Non-Protruding Hazard Detection for the Aged Vision-Impaired

Aravinda S. Rao*, Jayavardhana Gubbi†, Marimuthu Palaniswami* and Elaine Wong*

* Department of Electrical and Electronic Engineering, The University of Melbourne, Parkville, VIC 3010, Australia.

† Tata Innovation Labs, Tata Consultancy Services, Bengaluru, Karnataka - 560066, India.

aravinda@student.unimelb.edu.au,{palani, ewon}@unimelb.edu.au, j.gubbi@tcs.com

***Abstract*—Usage of the traditional white cane by the elderly with vision impairment is inefficient as many are also reliant on ambulatory aids such as wheelchairs and walking frames. The fall occurrence when using ambulatory aids is higher, contributed by *non-protruding hazards* such as potholes and drop-offs. Currently available technology for blind navigation, predominantly based on proximity sensing, is not designed to *detect non protruding hazards*. We address this critical need by developing a new optical laser system that combines innovative approaches in optical laser projection, vision-sensing, pattern recognition, and machine learning. Here, we present an overview of the system, including a new feature descriptor termed Histogram of Intersections, and results from our pilot study where pothole detection is achieved with over 90% accuracy.**

I. Introduction

The World Health Organization estimates that worldwide, there are currently about 285 million visually impaired people, with a staggering 82% aged 50 and above [1]. When vision loss is permanent and uncorrectable, one's quality of life of people is profoundly impacted. To ensure healthy aging and continued social interaction, developing confidence and skills to get around independently is critical. However, traditional mobility aids such as the white cane are unsuitable for many, especially those that are also reliant on ambulatory aids, e.g. wheelchairs, and walking frames. Many elderly with late onset vision impairment fall under this category.

In the last decade, there has been substantial research in the development of electronic travel aids (ETAs) to obtain visual and spatial information for blind navigation. Using a combination of sensors such as sound, ultrasound, and stereo-vision [2], most ETAs are efficient at detecting protruding obstacles but *lack the capability to detect non-protruding hazards*, e.g. potholes, cracks, uneven surfaces, drop-offs, descending stairs, and curbs. These hazards are major fall risks to those that rely on ambulatory aids [3]. Outside ETA research, several efforts in determining potholes and cracks have been implemented for asphalt-surfaced road maintenance [4]. These include 3D laser scanning and stereo-vision, 2D vision imaging to segment road anomalies, and using mechanical response through vibrations. Converting the above-mentioned approaches into ETA solutions is unsuitable due to the combination of high computational complexity, high cost, large equipment footprint, need for high illumination level and image quality, and the impracticality of requiring measuring devices to pass through road anomalies before detection data can be generated.

This work is supported by the Ian Potter Foundation - Health and Disability Grants.

Here, we propose an innovative optical laser-based detection system that addresses the above deficiencies to detect non-protruding hazards in addition to protruding obstacles. Our system relies only on monocular vision sensing and exploits the Hough transform with a new feature descriptor termed Histogram of Intersections (HoI) to achieve > 90% accuracy in pothole detection. Our system is also targeted for night time scenarios for round-the-clock navigation safety. To the best of our knowledge, this work represents the first steps towards establishing a low-cost and fast-responding system for non-protruding hazard detection for the blind.

II. System Overview

The schematic of our proposed optical laser-based detection system is shown in Fig. 1. Laser patterns from a structured laser are projected directly onto the surface in front of the ambulatory aid. Deformed patterns striking the surface are then recorded by a single HD camera to extract surface information. For the pilot study, algorithms and decision-making were implemented in OpenCV 2.4 on a Linux machine (Ubuntu 14.04 LTS). The underlying phenomenon to accurately detect safe/unsafe paths hinges on the capability to extract features from the captured video frames, i.e. breaks/fading in the intersection points of the deformed laser patterns. Steps taken to achieve the above are described next.

III. Feature Extraction

1) Preprocessing: Let $I(x, y; t)$ denote a video frame at time t, where $I(x, y)$ represents a particular pixel of the frame. Each video frame is firstly preprocessed by a bilateral filter with $\sigma_d = 5$ and $\sigma_r = 10$ to remove noise, preserve edges, and to reduce variations in the color pixels to enhance the sharpness of the video frames.

2) Hough Transform: Using Hough transform [5], laser lines that pass through the (x,y) coordinates in an image space are detected by exploiting known parameters (ρ,θ) in the parametric space. The output of the Hough transform provides a number of lines intersecting at any point (x, y) of $I(x, y)$ in the form of an accumulator matrix (ρ,θ). That is, the ρ and $\theta \in [0,\pi]$ combination indicates the number of intersecting sinusoidal lines present in the image space.

3) Histogram of Intersections (HoI): The intersecting points are then binned to create a histogram for each video frame. We use our novel feature description, HoI, to serve as the prime feature descriptor for this purpose. Let $(\rho,\theta)^T = R_\theta = [n_i,$

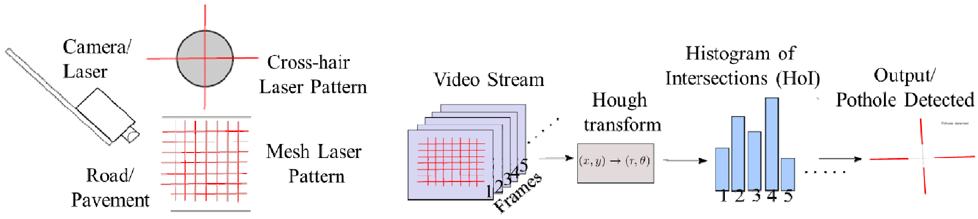

Fig. 1 Schematic of optical laser detection system for hazard detection.

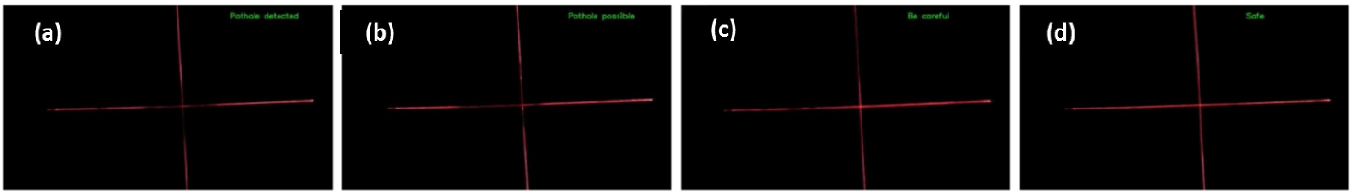

Fig. 2. Experiment results (crosshair) using HoI: (a) Pothole detected (|H| = 1), (b) Pothole possible (|H| = 2), (c) Be careful (|H| = 3), and (d) Safe (|H| = 4)

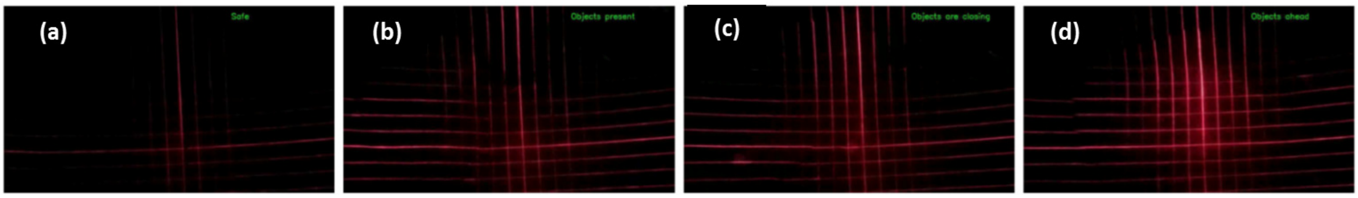

Fig. 3. Exp (mesh) using HoI: (a) safe (11 ≤ |H| < 60), (b) Objects present (61≤ |H| < 90), (c) Objects closing (91 ≤ |H| < 120), (d) Objects ahead (|H| > 121).

$n_{i+1},\ldots,n_N]$ where $n_i \in \Re$ indicates the number of lines intersecting for each ρ_i, θ_i combination. The HoI is therefore:

$$H_t = \sum_{i=1}^{N} R_\vartheta^{\,T} Q \qquad (1)$$

where $Q = [q_i, q_{i+1},\ldots,q_N]$, $q_i \in \Re$ is a row vector indicating the threshold values that determines as to which elements of an accumulator matrix is considered. For a video of υ frames, the HoI is given by:

$$\mathrm{HoI} = \mathrm{H} = [H_1, H_2, \ldots H_\upsilon]^T \quad (2)$$

4) Learning and Classification: Using the Gaussian Mixture Model (GMM) [6], a nonlinear relationship between the number of intersecting points and the type of event encountered was established to form classes of path cues, e.g. *Pothole Detected*, *Objects Closing*, *Safe*, *etc.*. A GMM consists of a finite number (K) of convex combination of Gaussians to form a parametric probability density function, and can be represented as:

$$p(\mathbf{x}|\theta) = \sum_{i=1}^{K} w_i \cdot g(\mathbf{x}|\mu_\mathrm{i}, \mathbf{\Sigma}_\mathrm{i}) \qquad (3)$$

where the Gaussian components are given by:

$$g(\mathbf{x}|\mu_\mathrm{i}, \mathbf{\Sigma}_\mathrm{i}) = \frac{1}{(2\pi)^{D/2}|\sum_i|^{1/2}} e^{\left\{\frac{1}{2}(\mathbf{X}-\mu_\mathrm{i})^T \sum_i^{-1}(\mathbf{X}-\mu_\mathrm{i})\right\}} \qquad (4)$$

Parameter X is a visible variable, K is the number of mixtures, and $\sum_i \omega_i = 1$. The mean value μ_i in (3) and (4) was used to distinguish amongst different path cue classes, and the initial number of clusters that was input to the GMM for learning was based on k-means clustering of real-life sample videos.

IV. Implementation and Results

Experiments were conducted using two laser projections: (a) cross-hair (Edmund Optics CW laser diode at 3mW and 635 nm); and (b) mesh (GS1 Laser Grid). Data was collected in HD mode (720p) using GoPro HERO4 Silver and with a narrow field of view. The laser-camera mount was ~ 1 m to the surface and angled at ~ 30–45° to surface normal. Example results obtained when using crosshair and mesh patterns are shown in Figs. 2 and 3, respectively. The number of intersecting points corresponding to each class of path cue is provided in the figure captions. In our pilot study, our proposed system achieved over 85% accuracy overall in classification output and over 90% in detecting potholes.

V. Summary

We presented the first solution in detecting non-protruding hazard detection by combining innovative approaches in vision-sensing, pattern recognition, and machine learning. Our solution is low cost and fast responding, requiring only one single video frame to compute a reliable decision.

References

[1] World Health Organization, “Visual impairment and blindness,” http://www.who.int/mediacentre/factsheets/fs282/en/, 2014, [Online: accessed on July 23, 2015]
[2] D. Dakopoulos, et. *al.*, “A wearable obstacle avoidance ETAs for blind, ” *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 40, no. 1, pp. 25-35, 2010.
[3] http://www.who.int/ageing/projects/AMRO-Jamaica.pdf [Online: accessed on 10 Sept, 2015]
[4] T. Kim and S.-K. Ryu, “Review and Analysis of Pothole Detection Methods,” *J. Emer. Trends in Comp. Inf. Sci*, vol. 5, no. 8, pp. 603-608, 2014.
[5] J. Illingworth and J. Kittler, “The adaptive hough transform,” *IEEE Trans Pattern Anal. Mach. Intell.*, vol. PAMI- 9, no. 5, pp. 690–698, 1987.
[6] K. P. Murphy, *Machine learning: a probabilistic perspective*. MIT press, 2012, ch. 11, Mixture models and the EM Algorithm.