Open Access

Self-Similarity Based Zebra-Crossing Detection for Intelligent Vehicle

Chao Wang[*], Chunxia Zhao and Huan Wang

School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China

Abstract: A zebra-crossing detection method for intelligent vehicle is proposed in this paper. The method is performed on a bird-eye view image called inverse perspective mapping image. The complete method includes two phrases. First, a morphological filer followed by horizontal projection is applied to fast extract candidate zebra-crossing regions, where the size and structure information of zebra-crossing are well utilized. Second, a recognition method based on self-similarity is presented to identify the candidate regions. Given a seed region of a zebra-crossing, the recognition method can infer overall zebra-crossing region by matching and growing. Experiments on a great number of real images which consist of several challenge scenes demonstrate the effectiveness and efficiency of the proposed approach.

Keywords: Intelligent vehicle, Zebra-crossing detection, Self-similarity, Inverse perspective mapping.

1. INTRODUCTION

The zebra crossing is a typical road marker for pedestrian crossing used in many countries around the world. It consists of multiple white strips, which are parallel to each other in a 3-D space. They intersect at a fixed point, called the vanishing point, when they are projected onto a 2-D image. A zebra crossing is an important perceptive object for intelligent vehicles (IV), since it is crucial to identify crossing roads. For instance, zebra-crossing detection can help the intelligent vehicle to automatically slow down its speed and trigger a special planning strategy in a cross road, or send an early warning signal to remind driver to be more careful.

For zebra-crossing detection, much work has been done for enhancing the safety and mobility of blind people while crossing a road [1-5]. In [1], a system called "cross-watch" is introduced; it can provide information about the location and orientation of crosswalks to a blind or visually impaired pedestrian holding a camera cell phone. In [5], a zebra-crossing detection approach is also designed for partially sighted people, which can tackle the ambiguity in distinguishing zebra crossings and stair-cases. However, these methods are only suitable for low-speed and rough detection. In [6], AdaBoost is applied to detect strips of zebra-crossing, a post-processing step is designed to group all strips to obtain final results. The method needs to employ large samples to train AdaBoost classifier. Mohammad *et al.* [15] use bipolar patterns to find candidate regions, the presence of a zebra crossing is inferred by analysis of structure and direction feature, including crossing width, bandwidth trends and so on. The method is further revised in [16] by introducing projection invariance recognition to adapt to various illuminations. Unfortunately, these methods are not specifically investigated for autonomous land vehicles and vehicle driving systems.

This paper proposes a robust zebra-crossing detection algorithm for autonomous land vehicle and driving assistance system. Firstly, inverse perspective map is created by utilizing camera calibration parameters to obtain a bird-eye view road image. Secondly, a course-to-fine detection process is applied to obtain candidate zebra-crossing region and finally true positives are recognized by extracting appearance and shape features. Experiments on real road image sequences which include many challenge scenes demonstrate the effectiveness and efficiency of the proposed method.

2. THE SELF-SIMILARITY MODEL

If some sub-regions of a zebra-crossing are considered, we may observe that they are self-similar. Specifically, see Fig. (**1**), suppose that a window is laid over a zebra-crossing and the sub-region inside the window consists of a pattern of three white marker patches and two road patches. Then we shift the window randomly. We may observe that the content inside the window changes but similar sub-regions are often observed; in particular, some similar sub-regions appear periodically if we shift the window along image rows, while others appear constantly along the main edge orientation (the direction of blue arrow). Note that there are a few shape and scale changes due to the perspective effect and we will consider this effect later. Anyway, the same evidence can also be observed if a sub-region with a different pattern is used.

Based on the above observation, a self-similarity based recognition method is proposed to identify a zebra-crossing. The method comprises two stages.

The first stage is a horizontal search and matching process, which aims to identify a zebra-crossing based on the similarity curve obtained from matching. We first set a prior pattern p. Without loss of generality, we assume that p is a pattern which consists of three white marker patches and two road patches. Then p is matched with each of the regions $\{q_i\}_{i=1}^{N}$ which lie in the same row and with the same size with

*Address correspondence to this author at the Nanjing University of Science and Technology, Researches on Intelligent Robot and Vehicle, E-mail: wangchao234@163.com
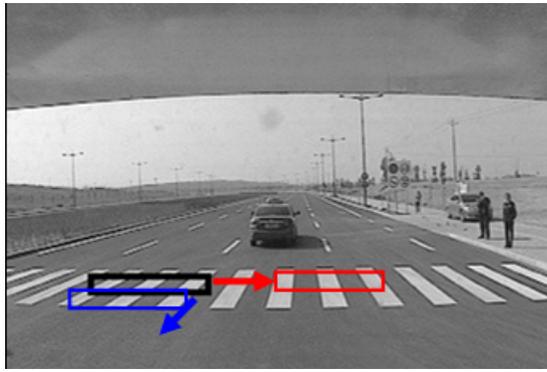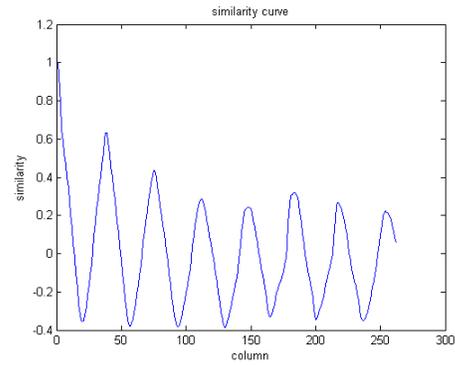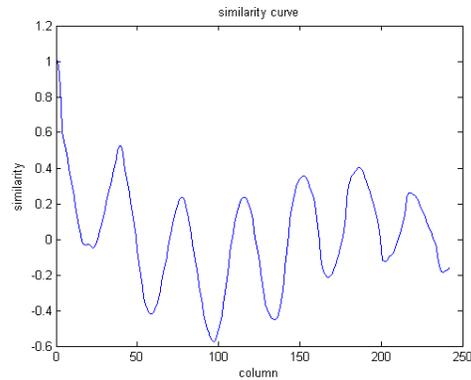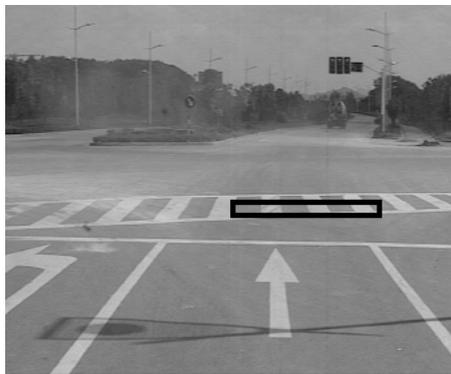
Fig. (1). A zebra-crossing comprises many self-similar sub-regions.

(a)

(b)

Fig. (2). Similarity curves under the effect of shadows (**a**) and dirty (**b**) are shown. The original images are shown on the left, where the prior patterns are displayed on respective images. Similarity curves are shown on the right, which are obtained by shifting the pattern horizontally and matching with other sub-regions with the same size.

p. All the similarity values are recorded according to the order of column coordination. Here the similarity function is defined based on the correlation coefficient:

$$sim(p,q_i)=\frac{(\overline{f}_p-u_p)\cdot(\overline{f}_{q_i}-u_{q_i})}{\left\|\overline{f}_p-u_p\right\|\cdot\left\|\overline{f}_{q_i}-u_{q_i}\right\|} \tag{1}$$

Where $\overline{f}_p$ and $\overline{f}_{q_i}$ are one-dimension intensity vectors of p and q_i individually, obtained by concatenating each

column of them; u_p and u_{q_i} denote their intensity mean. The main idea of using the correlation coefficient is to alleviate the effect of non-uniform illumination, for example, shadows, highlights.

Fig. (**2**) gives two images for illustration. One is under the effect of shadows and the other is under highlights. We select a sub-region as the prior patter p and let p match with other sub-regions along image rows. Their similarity curves are depicted individually. In Fig. (**2**), we can see
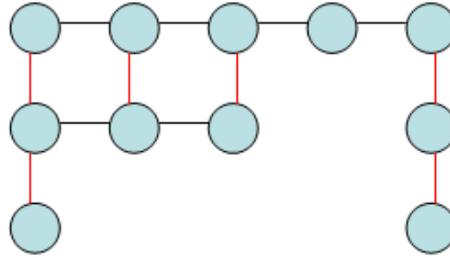
Fig. (3). A graph derived from the two stages.

clearly that the similarity curve changes periodically. Their magnitude changes but frequency keeps the same. This feature is of great use for zebra-crossing recognition since few regions show this merit in road scenarios except zebra-crossing region.

In the following, we design a specific method to identify the similarity curve. First, all the peaks in the curve are located. A Difference of Gaussian (DOG) filter convolves with the curve so that edge points are located. DOG filter is given as follows:

$$DoG(x) = \exp(-\frac{x^2}{2\sigma_1^{\ 2}}) - \exp(-\frac{x^2}{2\sigma_2^{\ 2}}) \qquad (2)$$

Where σ_1 and σ_2 are different scale factors for the two Gaussian functions, and $\sigma_1 < \sigma_2$. Then local maximums are further detected based on these edge points. Last, we calculate intervals of any two maximums and the difference of each maximum with its neighboring local minimum. It is required that the intervals should approximately be the same and the average different should be large than a threshold, say 0.5.

The second stage is to search and match along the main edge orientation θ of q_i, where θ can be obtained from the gradient histogram of q_i. For q_i and each of $\{R_j^{q_i}\}_{j=1}^n$ generated from the first scanning stage, they compare with their neighboring regions along θ and the regions whose similarities are higher than some predefined threshold T are constantly merged. The process stops until the similarity is lower than T.

Finally, a graph is obtained through the two stages, which may look like the graph shown in Fig. (**3**). In the graph, each node denotes a local similar region of the prior pattern. Although the remaining unexploited regions in this graph may be occluded by vehicles, pedestrians or dirty, the graph can be utilized to infer a zebra-crossing based on its structure and area. The formulation for identifying a zebra-crossing is given as follows:

$$s = \begin{cases} 1 & if \ \dfrac{N}{L} > T_1 \ \& \ L > T_2 \\ 0 & \\ & ohterwise \end{cases} \qquad (3)$$

Where N is the number of all nodes in the graph, and L is the maximum number of nodes in a row in the graph.

3. THE PROPOSED APPROACH

Our approach chiefly consists of three stages: an incoming road image is first transformed to an inverse perspective map (IPM), multiple candidate zebra-crossing regions are obtained in IPM based on morphologic filter and intensity projection. Finally, candidate zebra-crossing regions are further identified by the self-similarity model. Since the self-similarity model has been described in last section, we mainly introduce how to initialize and use this model properly.

### 3.1. Inverse Perspective Mapping

We first eliminate the perspective effect incurring in the captured road images to get the bird's eye view of the road [6-7], namely, inverse perspective mapping (IPM), in this process, all incoming images are remapped to a new image according to camera calibration parameters. This operation can make the following processing easier than the processing in the original road image. Let $M = (x, y.z)^T$ denote by a point in the surface of road, and $m = (u,v)^T$ denote by the corresponding point in image coordinate. Let $\tilde{M} = (x, y.z,1)^T$ and $\tilde{m} = (u,v,1)^T$ denote by the homogeneous coordinate of M and m. Here we assume that the road is flat and coincides with the plane $Z = 0$, hence $\tilde{M}$ can be simplified as $\tilde{M} = (x,y,1)^T$. Then we can obtain a 3×3 matrix H which can describe the relation of $\tilde{M}$ and $\tilde{m}$ is described as: $s \cdot \tilde{m} = H \cdot \tilde{M}$.

Where s is a non-zero scale factor. $H = A[r_1 \ r_2 \ t]$, which is a 3×3 matrix, the calculation of matrix H depends on the internal calibration parameter A and external calibration parameters r_1, r_2, t. The first two are rotation vector and the latter is translation vector. As to the concrete formulation, one can refer to [8, 9].
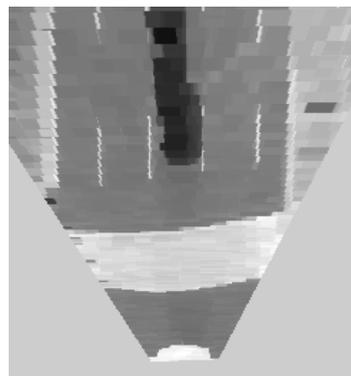
Based on the matrix H, an IPM image can be obtained. Specifically, Fig. (**4**) shows an IPM image, in this image, a pixel represents 5cm×5cm (here 'cm' represents centimeter and 'm' represents meter) region in world coordination.

(a)
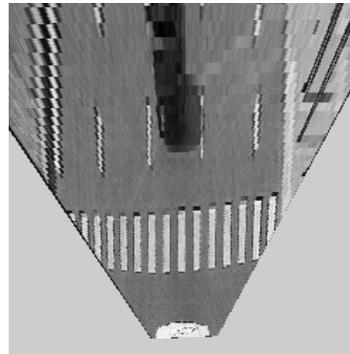
(b)

(c)

Fig. (4). Horizontal morphological close operation. (**a**) Captured road image (**b**) An IPM image, (**c**) Result image of horizontal morphological close operation with (**b**).

We set both the width and the height of the IPM image to 500 pixels, which means that the detection range in 3-D real world is that -12.5m to 12.5m in horizontal direction and 0 to 25m in vertical direction. Note that there are some invalid regions (gray regions) in the IMP image, which means the pixels in these regions do not find corresponding points in the captured image (Fig. **4a**). In the following, we use a mask image to indicate these regions and do nothing in these regions. It should be noted that the front part of the vehicle is also neglected by the mask technique. We can change the resolution of each pixel in the IPM image, for example, a pixel can represent 10cm×10cm or 1cm×1cm area in 3D space. However, it is shown that the 5cm×5cm obtains a good balance between the efficiency and the accuracy in our experiments (Fig. **4b**).

An advantage of using IPM is that strips of zebra-crossing are parallel in IPM image which is a robust characteristic to detect zebra-crossing. Another advantage is that some prior information about zebra-crossing's size can be utilized to recognize zebra-crossing. Usually, the size of zebra-crossing is standard, in China, the width of white strip is 45cm, the length is 500cm or 600cm, and the gap between two white neighborhood strips is 60cm. The prior information can be well used to remove most false positives.

3.2. Candidate Zebra-Crossing Detection in IPM

A zebra crossing in the IPM image is like a white rectangle region because it has dense vertical strips. Therefore, a straightforward strategy to find white regions in IMP image
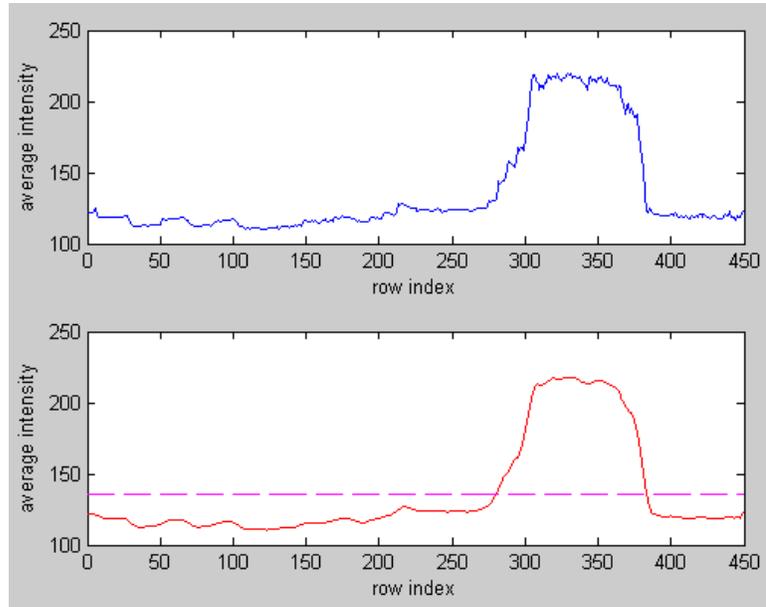
Fig. (5). Projection curve of horizontal intensity. The bottom red curve is the smoothed curve of the top one.

is used to find candidate zebra-crossing. This strategy includes two phrases, the Morphological close operation for enhancing the zebra crossing region and the Horizontal projection for localize white regions.

Morphological close operation is applied to the IPM to make the strips of zebra-crossing joined together, so as to enhance zebra crossing. Specifically, this can enhance the intensity of the zebra region with respect to other regions. Eq. 4 shows the formation of Morphological close operation, f is IPM image, b is 1×12 line-type structure cell, the size of b is in accordance with the fact of 12 pixel representing 60cm, as introduced in section 2.1. $\oplus$ and $\ominus$ represent dilation operation and erosion operation respectively, f' is the smoothed image.

$$f' = (f \oplus b) \ominus b \tag{4}$$

Thereafter, the horizontal projection is carried out to accumulate the intensity of pixels in each row using Eq. 5, here $f'(x,y)$ denotes the intensity of the pixel at position (x,y), the projection result is a 1D array P, and the P array is illustrated in Fig. (**1**).

$$P(y) = \frac{1}{W} \sum_{x=1}^{W} f'(x,y) \tag{5}$$

Then a Gaussian filter is further exploited to smooth P by using Eq.5, where the Gaussian function follows the distribution parameter $N(0,\sigma_1^2)$, as shown as:

$$P' = P * G \ , \ G = e^{-\frac{x^2}{\sqrt{2\pi\sigma_1^2}}} \tag{6}$$

We select an adaptive threshold T_0 to segment P', T_0 can be calculated adaptively by using $T_0 = u + \sigma$.

The variants u and σ are defined as the mean and the standard deviation of P', which can be presented in the form shown as:

$$u_2 = \frac{1}{H} \sum_{y=1}^{H} P'_y$$

$$\sigma_2 = \sqrt{\frac{1}{H-1} \sum_{y=1}^{H} (P'_y - u_2)^2} \tag{7}$$

Then we individually test each consecutive run-length which the value is larger than T_0, and further analyze the corresponding region R of each run-length in IPM image. If R satisfies following two criteria, it can be a candidate zebra-crossing region. First, the standard deviation of the region in unprocessed IPM image should be larger than a minimum threshold T_1 (we will set experimentally). Second, the height of the region should fall in the range of 80~140 pixels, which is equal to a range of 400cm~700cm in 3-D world coordination. As mentioned above, the former is aiming to filter the highlight region owing to severe illumination conditions, and the latter helps determine other markings which may have effect on zebra-crossing recognition.

Base on the above processing, we obtain a few candidate zebra-crossing regions, where the true zebra-crossing is included with a high probability. However, it is inevitable that some false alarm regions are still involved. These regions will be identified effectively in the following stage.
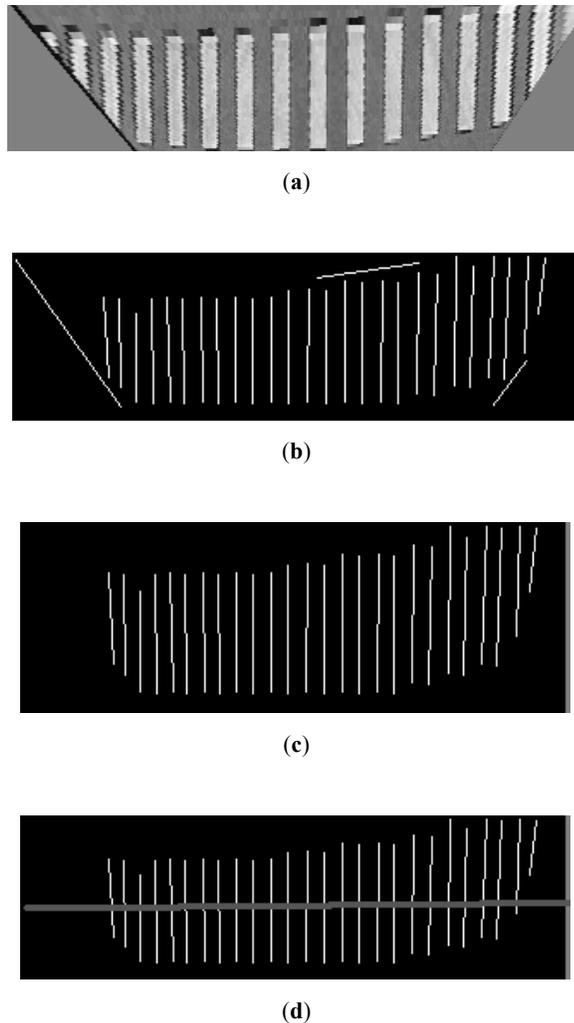
(a)

(b)

(c)

(d)

Fig. (6). Zebra-crossing recognition.

3.3. Zebra-Crossing Recognition *Via* Self-Similarity Model

The self-similarity model has been described in section 2. However, the selection of the prior pattern and the improvement of matching efficiency have not been concerned. In this section, we will resolve these problems by applying LSD algorithm. LSD is a linear time line segment detector with a control number of false detection. Once parameters are set, it requires no parameter turning while achieves a good performance. Fig. (**6b**) shows a detection result from the image in Fig. (**6a**).

We select parallel line segments from these line segments which satisfy three criteria:

(1) The direction of these line segments should be vertical or approximately vertical in IPM, since the strips of zebra-crossing are along the road direction.

(2) The distance of any two neighboring straight line segments should be approximately the same. This is effective to distinguish other road markings such as the region,

which consists of turning markings and parts of lane markers, The width of strips and the gap between two strips are usually not uniform.

(3) For any pair of line segments, they should largely overlap in the vertical direction of IPM.

We apply multi-RANSAC method [12] by involving the above conditions to select parallel line segments. Fig. (**6c**) shows detection results of parallel line segments.

If a lot of parallel line segments are detected, we only have to matching along the horizontal center line of them as shown in Fig. (**6d**). We randomly select six neighboring parallel lines with uniform intervals as a prior pattern to initialize the self-similarity model, and obtain the identification of the candidate sub-region.

In the end, we combine all the true sub-regions of a zebra-crossing and obtain their maximum boundary rectangle. The maximum boundary rectangle obtained is transformed to 2-D image coordinate for display using the inverse transformation of Eq.5.
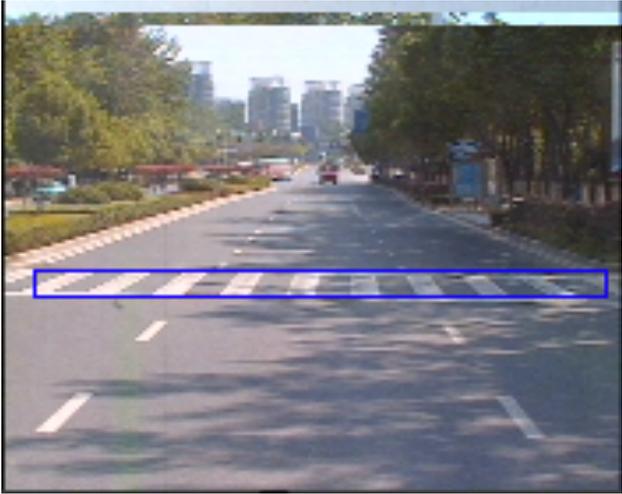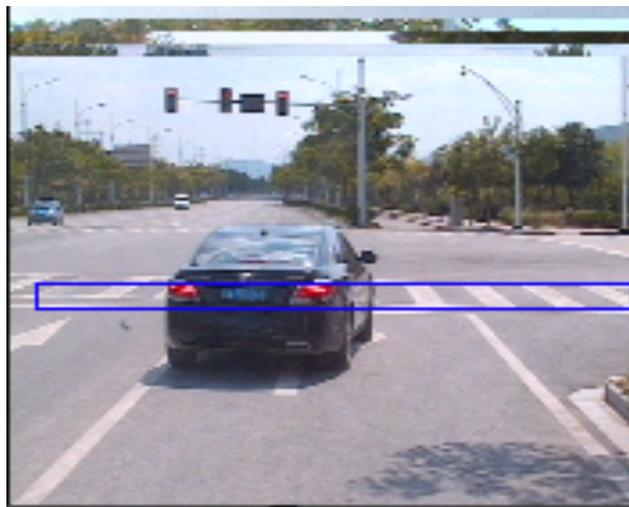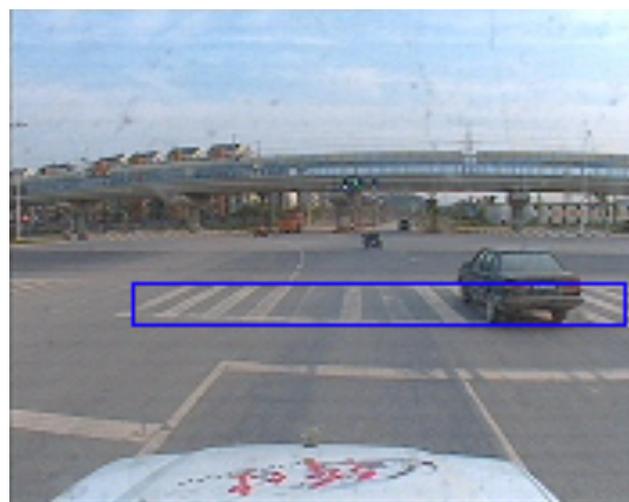
(**a**)
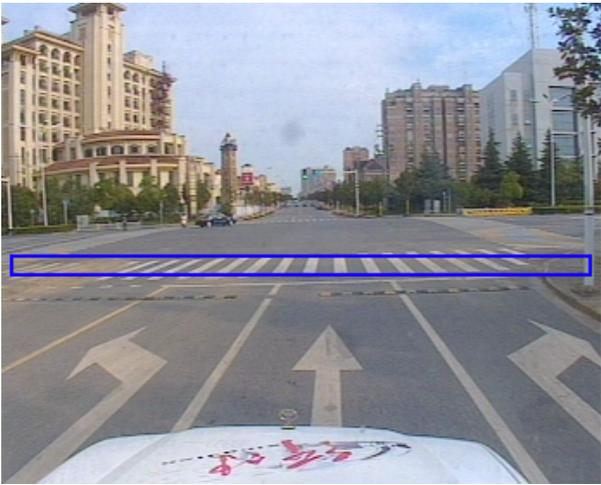
(**b**)

(**c**)

Fig. (**7**). Contd….

(**d**)

(**e**)

(**f**)

Fig. (**7**). Contd….

(**g**)

(**h**)

(**i**)

Fig. (**7**). Contd….

(**j**)

(**k**)

(**l**)

Fig. (**7**). Contd….

(**m**)

(**n**)

(**o**)

Fig. (7). Some detection results from our proposed method (**a-b**: Morning; c-d: Noon; **e-f**: Afternoon; **g-i**: Partial Occluded; **j-m**: False Negative cause of out-of-shape, fuzziness, distance; **n-o**: False Positive cause of overlap, reflection).

Table 1. Quantitative recognition result of zebra-crossing detection (%).

Scene	Morning	Noon	Afternoon	Partial Occluded	Reflection
Frames	940	1540	960	450	360
False Positive	13	18	21	8	14
False Negative	18	14	23	26	9
Precision	98.5%	98.8%	97.7%	98.1%	96.0%
Recall ratio	98.0%	99.1%	97.5%	94.1%	97.4%
F-Measure	98.2%	98.9%	97.6%	96.0%	96.7%

4. EXPERIMENTS

4.1. Experimental Setup

Experiments are conducted in a PC with CPU-AMD2.0, memory size-1G. The VC++6.0 software is used. We also use Zhang's method [1] to obtain the calibration parameter and further calculate the projection matrix for the inverse perspective mapping. The obtained projection matrix is as follows:

$$H = \begin{bmatrix} 0.0036 & 0 & -0.6420 \\ 0.0002 & -0.0039 & 1.4354 \\ 0 & 0.0017 & -0.1946 \end{bmatrix} \quad (8)$$

We also set $\sigma_1^2 = 4$ for the scale of Gaussian filter and $T_1 = 10$ for the minimum deviation threshold.

4.2. Dataset

A dataset with 500 images is constructed. All images are from real videos taken by a WAT-1000 camera rigidly mounted in the front of our autonomous lane vehicle. The dataset includes scenes with different weather and illumination conditions, as well as unclear and being partially occluded zebra-crossings. The image resolution is 352 by 288.

4.3. Evaluation Metrics

An overlapping ratio α is defined to judge whether or not a zebra-crossing has been detected.

$$\alpha = \frac{R_1 \cap R_2}{R_1 \cup R_2} \quad (9)$$

Where R_1 and R_2 denote the area of the Ground Truth region and the detected region respectively. Thus $\alpha > 0.8$ means a successful detection.

As for the evaluation of different methods, we employ three metrics, precision, recall and F-measure. Precision reflects the rightness of the detect results and recall indicates the ability of retrieving the desired information. F-measure is a balance between them. They are defined as:

$$Precision = \frac{TP}{TP+FP}, \ Recall = \frac{TP}{TP+FN},$$

$$Fmeausre = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \quad (10)$$

3.4. Results

Some typical detection results are reported in Fig. (**6**), here we use blue rectangle for representing zebra-crossing region. The images in the first row show detection results of zebra-crossing with different distance. Detection results in different scenes of variant illumination and partial occlusion are shown in the second and the third row respectively.

The quantitative result is shown in Table **1**. The average accuracy amounts to 98.9%. In our experiments, most of incorrect detections due to the distance is too far for the IPM to detect, and the original road markers are severely unclear or almost being occluded. Finally, we test the average time cost of our method with all images of the dataset. The average time cost of one frame is 57.24 ms, which shows that our proposed method is time efficient.

CONCLUSION

In this paper, we proposed a zebra-crossing detection method based on IPM. The method performs in a coarse-to-fine manner. Candidate zebra-crossing regions are first extracted via a morphology filter and the horizontal projection. A self-similarity based method is then used to identify the true zebra-crossing region. In the future, we will combine traffic light and traffic sign recognition techniques with our zebra-crossing detection approach for robust crossing road scene understanding.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflicts of interest.

ACKNOWLEDGEMENTS

REFERENCES

[1]   Volodymyr, C. James, and H. Y. Shen, "Detecting and location crosswalks using a camera phone," *Computer vision and pattern recognition workshops(CVPRW'08),* Anchorage,AK, USA , June 2008, pp.1-8.

[2]   Volodymyr, C. James, and H. Y. Shen, "Real-time walk light detection with a mobile phone," *Proceedings of the 12th international conference on computers helping people with special needs(ICCHP'10)*, Vienna, Austria, July 2010, pp.229-234.

[3]   Volodymyr, C. James, and H. Y. Shen, "Crosswatch: A camera phone system for orienting visually impaired pedestrians at traffic intersections," *Proceedings of the 11th international conference on computer helping people with special needs (ICCHP'08),* Linz, Austria, July 2008,   pp.1122-1128.

[4]   B. Jason, "Crosswalk detection utilizing lane following algorithms in OpenCV/Android[J]," https://wiki.csc.calpoly.edu, 2012.

[5]   S. Stephen, "Zebra-crossing detection for the partially sighted," *Proceedings of international conference on Computer vision and pattern recognition(CVPR'00),* Hilton Head Island, SC, June 2000, pp.211-217.

[6]   A. Borkar, M Hayes, and M.T. Stmth, "Detecting lane markers in complex urban environments," *IEEE 7th international conference on mobile Adhoc and sensor systems*, San Francisco, CA, Nov 2010, pp.727-732.

[7]   M. S. Javadi, M. A. Hannan, S. A. Samad and A. Hussain, "A robust vision-based lane boundaries detection approach for intelligent vehicles," *Information Technology Journal*, 2012, vol.11, no. 9, pp.1184-1192.

[8]   R. Y. Tsai, "An efficient and accurate camera calibration technique for 3D machine vision," *IEEE international conference of comput-*

[9]   er vision and pattern recognition,* London, UK, June 1986, pp. 364-374.

[9]   Z. Y. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. on pattern analysis and machine intelligence*, 2000, vol.22, no. 11, pp.1330-1334.

[10]   Y. Alex, R. Deepu, K. L. Maylor and R. Sausanto, "Object recognition by discriminative combinations of line segments, ellipses, and appearance features," *IEEE Transactions on pattern analysis and machine intelligence*, 2012, vol.34, no.9, pp.1758-1772.

[11]   G. Rafael, J. Jeremie, M. Jean and R. Gregory, "LSD: A fast line segment detector with a false detection control," *IEEE Transaction on pattern analysis and machine intelligent*, 2010, vol.32, no.4, pp.722-732.

[12]   M. Zuliani, C. S. Kenney and B. S. Manjunath, "The multi-RANSAC and its application to detect planner homographies," *International Conference on Image Processing*, Genoa, Italy, Sep 2005, 3: III-153-6.

[13]   P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *IEEE conf. on Computer vision and pattern recognition (CVPR'01)*, Kauai, HI, Dec 2001, pp.511-518.

[14]   T. Ojala, M. Pietikainen, and D. Harwood, "A comparative study of texture measures with classification based on feature distributions," *Pattern recognition*, 1996, vol.298, pp.51-59.

[15]   S.U. Mohammad and S. Tadayoshi, "Detection of pedestrian crossing using bipolarity feature-an image based technique," *IEEE Transactions on Intelligent transportation systems*, 2005, vol.6, no.4, pp.439-445.

[16]   S.U. Mohammad and S. Tadayoshi, "Robust zebra-crossing detection using bipolarity and projective invariant," *Proceedings of the Eighth International Symposium on Signal Processing and Its Applications, Philadelphia*, PA, USA, Aug 2005, pp.571-574.