Computer Vision and Artificial Intelligence Based Navigation on Top-view Virtual Plane

Hui Zeng¹, * and Jiaqi Luo²

¹School of Artificial Intelligence, Wuchang University of Technology, Wuhan 430223, Hubei, China
²School of Intelligent Manufacturing, Wuhan Guanggu Vocational College, Wuhan 430079, Hubei, China

*Corresponding author: 120100385@wut.edu.cn

Abstract. This paper shows a vision-based obstacle detection approach in an outdoor sidewalk environment. There is a major problem of applying single camera in outdoor navigation mission, that is the differentiation of various types of barriers and clutter pavement. To settle this issue, the property that pavement margins are sub-sampled while barrier margins are oversampled when the original image is mapped to the virtual plane of the top view. Morphology filters are applied to strengthen barrier margins as margin-blobs with greater size, while sparse margins from pavement are screened. On basis of the recognized barriers, secure walking areas are predicted by tracking a polar margin-blob histogram. The test of algorithm is made in various sidewalk scenes with complicated pavements with confirmed potency.

Keywords: Robotics, outdoor navigation, single camera, morphology filters.

1. Introduction

1.1. Related works

Recently, as many vision-based navigation systems are developed, blind people have been guided in some local environments [1]. Some of these systems use stereo vision approaches, where a depth map of the nearby context is produced by stereo cameras, and this map of depth is changed into stereo sound or haptic instruments for the blind’s self-navigation [2-4].

Despite the system based on stereo-vision, systems applying single camera were put forward. By comparing with stereo cameras, single camera can be maintained compactly and easily. Several of these systems based on mono-vision paid attention to the recognition of target pixels from background pixels. The critical part of this method is the extraction of features which can discriminate various types of obstacles in complex background. For instance, in NAVI system put forward by Sainarayanan et al. [5], a fuzzy learning vector quantization (LVQ) neural network is used for classifying target and background pixels. Despite a good prospect of the rate of classification in an indoor context, the LVQ classifier is educated with an assumption that the background color is lighter than the barrier color, which may not be applicable in outdoor environment.
1.2. Suggested Approach Summary
In the suggested approach, a camera is fastened to the belly of blind users while looking down at the road ahead. The fundamental concept of barrier recognition is to distinguish from barrier pixels with pavement pixels. By comparing with Sainarayanan’s approach using pixel-wise characteristics, it is found that margin-based characteristics discriminate barriers with road pavement. Through mapping the original figure to the virtual plane of the top view, the road pavement margins in the nearby area are sub-sampled, while barrier margins located in the far area are over-sampled. Then, the morphology filter is adopted to enhance the effect of the non-uniform resampling on association and ratio of edges. For a safe walking area, the histogram of a polar margin-blob is calculated. Part of histogram where the largest valley happens is recognized and trailed for the biggest region with no barrier margin-blob.

2. Obstacle detection
2.1. Top-view plane mapping
Inverse perspective mapping (IPM) [6] is adopted to map the original figure to a virtual plane of the top view.

![Figure 1. Camera model for top-view change.](image)

To be mathematical, the modelling of IPM is the projection of a 3D Euclidean space $W$, with factors $(x, y, z) \in \mathbb{R}^3$, onto a 2D planar subspace of $\mathbb{R}^3$, indicated by $I$, with factors $(u, v) \in \mathbb{R}^2$.

The change from $W$ to $I$ is given as:

\[
\begin{align*}
  u(x,0,z) &= \frac{\arctan \left( \frac{h \sin r(x,0,z)}{z - d} \right) - (\theta_b - \tilde{\theta}_v)(m-1)}{2\tilde{c}_v} \\
  v(x,0,z) &= \frac{\arctan \left( \frac{z - d}{x - f} \right) - (r_0 - \tilde{\theta}_u)(n-1)}{2\tilde{c}_u}
\end{align*}
\]

(1)

The significant coefficients in (1) are shown below:

① Camera position $C = \{l, h, d\}$ in global coordinate system is used to define camera viewing point.

② Two angles $r_0$ and $\theta_0$ describe camera viewing orientation, as shown in Fig.1.

③ Camera angular aperture means $2\alpha_u$ in the orientation of row and $2\alpha_v$ in the orientation of column. $\alpha_u = \arctan(u_0/f)$, $\alpha_v = \arctan(v_0/f)$, $(u_0, v_0)$ means the focal center of the camera, and $f$ refers to focal length.

④ Camera resolution means $n \times m$. 
Fig. 2 shows top-view mapping result obtained by applying (1) to a perspective-view figure. On the original figure margin map shown in (c), barrier’s edges with those pavement margins around can be discriminated hardly. Nevertheless, on the top-view margin map displayed in (d), barrier margins are strengthened by oversampling, while sub-sampling suppress pavement margins.

2.2. Obstacle margin-blob extracting

After top-view re-sampling, the barrier margins are strengthened with regard to scale and connectivity. For the further emphasis of this effect, edge-blobs with large size is extracted by combining morphology operations with connected component analysis. These margin-blobs are identified as candidate barrier representatives.

Here, pavement margin sections with an opening operation are removed by a 3×3 rectangular structure element, then the closing procedure will be conducted to fill the gaps within existing
foreground pixels. Then, the associated foreground pixels are grouped into blobs through an association part labeling operation. Blobs smaller than the predefined threshold will be abandoned.

According to Fig.3(c), opening operation eliminates a lot of small edge segments from pavement, and closing procedure fixes the foreground blob shape. Finally, according to Fig.3(d), only three main margin-blobs corresponding to reasonable barrier areas are chosen.

3. Safe waling area prediction

3.1. Polar margin histogram

On basis of the identified barriers, a polar margin histogram is established on the top-view plane for estimating safe walking region. According to Fig.4(c), on the margin map, from the right bound to the left bound, the sampling of radial orientations (labelled with red dash line) about the convergence point C is made. For the radial direction of each sample, the counting of margin-blob pixels lying along this orientation is made. Through the accumulation of all the radial orientations sampled, a polar margin histogram can be established according to Fig.4(d).

In this figure, the horizontal axis means the radial direction of sampling by angles, and the vertical axis refers to the margin pixels number along each sampling orientation angle. The bins with large values demonstrate the orientations where barriers happen, while that with zero values amount to the orientations with no barrier. Hence, the bins with zero values are used to eliminate safe-area.

![Figure 4. Safe-area prediction.](image)

3.2. Polar margin histogram pursual

For initialization pursual, the grouping of connected zero value bins in frame $t$ as $V_t^i$ is made. Which is followed by organization in accordance with their group size $|V_t^i|$. Then, the greatest bin group $MAX(|V_t^i|)$ is chosen as the pursual group $V_t^i$ in frame $t$. In the next frame, the zero-value bin group $V_t^i$ most near the pursual group $V_t^{i-1}$ in frame $t-1$ is chosen as the pursual group $V_t^i$ in frame $t$. In case of smaller pursual group $|V_t^i|$ than a threshold, pursual will be ceased and re-started from the start. To evaluate the safe-area more steadily, a one dimensional Kalman filter is used to smooth the output bounding orientation values.

4. Experimental results

For the evaluation of the barrier detection precision, the algorithm is tested on several side-walk video clips, with complicated road side construction and disordered road surface. Take 2000 frames for test.
All the key barrier positions are marked on the top-view figures of these frames manually. If the barrier is identified within a 10 pixels deviation from the true location, it is right identification, otherwise it is wrong. If the barrier is not identified, it is deemed to be miss. If several road surface clutters are recognized as barriers, it is deemed to be false. Table 1 shows the barrier detection outcome.

| Total barriers | Correct Detection | False Detection | Miss Detection |
|----------------|-------------------|-----------------|---------------|
| 9486           | 7788 (82.1%)      | 1660 (17.5%)    | 38 (0.4%)     |

5. Conclusion
A vision-based navigation algorithm for the blind in outdoor context is proposed. By comparing with the approach based on general stereo-vision, the suggested system deals with this issue applying just single camera. Future work will involve developing data transformation scheme to process the data got from the figure to language domain, and convey the navigation instructions via audio information to the blind user in a suitable way.

Acknowledgements
Scientific Research Project of Education Department of Hubei Province (B2020258) supported this research.

References
[1] D. Dakopoulos and N. G. Bourbakis, “Wearable obstacle avoidance electronic travel aids for blind: a survey,” IEEE Trans. Syst., Man, And Cybern., vol. 40, no. 1, pp. 25–35, Jan. 2010.
[2] L. A. Johnson and C. M. Higgins, “A navigation aid for the blind using tactile-visual sensory substitution,” in Proc. 28th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., NewYork, 2006, pp. 6298–6292.
[3] N. Bourbakis, “Sensing 3D dynamic space for blind,” IEEE Engineering in Medicine and Biology Magazine, Vol. 27, No. 1, 2008, pp. 49-55.
[4] D. Dakopoulou, S. K. Boddhu, and N. Bourbakis, “A 2D vibration array as an assistive device for visually impaired,” in Proc. 7th IEEE Int. Conf. Bioinf. Bioeng., Boston, 2007, vol. 1, pp. 930–937.
[5] G. Sainarayanan, R. Nagarajan, and S. Yaacob, “Fuzzy image processing scheme for autonomous navigation of human blind,” Appl. Softw. Comput., vol. 7, no. 1, pp. 257–264, Jan. 2007.