MVA2005 IAPR Conference on Machine VIsion Applications, May 16-18, 2005 Tsukuba Science City, Japan
3-29

Detection of Pedestrian Crossing using Bipolarity and Projective Invariant

Mohammad Shorif Uddin
Dept. of Mech. and System Eng.
Kyoto Institute of Technology
Sakyo-ku, Kyoto 606-8585, Japan
E-mail: shorif@yahoo.com

Tadayoshi Shioyama
Dept. of Mech. and System Eng.
Kyoto Institute of Technology
Sakyo-ku, Kyoto 606-8585, Japan
E-mail: shioyama@kit.ac.jp

Abstract

A novel image-based approach for detecting pedestrian crossings to enhance the safety and mobility of blind people while crossing a road is described. This is achieved through bipolarity-based segmentation and projective invariant-based recognition. A thorough evaluation of the proposed approach was conducted using 100 real images with and without crossings. It was found that the proposed technique perform with 96% accuracy and with no dangerous error.

1 Introduction

According to the World Health Organization statistics, approximately 40 million people are blind all over the world [1]. Independent navigation is the main barrier for these vision-disabled people. Usually, blind people use a white cane as a travel aid. The range of detection of special patterns or obstacles using a cane is very narrow. To improve the versatility of the white cane, various devices have been developed such as the SONICGUIDE [5], the Mowat sensor [2], the Laser cane [4] and the Navbelt [9]. However, these devices are not able to assist the blind at a pedestrian crossing. Some traffic lights have beepers, which prompt the blind person to cross the road, when it is safe to do so. But, such equipment is not available at every crossing; perhaps, it would take too long for such equipment to be installed and maintained at every crossing. Blind people obviously can not see, but can hear. The arrival of fast and cheap digital portable laptop computers with multimedia computing to convert audio-video streams in real time opens new avenues to develop an intelligent navigation system for blind people.

For a safe road crossing, at first a blind person needs the information about the location of a crossing i.e. whether his frontal area is a crossing or not, then an idea about its length and the state of traffic lights. Previously, we developed crossing length measurement techniques [6], [7], [11], which are based on an assumption that the image contains a crossing. Therefore, detection of the existence of a crossing is an important preprocess.

Stephen Se [10] first proposed a pedestrian crossing detection by grouping lines and checking for concurrency using the vanishing point constraint. However, a thorough evaluation of this technique is not performed yet and also the technique is working slow and far from real time. Meijer's [8] "vOICe" consisting of a head-mounted camera, stereo headphones and a laptop, is the only commercially available vision based travel aid that uses 1-to-1 image-to-sound mapping. Though it can recognize the walls, doors etc., but pedestrian crossing detection technique is still absent in it.

In this paper, the existence of crossing is detected by bipolarity-based segmentation and projective invariant-based recognition. To confirm the effectiveness of the proposed method, experiment is performed using 100 real images with and without crossings.

2 Detection Principle

In a crossing, the usual black road surface is painted with constant width periodic white stripes. In Japan, the width of each white or black band is 45 cm. Some images of real road scenes containing pedestrian crossings are shown in Figs. 2(a)–2(e). The size of images is (width $\times$ height) = (640 $\times$ 480) pixels. The crossing pattern can be treated as bipolar region.

The proposed technique includes three major steps. First, extraction of the candidate crossing area based on bipolarity feature. Then checking the candidate area for the appropriateness of its position in the image as well as crossing direction, as the technique is searching for a frontal crossing. Second, extraction of feature points on the central vertical line of the candidate area. Third, taking the final decision of crossing or not crossing using projective invariant. As bipolarity and projective invariant features are used as the main tools in the proposed algorithm, we describe successively these two in the following subsections.

2.1 Bipolarity

We denote the intensity distribution of an image block as $p_0(x)$. If the block contains only black and white pixels, then $p_0(x)$ can be written as $p_0(x) = \alpha p_1(x) + (1-\alpha)p_2(x)$, where $0 \le \alpha \le 1$, $p_1(x)$ is the intensity distribution of black pixels and $p_2(x)$ is the intensity distribution of white pixels. Let define some variables as:

$$\mu_i = \int_{-\infty}^{\infty} x p_i(x)dx, \quad \sigma_i^2 = \int_{-\infty}^{\infty} (x-\mu_i)^2 p_i(x)dx, \quad (i = 0,1,2), \tag{1}$$

where μ_i and $\sigma_i{}^2$ represents mean and variance, respectively. Using the above relations, we can write $\sigma_0{}^2$ as

$$\sigma_0^2 = \alpha\sigma_1^2 + (1-\alpha)\sigma_2^2 + \alpha(1-\alpha)(\mu_1-\mu_2)^2. \quad (2)$$

Equation (2) shows that the total variance consists of the weighed sum of variances and the difference of means. If $\sigma_0^2 \approx \alpha(1-\alpha)(\mu_1-\mu_2)^2$, then $p_0(x)$ can be said almost bipolar. So, we define the bipolarity γ as

$$\gamma \equiv \frac{1}{\sigma_0^2}\left\{\alpha(1-\alpha)(\mu_1-\mu_2)^2\right. \quad . \quad (3)$$

Equation (3) implies that $0 \leq \gamma \leq 1$. If $\gamma = 1$, there are α, p_1 and p_2 such that $\sigma_1 = \sigma_2 = 0$. This means $p_1(x) = \delta(x-\mu_1)$ and $p_2(x) = \delta(x-\mu_2)$. So, $\gamma = 1$ corresponds to perfect bipolarity and $\gamma = 0$ represents the absence of bipolarity. Let,

$$d = |\mu_1-\mu_2|, \quad \sigma_1 \leq \frac{|\mu_1-\mu_2|}{n} \ \text{and}\ \sigma_2 \leq \frac{|\mu_1-\mu_2|}{n}.$$

Then, (3) implies

$$\gamma \geq \frac{\alpha(1-\alpha)d^2}{d^2\left\{\frac{1}{n^2}+\alpha(1-\alpha)\right.} = 1 - \frac{1}{1+\alpha(1-\alpha)n^2}. \quad (4)$$

Therefore, using (4), one can estimate a lower bound of γ taking α as a parameter. For example, γ must be greater than 0.8 when there are p_1 and p_2 such that $\alpha = 0.5$ and $n \geq 4$.

2.2 Projective Invariant

According to projective transformation [3], the cross ratio of four collinear points' Euclidean distances

$$I \equiv \frac{l_{12}l_{34}}{l_{13}l_{24}}, \quad (5)$$

is invariant, where l_{ij} is the Euclidean distance between two points i and j.

Earlier we mentioned that a pedestrian crossing is characterized by equal width periodic white and black stripes. Let denote the width of each crossing band is b. Consider feature points, which are edge points of white bands, then using (5) we can write the projective invariant for four consecutive feature points of a pedestrian crossing as

$$I \equiv \frac{l_{12}l_{34}}{l_{13}l_{24}} = \frac{b\cdot b}{2b\cdot 2b} = 0.25. \quad (6)$$

If there are n feature points, we can find $(n-3)$ projective invariants using (5). For each projective invariant $I(k),\ k = 1,2,\cdots,n-3$, we check whether $I(k)$ is within 10% of 0.25. That means,

$$\left|\frac{I(k)}{0.25} - 1.00\right| < 0.10. \quad (7)$$

We have used this 10% tolerance on the basis of experimental data.

3 Method for Detection of Existence of Crossing

A block diagram of the proposed technique is shown in Fig. 1. At first, the color image is converted to a gray scale image. Then the image is partitioned into equal-size rectangular blocks of size (16×16) pixels.

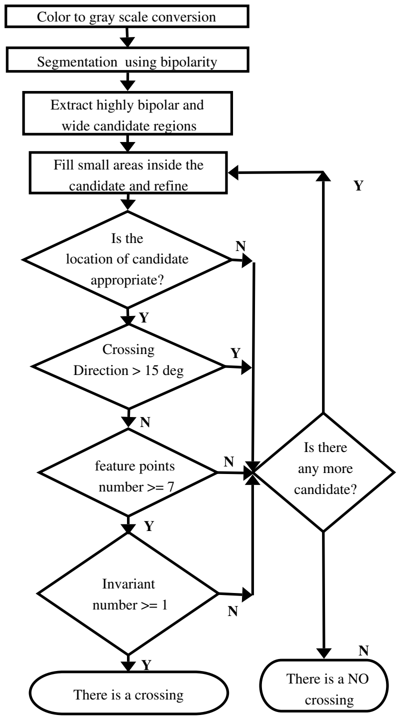

Figure 1: Flow diagram of the proposed method.

We follow the following two steps to find the crossing region candidates.

1. This step identifies homogenous bipolar regions. Segmentation is carried out by merging neighboring blocks of similar pattern on the basis of intensity distributions.
2. Only keep regions that are sufficiently bipolar and wide. Calculate the bipolarity of each segmented region. First, determine the largest region that have bipolarity higher than 0.80. Then extract the candidate regions those have bipolarity greater than 0.80 and areas more than 50% of the largest region's area.

   Then, we follow the following steps in checking each candidate region.
3. This section refines the segmentation. First, check the largest area candidate region. If there exists a small region of different label within this candidate region, then fill it with same label if its area is less than 5% of the candidate area. Then refine the bottom boundary region of the crossing area. The refinement is done by assigning same label to the pixels on a different label horizontal line if its leftmost and rightmost pixels lie in the candidate region.
4. Make sure the region is centered in the image. Check the position (either it is in appropriate position or too

left or too right or too far w. r. t. the observer) of the candidate region. If the total candidate region lies in a position, which is situated within 20% of image width from the left or right boundaries, then the position of the candidate region is treated too left or too right, respectively. If y-coordinate of the bottommost point of the candidate region is situated more than 40% of image height from the bottom boundary, then the position of the candidate region is treated too far from the observer. If the candidate region is not in an appropriate location then decide there is no crossing for this candidate and go to examine the next candidate region (if any).

5. As a blind person is interested in detecting the frontal crossing not the sided one, therefore, estimation of crossing direction is important. Calculate the crossing direction [6], [7] using the image content at the location of the candidate region with the help of the two-dimensional fast Fourier transform. If the crossing direction is greater than a threshold (15° is used here on the basis of experimental data) then decide there is no crossing for this candidate, as the crossing direction is too steep and go to examine the next candidate region (if any).

6. This section extract feature points and then take the decision on the basis of projective invariants. Binarize the original image content at the location of the candidate region using the mean. Use median filtering of window size (5×5) pixels on the binarized image to eliminate sporadic noise. Determine the center position of the candidate region. Then extract the feature points (which are the edge points of the white bands) on the vertical line passing through the center point of the candidate region. If there are at least 7 feature points, then go for checking the projective invariant criterion. This condition ensures that there are more than 3 white bands exist in the crossing region. Some road marking (not crossing) consists of three painted stripes. An image of this type is shown in Fig. 2(f). Hence the above condition will safeguard against false positive detection result. If there is at least one invariant satisfying the invariant condition (7) then decide there is a crossing, otherwise go to examine the next candidate region (if any).

4 Results and Discussions

To evaluate the performance of the proposed method we used 100 real images with different backgrounds taken by a commercial digital camera. Among them, 66 images include crossing and the rest 34 are of no crossing. Fig. 2 shows some samples of experimental images. Fig. 3 presents the different steps of an experimental image of Fig. 2(a) in detecting the existence of a crossing.

The complete result of the detection of the existence of crossing is summarized in Table 1. From this table, we see that the proposed algorithm is quite successful in detecting the existence of crossings from real road images. The method has not made any dangerous (false positive) error such that it decides the existence of a crossing for a scene without crossing. For a scene containing crossing where the white paintings on the crossing are damaged or the scene contains too few crossing bands - very short crossing (less than 4 white bands) then it decides nonexistence of crossing (i.e. false negative). These have happened only in 4 cases. Among them, 2 cases contain less than four white bands - an image of these type is shown in Fig. 2(c). A case shown

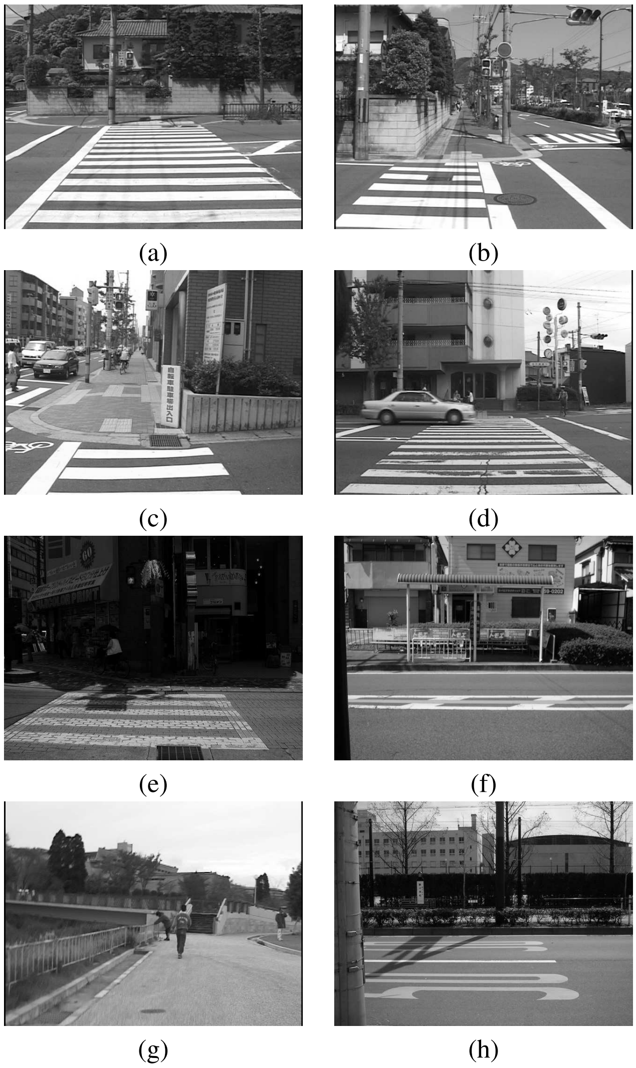

Figure 2: Some experimental images: crossing is detected in (a) and (b), false negative is obtained in (c)–(e), no crossing is detected in (f)–(h).

Table 1: Crossing Detection Result Summary.

Decision	Image with crossing	Image without crossing
crossing	62 (Ok)	0 (False positive)
No crossing	4 (False negative)	34 (Ok)
Total number of images	66	34

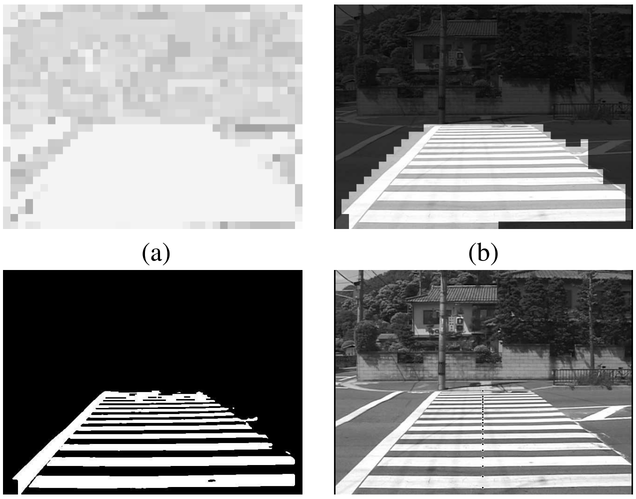
(a) (b) (c) (d)

Figure 3: Different steps in detecting the existence of a crossing for the image of Fig. 2(a): (a) bipolarity of each segmented region, (b) candidate crossing area, (c) binarized crossing area (d) extracted feature points marked on the image.

in Fig. 2(d) where the white paintings on the crossing are damaged, gives false negative result. The rest case shown in Fig. 2(e) also gives false negative result. The segmentation step could not extract any highly bipolar wide area in this image, as the white paintings of crossing are done on unusual pattern of road surface.

If the crossing is occluded by vehicles or other obstructions then the method fails to detect the crossing. Therefore, detection of vehicles needs to be included with the present system as a preprocess. If there are vehicles on the image, then a message will be delivered to the user to take another image. In principle, as the technique is using bipolarity, so it is robust from the viewpoint of different illumination environments, such as cloudy, sunny or rainy situations at morning, noon, afternoon and even evening, if the illumination is uniform in the image. In the rainy situation, the method is supposed to be applicable provided crossing stripes are not occluded by water. Though in the present paper we could not include images of different illuminations (i.e. images of evening or night or rainy day), in our future paper we will report about these situations in details.

5 Conclusions

In this paper, a simple image-based method for the detection of existence of crossing has been proposed. Experimental results with real road scenes confirmed its effectiveness. The method has not made any dangerous (false positive) error such that it decides the existence of a crossing for a scene without crossing. For a crossing scene where the white paintings on the crossing are damaged or the scene contains too few crossing bands, then it decides nonexistence of crossing (i.e. false negative). These have happened only in 4 cases among 100 experimental images. We hope this technique will help in improving the mobility of blind people. Our final goal is to put the complete road crossing system in a device like a mobile phone set.

Acknowledgements

The authors gratefully acknowledge the suggestions and comments of Professor Yasuo Yoshida and Mr. Tadashi Matsuo of Kyoto Institute of Technology, Japan. They are also grateful for the support of Japan Society for the Promotion of Science under Grants-in-Aid for Scientific Research (No. 16500110 and No. 03232).

References

[1] B. Thylefors, A. D. Negrel, R. Pararajasegaram, K. Y. Dadzie, "Global data on blindness," Bulletin of World Health Organization **73**, pp.115-121, 1995.

[2] D. L. Morrissette et al., "A follow-up study of the Mowat sensor's applications, frequency of use, and maintenance reliability," *J. Vis. Impairment and Blindness* **75**, pp. 244-247, (1981).

[3] I. Weiss, "Geometric invariants and object recognition," Int. J. Comp. Vision, **10**, no. 3, pp. 207-231, 1993.

[4] J. M. Benjamin, "The new C-5 laser cane for the blind," *in Carnahan Conf. on Electronic Prosthetics*, pp. 77-82, (1973).

[5] L. Kay, "A sensor and to enhance spatial perception of the blind, engineering design and evaluation," *Radio Electron. Eng.* **44**, pp. 605-629, (1974).

[6] M. S. Uddin and T. Shioyama, "Measurement of the length of pedestrian crossings - a navigational aid for blind people," *Proc. IEEE Intl. Conf. on Intelligent Transportation Systems (ITSC2004)*, Washington, USA, Oct. 2004 pp. 690-695.

[7] M. S. Uddin and T. Shioyama, "Measurement of pedestrian crossing length using vector geometry - an image based technique," *IEEE International Midwest Symposium on Circuits and Systems (MWSCAS 2004)*, Hiroshima, Japan 25-28 July 2004, Vol. I, pp. 229-232.

[8] P. B. L. Meijer, "Vision technology for the totally blind," [Online] http://www.seeingwithsound.com/

[9] S. Shoval, J. Borenstein, Y. Koren "Auditory guidance with the navbelt-a computerized travel aid for the blind," *IEEE Trans. Syst. Man Cybern.* **C28**, pp. 459-467, (1998).

[10] Stephen Se, "Zebra-crossing detection for the partially sighted," *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition (CVPR)*, South Carolina, USA, Jun. 2000, vol. 2, pp. 211-217.

[11] T. Shioyama, H. Wu, N. Nakamura, S. Kitawaki, "Measurement of the length of pedestrian crossings and detection of traffic lights from image data," *Meas. Sci. Technol.* **13**, pp. 1450-1457, (2002).