Available online at www.sciencedirect.com
ScienceDirect
October 2016, 23(5): 88–96
www.sciencedirect.com/science/journal/10058885
The Journal of China Universities of Posts and Telecommunications
http://jcupt.bupt.edu.cn

Vision-based positioning system

Song Meina (✉), Ou Zhonghong, E Haihong, Song Junde, Zhao Xuejun

School of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876, China

Abstract

Conventional outdoor navigation systems are usually based on orbital satellites, e.g., global positioning system (GPS) and global navigation satellite system (GLONASS). The latest advances from wearable, e.g., BaiduEye and Google Glass, have enabled new approaches to leverage information from the surrounding environment. For example, they enable the change from passively receiving information to actively requesting information. Thus, such changes might inspire brand new application scenarios that were not possible before. In this work, we propose a vision-based navigation system based on wearable like Baidu Eye. We discuss the associated challenges and propose potential solutions for each of them. The system utilizes crowd sensing to collect and build a traffic signpost database for positioning reference. Then it leverages context information, such as cell identification (Cell ID), signal strength, and altitude combined with traffic sign detection and recognition to enable real-time positioning. A hybrid cloud architecture is proposed to enhance the capability of sensing devices (SD) to realize the proposed vision.

Keywords vision-based, positioning system, wearable, machine vision

1 Introduction

The outdoor navigation systems have been dominated by the space-based satellite systems in the past decades. Amongst the available solutions, GPS has the widest coverage and gains the most adoption. The advantages of GPS-like navigation systems include providing the best-in-class positioning accuracy, and are resilient from weather conditions. Nevertheless, the disadvantages of such systems are also obvious:

1) In urban areas with high buildings blocking the line-of-sight to satellites, i.e., the so-called 'urban canyons', the positioning error can be up to hundreds of meters [1].

2) The initial acquisition time is in the scale of tens of seconds (10 s~20 s) from our experiments [2], which is consistent with previous work [3].

In this paper, we propose a vision-based positioning system to solve the problems mentioned above, which we name human-centric positioning system. A human-centric positioning system revolves around human rather than machine, and the knowledge of human and the context information can be injected into the system to improve the quality of experience of human. Wearable devices, such as BaiduEye and Google Glass, offer great potential towards this vision. Such devices integrate first-person image capturing (Google Glass can capture a picture simply by winking eyes twice), video shooting, sensing, processing, and communication capabilities into the same package. Combined with real-time interpretation, a system can be built to realize the human-centric positioning system mentioned above.

The human-centric positioning system has a set of advantages. Firstly, it is analogous to realistic scenarios and closer to human behaviors (people only ask questions when they have doubts). Secondly, it is more preferable and less intrusive from privacy perspective. Thirdly, it is energy-efficient because it only needs to be triggered for the time period when the user has explicit needs, which is essential for battery-powered devices, e.g., smartphones and wearable. Lastly, it can be seamlessly integrated with existing GPS-based systems to improve positioning

Received date: 14-12-2015
Corresponding author: Song Meina, E-mail: mnsong@bupt.edu.cn
DOI: 10.1016/S1005-8885(16)60062-6

accuracy.

We can image a hypothetical scenario of human-centric navigation system: Edward is wearing a BaiduEye (or Google Glass) and driving towards a place of interest. When he is close to the destination, he is confused which exit to take from the freeway, because there are several exits near by. By winking his eyes towards certain traffic signpost, the positioning system instantly locates his position and chooses the correct exit for him accordingly. When it is getting dark and Edward passes by some hotel advertisements, he recalls that he has forgotten to book a hotel room. By winking his eyes towards the advertisements once again, a hotel room with the highest performance/cost ratio is ready for him to check-in.

Despite the advantages mentioned above, realizing such a vision is highly challenging. The challenges arise from various aspects:

1) How to collect traffic signposts to build a database that contains sufficient number of signposts of decent-quality for accurate positioning?

2) How to accurately position a user that can meet the stringent real-time requirements?

3) How to enable the vision using smartphones and wearable whose capability is still relatively weak regardless of the recent rapid advancements?

In this paper, utilize a multidisciplinary approach to solve the problems mentioned above. Specifically, we utilize mobile crowd sensing to collect and build traffic sign database. Use wearable like BaiduEye or Google Glass to express explicitly user needs (capturing pictures of traffic signs). Leverage computer vision techniques for traffic sign detection and text recognition. Integrate user context information (e.g., Cell ID, wireless fidelity (WiFi) and altitude) to significantly improve the detection and recognition speed. Offload computation from mobile devices to one-hop cloud to enhance the capability of SD. In the following sections, we will discuss the challenges in detail, and lay out the potential solutions towards each of them.

2 Related work

The related work with his paper can be roughly divided into two categories. One is the general traffic sign detection and recognition scheme, the other is related to utilizing wearable to enhance human capabilities, specifically, utilizing vision-based techniques to provide a positioning system.

1) Traffic sign detection and recognition scheme. Traffic sign detection and recognition scheme has experienced significant improvement in the past decade. This can be shown by two competitions organized in Refs. [4–5]. In general, the best-in-class approaches can achieve better than 95% detection and recognition rate. For some sub-glass traffic signs, the detection and recognition rate can reach 100%. This provides us with solid state-of-the-art techniques to realize our vision. Nevertheless, all of these existing work focuses on detecting/recognizing a single traffic sign only, e.g., give way or speed limit sign, While we are interested in the overall shape of the compound signpost and the relative locations of its child signposts. Thus, novel algorithms have to be designed.

2) Utilizing wearable. The work from Ha et al. [1] is the most relevant to this work. Ha et al. [1] build an assistive system based on Google Glass devices for users with cognitive decline. Nevertheless, cognitive assistance is a completely different application domain than the positioning system proposed in this work. The real-time requirement from this work is also much more stringent than the cognitive assistance system. There is a landmark-based system for pedestrian navigation using geotagged photos [3]. Nevertheless, the landmarks used in Ref. [3] are usually high buildings etc., which limits the applicability of such systems because such landmarks are not always available. In this paper, we utilize traffic signposts for positioning, which are easily available on roadsides. Furthermore, the real-time requirement from pedestrian scenario is completely different from vehicular positioning system. Thus, novel algorithms and new techniques have to be developed to solve the research challenges in this work.

3 Building traffic signpost database

Collecting and building a database from traffic signposts is valuable for various applications, including highway maintenance, traffic sign recognition, driver-assistance system, and intelligent autonomous vehicles. Building such a database, however, is time-consuming and needs a large amount of work. Firstly, the completeness of database should be as complete as possible. Namely, it should contain most (if not all) of the traffic signposts available along or above roadside to provide good