{score} .

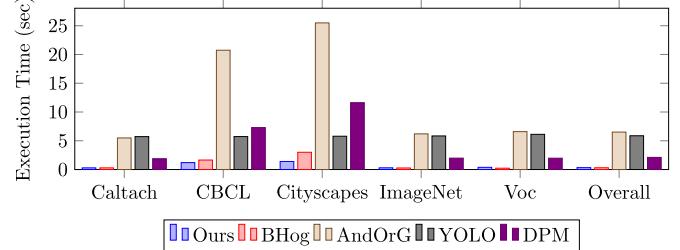


Fig. 23. The performance comparison over the datasets excluding the front and rear view bicycle image in terms of execution time.

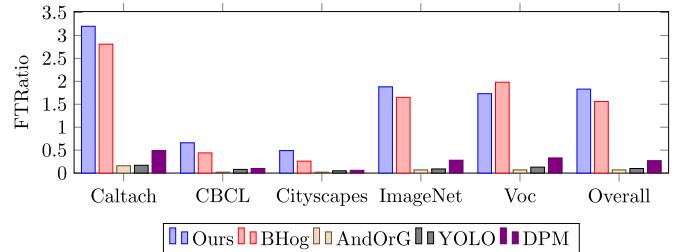


Fig. 24. The performance comparison over the chosen datasets excluding the front and rear view bicycle image in terms of the ratio of F_{score} to execution time .

applied our proposed algorithm to detect the bicycle in the image. These bicycle poses are different in the orientation with the step of 15° around to the three axes respectively. The degree of the orientation is from +75° to -75° around the y -axis, from +60° to -60° around the z -axis and from +45° to -45° around the x -axis. The range of the orientation angles around the x -axis to evaluate is reasonably narrower because of the limitation from the pedal touching the ground in this case. As shown in Fig. 25, the algorithm performs well to extract the bicycles in different orientations successfully except to the bicycles in the $\pm 75^\circ$ orientation around the y -axis.

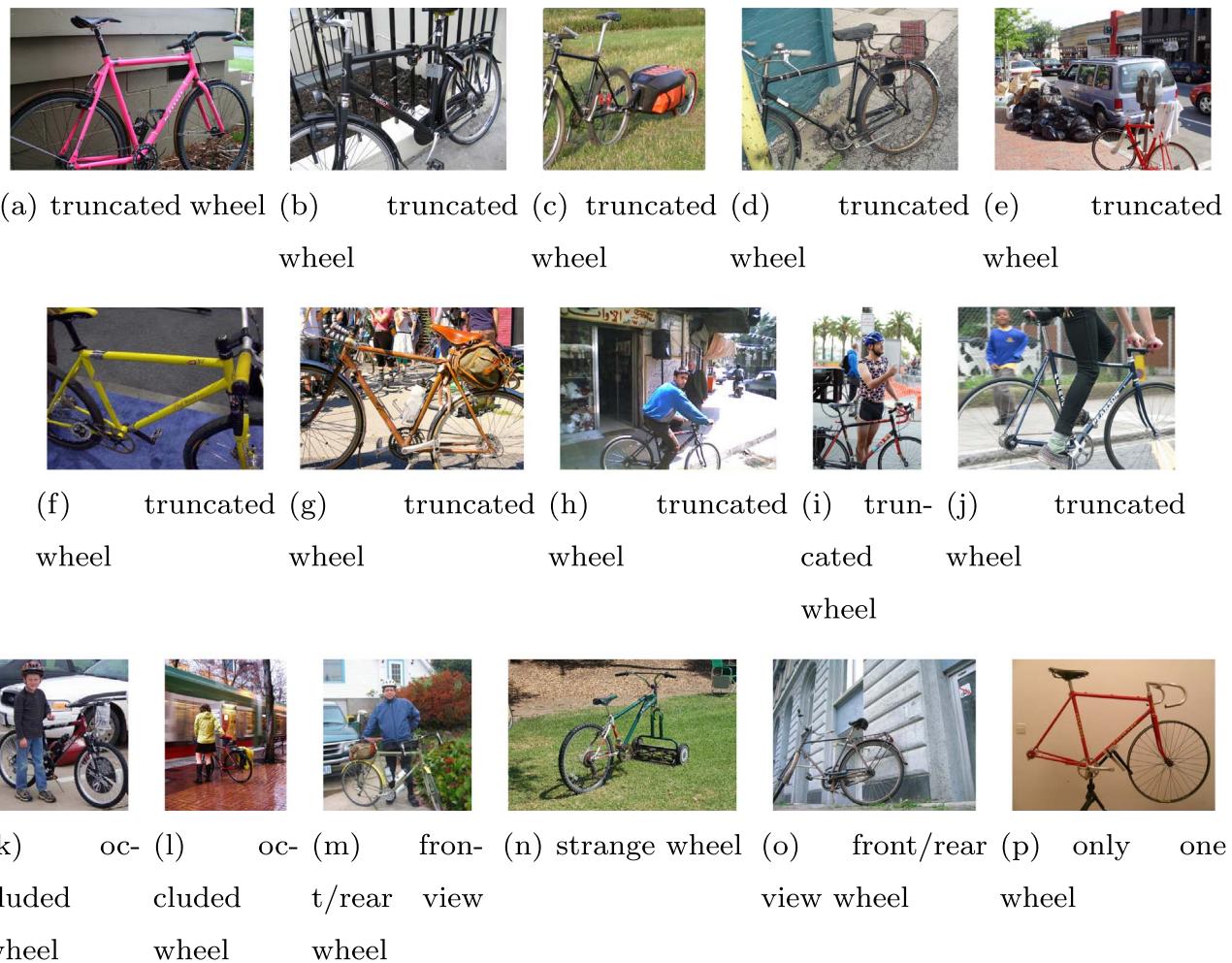
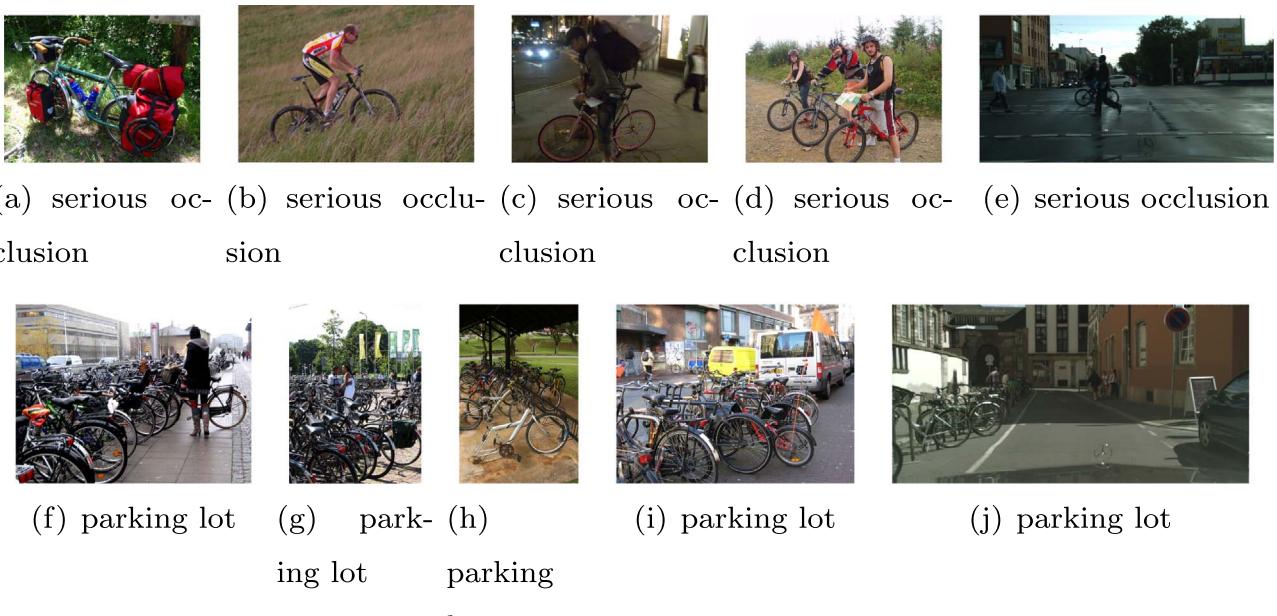
4.4. Discussion and limitation

The false detections will be discussed in the following two sections. The reasons to cause the false negatives and the false positives were explored and discussed. The reasons for the false negatives were the false negative of ellipse detection on any ellipse in a pair of wheels to compose a bicycle, the serious occlusion, the special style of bicycle and the unusual orientation of the bicycle.

4.4.1. The false negatives

Several situations caused the proposed bicycle detector to miss the target. The first one is the false negative of any ellipse in a pair of wheels to compose a bicycle when the wheel was truncated, occluded or in front/rear view as shown in Fig. 26. These situations made the demand of three arcs to detect the two ellipses presenting the pair of wheels not met such that the proposed geometric bicycle model was not held. Fig. 26(p) is an extreme example which exists the bicycle with

**Fig. 25.** The evaluation of the bicycle poses.

**Fig. 26.** The false negatives due to the false negative of any ellipse in a pair of wheels.**Fig. 27.** The false negatives due to the serious occlusion.

only one wheel in the image. Fig. 26(n) is a special case that the front wheel is strange.

The second situation is the serious occlusion. As aforementioned,

the proposed bicycle detector is robust to the occlusion when the occlusion was slight as shown in the visual detection results. The slight occlusion means that the coverage of the occlusion on the wheel is less

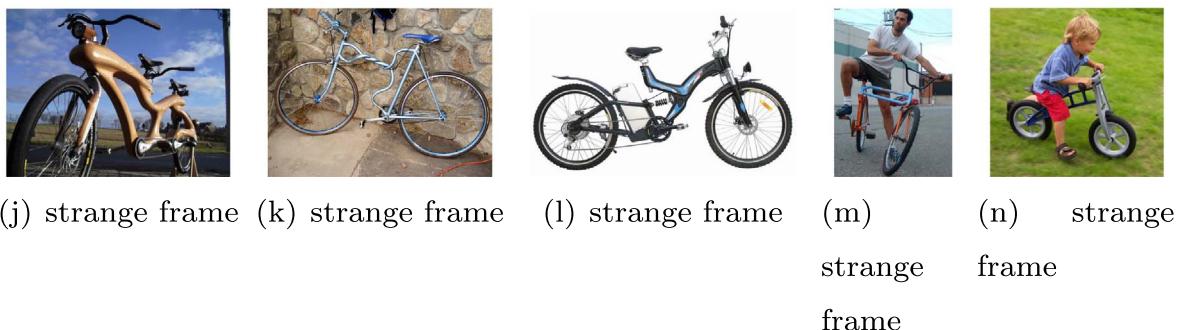


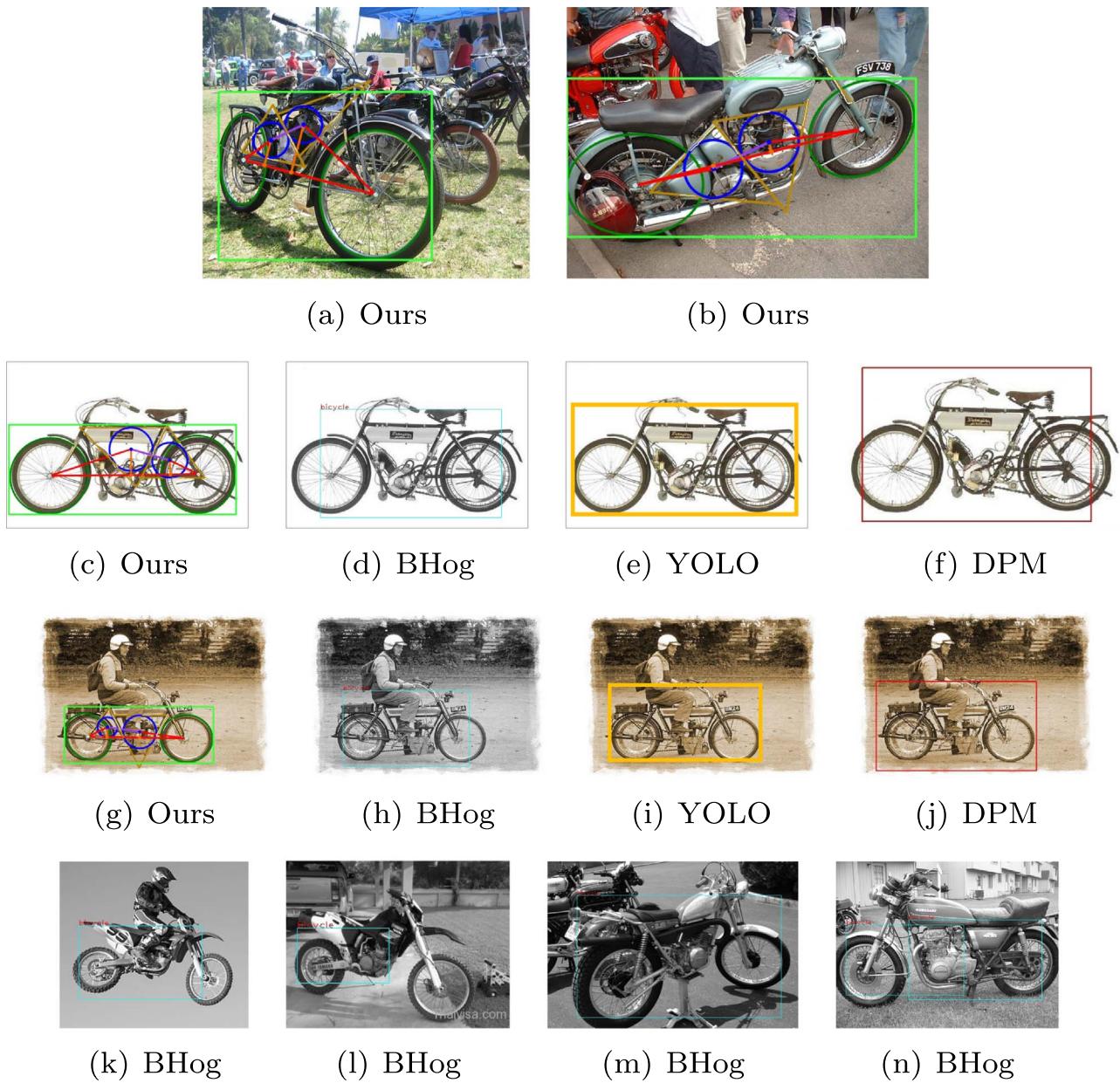
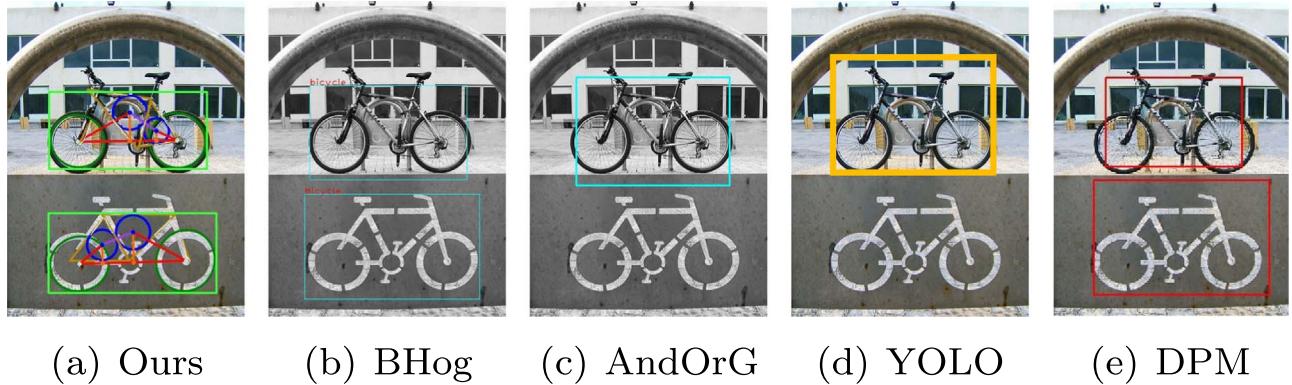
Fig. 28. The false negatives due to their special bicycle frames.



Fig. 29. The false negatives due to the unusual orientation.

than one-quarter of the ellipse contour and the main contours of the bicycle frame are visible enough. On the other hand, the serious occlusion means the occlusion blocked the tube detection or the wheel

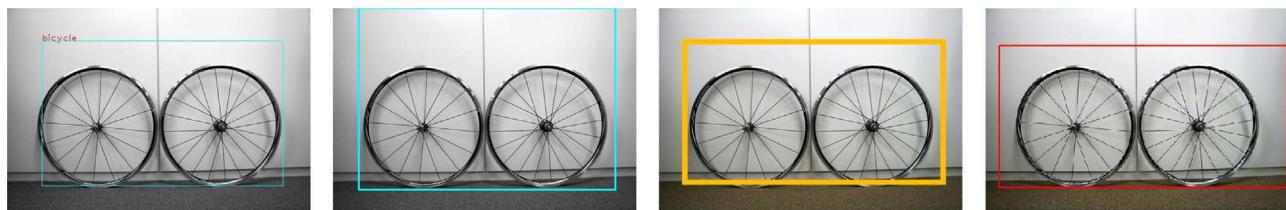
detection as shown in the top row of Fig. 27. The most serious occlusion was observed in the parking lots as shown in Fig. 27(f)–(j) which is the most challenging case for all bicycle detectors.

**Fig. 30.** The false positives on the motorcycle images due to the similar structure.**Fig. 31.** The false positives due to the printed pattern.

These two above situations derived that the main reason of the low recall of our proposed method is the truncation condition. This limitation could be improved with a more robust ellipse detection

method to ease the effect of the heavy loss of contour to form the ellipse. This is the future work we planned to develop.

The third situation is the special style of bicycle especially with



(a) BHog

(b) AndOrG

(c) YOLO

(d) DPM

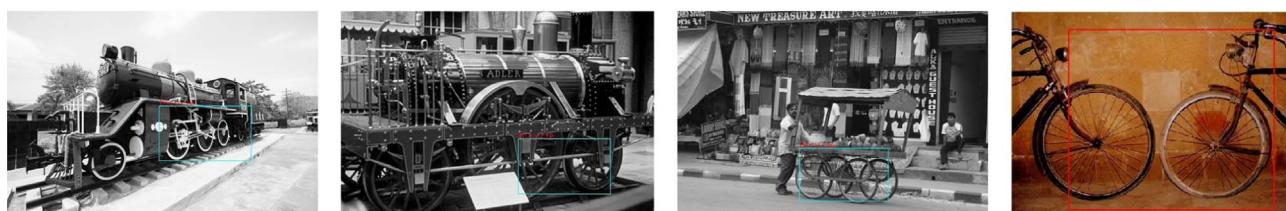
Fig. 32. The false positives on the two wheel only image.

(a) input

(b) input

(c) input

(d) YOLO

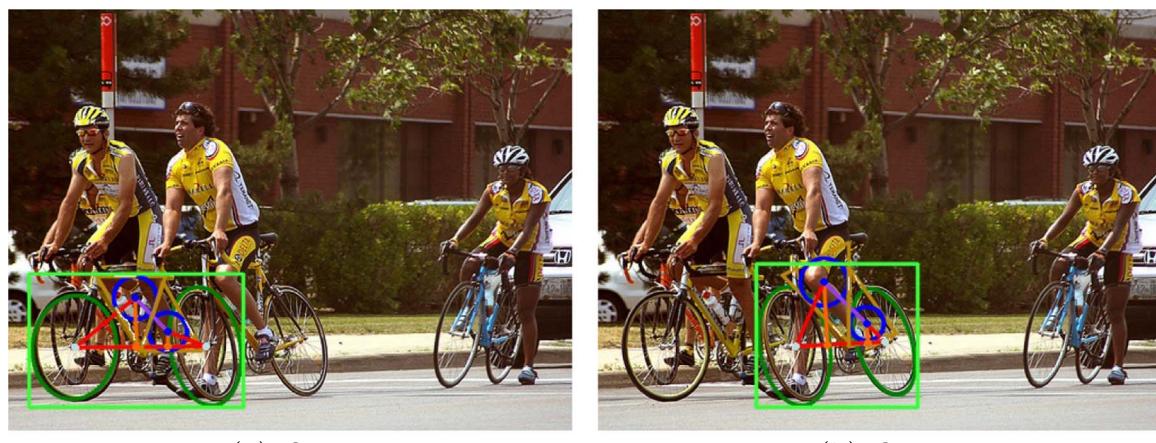


(e) BHog

(f) BHog

(g) BHog

(h) DPM

Fig. 33. The false positives on other images.

(a) Ours

(b) Ours



(c) BHog

(d) YOLO

(e) DPM

Fig. 34. The false positives on the bicycle image having the wheel occluded by a wheel.



Fig. 35. The false negatives due to the motion blur.

respect to the bicycle frame as shown in Fig. 28. Although the mainly popular bicycles are touring bicycle, mountain bicycle and Dutch bicycle, there are still folding bicycle, child bicycle and others. Fig. 28(a)–(c) shows the examples of folding bicycle. The folding bicycle in Fig. 28(a) lacks the rear triangle, the one in Fig. 28(b) lacks the front triangle intersecting the rear ellipse and the other one in Fig. 28(c) even lacks both of the triangle shapes of the bicycle frame. Similar situation occurred as shown in Fig. 28(e)–(i). Furthermore, there were bicycle images as shown in Fig. 28(j)–(n) with strange bicycle frames which are difficult to be formulated mathematically because their shapes are irregular. Note that although the missed child bicycle is shown as an example of the special bicycle frames in Fig. 28(n), the frame of a child bicycle is often similar to those of the popular bicycles with the attached pair of auxiliary wheels and could be detected successfully when the occlusion was slight as shown in Fig. 9(y). However, for a child bicycle, the occlusion was usually more serious because of the originally small size such that the rider occluded the bicycle frame more. The fourth situation is the unusual orientation as shown in Fig. 29.

4.4.2. The false positives

This section introduces the analysis of the quality of detection results with respect to the precision and recall. First, the samples of the false positives from our proposed detector on the images containing the motorcycles were found as shown in Fig. 30(a)–(c) and (g) because the geometric structure satisfied the proposed geometric bicycle model. The four false positives happened because the mechanism of this kind of motorcycle was based on the bicycle frame except to the space to attach the motor and the controller. The latter two false positives were also found in the detection results from BHog, YOLO and DPM methods. In the fourth row of Fig. 30, more false positives on the motorcycle images were obtained from BHog method, but these did not occur with our proposed method. This makes BHog method suffers lower precision. These four images present the main styles of the present motorcycles.

Besides, the printed pattern to express the logical concept was detected as a bicycle from our result as shown in Fig. 31(a), and so did the BHog and DPM methods as shown in Fig. 31(b) and (e). This is because the visual shape satisfied our proposed model and the appearance features matched the pre-trained classifier in BHog and DPM.

However, it is interesting that the image containing only two wheels led all the comparing methods to produce the false positive as shown in Fig. 32(a)–(d). Oppositely, our method resisted this kind of false positives due to the leakage bicycle frame for the proposed geometric bicycle model. Similar condition held in other results comparing to the methods. This is because there is almost no geometric shape like a bicycle in nature and the geometric shape of man-made objects similar to a bicycle did not meet the proposed bicycle model. Fig. 33 demonstrates some results in this case. Fig. 33(a)–(c) were detected to exist bicycles by BHog and DPM methods, and they did not happen in our detection results.

In the case of occlusion, Fig. 34 shows the false positive from our proposed detector on the bicycle image having the wheel occluded by a

wheel. The rear wheel of the left bicycle was occluded by the leg of the rider and the front wheel of the adjacent bicycle such that our detector treated the front wheel of the adjacent bicycle as the rear wheel of the left bicycle. In fact, this was a false positive. The BHog and YOLO methods performed good on this image, but there was an extra wrong bounding box in the detection results from YOLO method. However, there is no more information to make sure which wheel was adopted to support the detection result. This observation figured out the challenge of the wheel pairing in the images which have multiple bicycles occluding each other as shown, for example, in Fig. 27(f)–(j).

At the end, the motion blur also affects the performance of the detector. In Fig. 35, the images were blurred because of the motion happened when the image acquisition performed. However, this could be overcome by the de-blurring techniques [9,37,57,67] and the details are out of the scope of this paper.

5. Conclusion

A bicycle detection algorithm and a geometric model of bicycle for single side-view image are proposed and introduced in this paper. The experimental results showed the practicability and the performance of the proposed method. Without the demand of training process, the accuracy rate of this proposed algorithm is still acceptable and comparable. The detected bicycles with the reliability values were extracted in this proposed algorithm. In the future, we will evaluate more clustering methods for them and choose one representative bicycle of the same clustering of bicycles to make this algorithm account for the occlusion problem of multiple bicycles. Moreover, we will use the other edge detection algorithm, an improved ellipse detection algorithm and certain preprocessing technique to improve the detection accuracy.

References

- [1] J.-K. Kim, S. Kim, G.F. Ulfarsson, L.A. Porrello, Bicyclist injury severities in bicyclermotor vehicle accidents, *Accid. Anal. Prev.* 39 (2) (2007) 238–251. <http://dx.doi.org/10.1016/j.aap.2006.07.002> (URL <<http://www.sciencedirect.com/science/article/pii/S000145750600128X>>).
- [2] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005, vol. 1, 2005, pp. 886–893. <http://dx.doi.org/10.1109/CVPR.2005.177>
- [3] P. Felzenszwalb, R. Girshick, D. McAllester, D. Ramanan, Object detection with discriminatively trained part-based models, *IEEE Trans. Pattern Anal. Mach. Intell.* 32 (9) (2010) 1627–1645. <http://dx.doi.org/10.1109/TPAMI.2009.167>.
- [4] R.B. Girshick, From rigid templates to grammars: object detection with structured models, Citeseer, 2012.
- [5] S. Rogers, N. Papanikopoulos, Counting bicycles using computer vision, in: 2000 IEEE Intelligent Transportation Systems, 2000. Proceedings, 2000, pp. 33–38. <http://dx.doi.org/10.1109/ITSC.2000.881014>
- [6] J. Zhang, Z.-Q. Zhu, P. Liu, W.-H. Zhang, Detection of contours of wheels based on improved codebook, in: 2013 International Conference on Machine Learning and Cybernetics (ICMLC), vol. 2, 2013, pp. 510–515. <http://dx.doi.org/10.1109/ICMLC.2013.6890348>
- [7] Y.K. Liu, B. Alik, An efficient chain code with Huffman coding, *Pattern Recognit.* 38 (4) (2005) 553–557. <http://dx.doi.org/10.1016/j.patcog.2004.08.017> (URL <<http://www.sciencedirect.com/science/article/pii/S0031320304003723>>).
- [8] Z. Qui, D. Yao, Y. Zhang, D. Ma, X. Liu, The study of the detection of pedestrian and bicycle using image processing, in: 2003 IEEE Intelligent Transportation Systems, 2003. Proceedings, vol. 1, 2003, pp. 340–345. vol.1. <http://dx.doi.org/10.1109/ITSC.2003.1251974>

- [9] Y. Fujimoto, J. Hayashi, A method for bicycle detection using ellipse approximation, in: 2013 Proceedings of the 19th Korea-Japan Joint Workshop on Frontiers of Computer Vision, (FCV), 2013, pp. 254–257. <http://dx.doi.org/10.1109/FCV.2013.6485499>
- [10] K. Takahashi, Y. Kuriya, T. Morie, Bicycle detection using pedaling movement by spatiotemporal Gabor filtering, in: TENCON 2010–2010 IEEE Region 10 Conference, 2010, pp. 918–922. <http://dx.doi.org/10.1109/TENCON.2010.5686551>
- [11] Y. Lee, T. Kim, S. Lee, J. Shim, Spatial regions periodicity based detection of two-wheeler using histogram of oriented gradients, *Int. J. Multimed. Ubiquitous Eng.* 10 (4) (2015) 325–336.
- [12] H. Jung, Y. Ehara, J.K. Tan, H. Kim, S. Ishikawa, Detection of a bicycle in video images using msc-hog feature, *Int. J. Innov. Comput. Inf. Control* 10 (2) (2014) 521–533.
- [13] R.E. Schapire, Y. Singer, Improved boosting algorithms using confidence-rated predictions, *Mach. Learn.* 37 (3) (1999) 297–336. <http://dx.doi.org/10.1023/A:10071614523901>.
- [14] M.K. Kocamaz, J. Gong, B.R. Pires, Vision-based counting of pedestrians and cyclists, in: 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), 2016, pp. 1–8. <http://dx.doi.org/10.1109/WACV.2016.7477685>
- [15] B.K. Horn, B.G. Schunck, Determining optical flow, *Artif. Intell.* 17 (1) (1981) 185–203. [http://dx.doi.org/10.1016/0004-3702\(81\)90024-2](http://dx.doi.org/10.1016/0004-3702(81)90024-2).
- [16] H. Bay, A. Ess, T. Tuytelaars, L.V. Gool, speeded-up robust features (surf), *Comput. Vis. Image Underst.* 110 (3) (2008) 346–359. <http://dx.doi.org/10.1016/j.cviu.2007.09.014> (similarity Matching in Computer Vision and Multimedia. URL <http://www.sciencedirect.com/science/article/pii/S1077314207001555>).
- [17] G. Somasundaram, V. Morellas, N. Papaniolopoulos, Counting pedestrians and bicycles in traffic scenes, in: 2009 Proceedings of the 12th International IEEE Conference on Intelligent Transportation Systems, 2009, pp. 1–6. <http://dx.doi.org/10.1109/ITSC.2009.5309690>
- [18] A. Bosch, A. Zisserman, X. Munoz, Representing shape with a spatial pyramid kernel, in: Proceedings of the 6th ACM International Conference on Image and Video Retrieval, CIVR '07, ACM, New York, NY, USA, 2007, pp. 401–408. <http://dx.doi.org/10.1145/1282280.1282340>
- [19] D.G. Lowe, Object recognition from local scale-invariant features, in: Proceedings of the Seventh IEEE International Conference on Computer Vision, 1999, vol. 2, 1999, pp. 1150–1157. <http://dx.doi.org/10.1109/ICCV.1999.790410>
- [20] H. Hu, P. Tao, Z. Gao, Q. Wang, Z. Li, Z. Qu, Vision-based bicycle detection using multiscale block local binary pattern, *Math. Probl. Eng.* (2014).
- [21] K. Yang, C. Liu, J.Y. Zheng, L. Christopher, Y. Chen, Bicyclist detection in large scale naturalistic driving video, in: Proceedings of the 17th International IEEE Conference on Intelligent Transportation Systems (ITSC), 2014, pp. 1638–1643. <http://dx.doi.org/10.1109/ITSC.2014.6957928>
- [22] F.F. Shahraki, A.P. Yazdanpanah, E.E. Regentova, V. Muthukumar, Bicycle detection using HOG, HSC and MLBP in: Advances in Visual Computing: Proceedings of the 11th International Symposium, ISVC 2015, Las Vegas, NV, USA, December 14–16, 2015, Proceedings, Part II, Springer International Publishing, Cham, 2015, pp. 554–562. http://dx.doi.org/10.1007/978-3-319-27863-6_51
- [23] W. Tian, M. Lauer, Fast cyclist detection by cascaded detector and geometric constraint, in: 2015 IEEE Proceedings of the 18th International Conference on Intelligent Transportation Systems, 2015, pp. 1286–1291. <http://dx.doi.org/10.1109/ITSC.2015.211>
- [24] W. Tian, M. Lauer, Fast and robust cyclist detection for monocular camera systems. <http://dx.doi.org/10.1.1.721.7825>
- [25] T. Watanabe, S. Ito, K. Yokoi, Co-occurrence Histograms of Oriented Gradients for Pedestrian Detection, Springer, Berlin, Heidelberg, 2009, pp. 37–47. http://dx.doi.org/10.1007/978-3-540-92957-4_4
- [26] G. Yanlei, S. Kamijo, Bicyclist recognition and orientation estimation from on-board vision system, *Int. J. Autom. Eng.* 6 (2) (2015) 67–73. http://dx.doi.org/10.20485/jxaeiae.6.2_67
- [27] I. Laptev, Improving object detection with boosted histograms, *Image Vis. Comput.* 27 (5) (2009) 535–544. <http://dx.doi.org/10.1016/j.imavis.2008.08.010>.
- [28] X. Wang, L. Lin, L. Huang, S. Yan, Incorporating structural alternatives and sharing into hierarchy for multiclass object recognition and detection, in: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013.
- [29] H. Cho, P. Rybski, W. Zhang, Vision-based bicycle detection and tracking using a deformable part model and an ekf algorithm, in: 2010 Proceedings of the 13th International IEEE Conference on Intelligent Transportation Systems (ITSC), 2010, pp. 1875–1880. <http://dx.doi.org/10.1109/ITSC.2010.5624993>
- [30] H. Cho, P.E. Rybski, W. Zhang, Vision-based bicyclist detection and tracking for intelligent vehicles, in: 2010 IEEE Intelligent Vehicles Symposium (IV), 2010, pp. 454–461. <http://dx.doi.org/10.1109/IVS.2010.5548063>
- [31] H.H. Chen, C.-C. Lin, W.-Y. Wu, Y.M. Chan, L.C. Fu, P.Y. Hsiao, Integrating appearance and edge features for on-road bicycle and motorcycle detection in the nighttime, in: Proceedings of the 17th International IEEE Conference on Intelligent Transportation Systems (ITSC), 2014, pp. 354–359. <http://dx.doi.org/10.1109/ITSC.2014.6957716>
- [32] B.-F. Lin, Y.-M. Chan, L.-C. Fu, P.-Y. Hsiao, L.-A. Chuang, S.-S. Huang, M.-F. Lo, Integrating appearance and edge features for sedan vehicle detection in the blind-spot area, *IEEE Trans. Intell. Transp. Syst.* 13 (2) (2012) 737–747. <http://dx.doi.org/10.1109/TITS.2011.2182649>.
- [33] J. Yan, Q. Ling, Y. Zhang, F. Li, F. Zhao, An adaptive bicycle detection algorithm based on multi-Gaussian models, *J. Comput. Inf. Syst.* 9 (2013) 10075–10083.
- [34] S. Messelodi, C.M. Modena, G. Cattoni, Vision-based bicycle/motorcycle classification, *Pattern Recognit. Lett.* 28 (13) (2007) 1719–1726. <http://dx.doi.org/10.1016/j.patrec.2007.04.014>.
- [35] H. Ardö, M. Nilsson, A. Laureshy, A. Persson, Reduced Search Space for Rapid Bicycle Detection, 2nd International Conference on Pattern Recognition Applications and Methods (ICPRAM 2013), SciTePress, 2013, p. 6 <http://dx.doi.org/10.5220/0004264804530458>.
- [36] A. Moro, E. Mumolo, M. Nolich, K. Umeda, Real-time gpu implementation of an improved cars, pedestrians and bicycles detection and classification system, in: 2011 Proceedings of the 14th International IEEE Conference on Intelligent Transportation Systems (ITSC), 2011, pp. 1343–1348. <http://dx.doi.org/10.1109/ITSC.2011.6083004>
- [37] J. Dukesherer, C. Smith, A hybrid Hough–Hausdorff method for recognizing bicycles in natural scenes, in: 2001 IEEE International Conference on Systems, Man, and Cybernetics, vol. 4, 2001, pp. 2493–2498. <http://dx.doi.org/10.1109/ICSMC.2001.972932>
- [38] M. Leung, T. Huang, Detecting wheels of vehicle in stereo images, in: Proceedings of the 10th International Conference on Pattern Recognition, 1990, vol. 1, 1990, pp. 263–267. <http://dx.doi.org/10.1109/ICPR.1990.118108>
- [39] T. Ardeshtiri, F. Larsson, F. Gustafsson, T. Schon, M. Felsberg, Bicycle tracking using ellipse extraction, in: 2011 Proceedings of the 14th International Conference on Information Fusion (FUSION), 2011, pp. 1–8.
- [40] B. Sun, K. Saenko, From virtual to reality: fast adaptation of virtual object detectors to real domains., in: BMVC, vol. 1, 2014, p. 3.
- [41] P.L. Rosin, T. Ellis, Frame-based system for image interpretation, *Image Vis. Comput.* 9 (6) (1991) 353–361. [http://dx.doi.org/10.1016/0262-8856\(91\)90002-7](http://dx.doi.org/10.1016/0262-8856(91)90002-7).
- [42] G. Loy, N. Barnes, Fast shape-based road sign detection for a driver assistance system, in: 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2004 (IROS 2004). Proceedings, vol. 1, 2004, pp. 70–75. <http://dx.doi.org/10.1109/IROS.2004.1389331>
- [43] J. Khan, S. Bhuiyan, R. Adhami, Image segmentation and shape analysis for road-sign detection, *IEEE Trans. Intell. Transp. Syst.* 12 (1) (2011) 83–96. <http://dx.doi.org/10.1109/TITS.2010.2073466>
- [44] G. Loy, A. Zelinsky, Fast radial symmetry for detecting points of interest, *IEEE Trans. Pattern Anal. Mach. Intell.* 25 (8) (2003) 959–973. <http://dx.doi.org/10.1109/TPAMI.2003.1217601>.
- [45] O. Achler, M. Trivedi, Vehicle wheel detector using 2d filter banks, in: 2004 IEEE Intelligent Vehicles Symposium, 2004, pp. 25–30. <http://dx.doi.org/10.1109/IVS.2004.1336350>
- [46] V. Beran, A. Herout, I. Řezníček, Video-based bicycle detection in underground scenarios, in: Proceedings of the 17th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision, 2009, pp. 103–107.
- [47] C. Basca, M. Talos, R. Brad, Randomized Hough transform for ellipse detection with result clustering, in: The International Conference on Computer as a Tool, 2005. EUROCON 2005, vol. 2, 2005, pp. 1397–1400. <http://dx.doi.org/10.1109/EURCON.2005.1630222>
- [48] R.A. McLaughlin, Randomized Hough transform: improved ellipse detection with comparison, *Pattern Recognit. Lett.* 19 (34) (1998) 299–305. [http://dx.doi.org/10.1016/S0167-8655\(98\)00010-5](http://dx.doi.org/10.1016/S0167-8655(98)00010-5)
- [49] S.-C. Zhang, Z.-Q. Liu, A robust, real-time ellipse detector, *Pattern Recognit.* 38 (2) (2005) 273–287. <http://dx.doi.org/10.1016/j.patcog.2004.03.014> (URL <http://www.sciencedirect.com/science/article/pii/S0031320304001372>).
- [50] L. Libuda, I. Grothues, K.-F. Kraiss, Ellipse detection in digital image data using geometric features, in: Advances in Computer Graphics and Computer Vision: International Conferences VISAPP and GRAPP 2006, Setúbal, Portugal, February 25–28, 2006, Revised Selected Papers, Springer, Berlin, Heidelberg, 2007, pp. 229–239. http://dx.doi.org/10.1007/978-3-540-75274-5_15
- [51] D.K. Prasad, M.K. Leung, S.-Y. Cho, Edge curvature and convexity based ellipse detection method, *Pattern Recognit.* 45 (9) (2012) 3204–3221. <http://dx.doi.org/10.1016/j.patcog.2012.02.014> (best Papers of Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA'2011). URL <http://www.sciencedirect.com/science/article/pii/S0031320312000763>).
- [52] J.-P. He, Y. Ma, Triangle detection based on windowed Hough transform, in: International Conference on Wavelet Analysis and Pattern Recognition, 2009. ICWAPR 2009, 2009, pp. 95–100. <http://dx.doi.org/10.1109/ICWAPR.2009.5207484>
- [53] H.-M. Liu, Z.-H. Wang, Detection of arbitrary triangle, in: 2011 4th International Congress on Image and Signal Processing (CISP), vol. 2, 2011, pp. 893–897. <http://dx.doi.org/10.1109/CISP.2011.6100305>
- [54] H. Demirezen, M. Baran, Triangle detection with color information, in: 2010 International Conference on Applied Electronics (AE), 2010, pp. 1–4.
- [55] J. Zhu, J. Yu, C. Wang, F.Z. Li, Colour combination attention for object recognition, *IET Image Process.* 8 (9) (2014) 539–547. <http://dx.doi.org/10.1049/iet-ipr.2013.0431>.
- [56] J.R. Lv, X.G. Luo, An approach for triangle traffic sign detection based on improved radon transform, in: 2012 Proceedings of the Fourth International Conference on Computational and Information Sciences (ICCIS), 2012, pp. 176–178. <http://dx.doi.org/10.1109/ICCIS.2012.62>
- [57] M. Fornaciari, A. Prati, R. Cucchiara, A fast and effective ellipse detector for embedded vision applications, *Pattern Recognit.* 47 (11) (2014) 3693–3708. <http://dx.doi.org/10.1016/j.patcog.2014.05.012>
- [58] P.L. Rosin, Measuring shape: ellipticity, rectangularity, and triangularity, *Mach. Vis. Appl.* 14 (3) (2003) 172–184. <http://dx.doi.org/10.1007/s00138-002-0118-6>.
- [59] G. Bradski, A. Kaehler, Learning OpenCV: Computer Vision with the OpenCV Library, O'Reilly, 2008.
- [60] G. Griffin, A. Holub, P. Perona, Caltech-256 object category dataset. Technical Report 7694, California Institute of Technology, 2007. <http://authors.library.caltech.edu/7694>

- [61] J. Deng, W. Dong, R. Socher, L.J. Li, K. Li, L. Fei-Fei, Imagenet: A large-scale hierarchical image database, in: IEEE Conference on Computer Vision and Pattern Recognition, 2009. CVPR 2009, 2009, pp. 248–255. <http://dx.doi.org/10.1109/CVPR.2009.5206848>
- [62] G.A. Miller, Wordnet: a lexical database for English, Commun. ACM 38 (11) (1995) 39–41. <http://dx.doi.org/10.1145/219717.219748>.
- [63] M. Everingham, S.M.A. Eslami, L. Gool, C.K.I. Williams, J. Winn, A. Zisserman, The Pascal visual object classes challenge: a retrospective, Int. J. Comput. Vis. 111 (1) (2014) 98–136. <http://dx.doi.org/10.1007/s11263-014-0733-5>.
- [64] S. Bileschi, Cbcl streetscenes challenge framework, 2007.
- [65] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, B. Schiele, The cityscapes dataset for semantic urban scene understanding, 2016, CoRR abs/1604.01685. [arXiv:1604.01685](https://arxiv.org/abs/1604.01685)
- [66] J. Redmon, S.K. Divvala, R.B. Girshick, A. Farhadi, You only look once: Unified, real-time object detection, 2015, CoRR abs/1506.02640.
- [67] S. Xiang, G. Meng, Y. Wang, C. Pan, C. Zhang, Image deblurring with matrix regression and gradient evolution, Pattern Recognit. 45 (6) (2012) 2164–2179. <http://dx.doi.org/10.1016/j.patcog.2011.11.026> (brain Decoding, URL (<http://www.sciencedirect.com/science/article/pii/S0031320311004808>)).



Yen-Bor Lin received the B.S. degree in computer science engineering from Yuan Ze University, Chung-Li, Taiwan, in 2006 and the M.S. degree in computer science and information engineering from National Cheng Kung University, Tainan, Taiwan, in 2008. He is currently working toward the Ph.D. degree in the Department of Computer Science and Information Engineering. His research interests include image processing, embedded system design, biomedical signal processing, and cryptography.



Chung-Ping Young (S97M06) received the B.S. degree in electronic engineering from Chung Yuan Christian University, Zhongli, Taiwan, in 1985 and the M.S. and Ph.D. degrees in electrical engineering from the University of Missouri, Columbia, in 1994 and 1997, respectively. From 1994 to 1997, he was a Research Assistant with the University of Missouri, where he worked on power measurement. From 1998 to 2003, he was with Wistron InfoComm and Phoenix Technologies as a Senior Engineer and Principal Engineer. Since 2003, he has been with the Department of Computer Science and Information Engineering, National Cheng Kung University, Tainan, Taiwan, where he is currently an Associate Professor. His fields of interest include embedded vision, virtual instrumentation, telematics, and wireless sensor networks.