Context-aware obstacle detection for navigation by visually impaired

Pedram Gharani, Hassan A. Karimi

PII: S0262-8856(17)30098-7
DOI: doi:10.1016/j.imavis.2017.06.002
Reference: IMAVIS 3624

To appear in: *Image and Vision Computing*

Received date: 22 January 2016
Revised date: 9 March 2017
Accepted date: 16 June 2017

Please cite this article as: Pedram Gharani, Hassan A. Karimi, Context-aware obstacle detection for navigation by visually impaired, *Image and Vision Computing* (2017), doi:10.1016/j.imavis.2017.06.002

Context-aware Obstacle Detection for Navigation by Visually Impaired

Pedram Gharani and Hassan A. Karimi
Geoinformatics Laboratory, School of Information Sciences, University of Pittsburgh, USA

This paper presents a context-aware smartphone based visual obstacle detection approach to aid visually impaired people in navigating indoor environments. The approach is based on processing two consecutive frames (images), computing optical flow, and tracking certain points to detect obstacles. The frame rate of the video stream is determined using a context-aware data fusion technique for the sensors on smartphones. Through an efficient and novel algorithm, for each consecutive frame, a point dataset on the frames is designed and evaluated to check whether the points belong to an obstacle. In addition to determining the points based on the texture in each frame, our algorithm also considers the heading of user movement to find critical areas on the image plane. We validated the algorithm through experiments where the proposed algorithm was compared against two comparable algorithms, one sparse and one that was predefined and grid-based. The experiments were conducted in different indoor settings, and results based on precision, recall, accuracy, and f-measure were compared and analyzed. The results show that, in comparison to the other two widely used algorithms for this process, our algorithm is more precise, and the parameters also indicate some other improvements. We also considered time to contact parameter for clustering the points and presented the improvement of the performance of clustering by using this parameter.

Keywords: Obstacle detection, optical flow, context-aware

1 Introduction

Moving and traveling in unfamiliar environments are basic-and challenging-daily activities for visually impaired people (Marston and Golledge, 2003). Lack of preview, low or no knowledge of the environment, and limited access to information related to orientation are three major limitations to their independent movement and overall travel (Golledge and Marston, 1997; Harper et al., 2003; Helal et al., 2001; Marston and Golledge, 2003; Roentgen et al., 2008). To address these shortcomings, a large number of products, systems, and assistive devices have been developed which could be classified into two major groups: obstacle detection or micro-navigation systems and macro-navigation systems (Katz et al., 2012). Micro-navigation systems enable people who are blind or visually impaired (BVI) to safely navigate through the environment. Macro-navigation systems generally deal with localization, path planning, representation, and interaction (Fallah et al., 2013). In other words, macro-navigation systems assist BVI people by localizing them and providing turn-by-turn navigation instructions, while micro navigation systems detect the objects that can be a barrier for the BVI people.

There are numerous approaches and techniques helping people who are BVI handle navigation challenges. Canes and guide dogs are two commonly used approaches to assist BVI users. However, a cane has some shortcomings, like no protection against obstacles in upper part of the body, as well as a limited preview about the ambient environment (Farmer and Smith, 1997). In general, both guide dogs and canes are not able to detect overhead objects (Cui et al., 2010). A more flexible approach lies in using Electronic Travel Aids (ETA) (Bousbia-Salah et al., 2011; Ganz et al., 2014, 2011; Guerrero et al., 2012; Ran et al., 2004; Zheng et al., 2014). There are different types of ETAs. Some are based on optical triangulation, such as LaserCane (Benjamin et al., 1973), while others use an acoustic-triangulation method, such as GuideCane (Borenstein, 2001). Another example is the Minoru 3D webcam ("Minoru 3D webcam," 2010) plus a laptop, an ad-hoc ETA based on stereo vision, which can recover a full depth map.

In this paper, we focus on addressing the problem of obstacle detection in indoor environments using computer vision techniques for processing the frames and sensors (accelerometers, gyroscope, magnetometer) on smartphones for activity recognition and contextual-aware frame extraction. Using these sensors and a single camera, we are not able to directly measure or estimate depth. In the area of robotics, depth estimation and geometric modeling of the ambient environment are possible by using stereo camera (imagery) (Cheng et al., 2010; Murray and Little, 2000; Nakju et al., 2004), and some other sensors such as