Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 115 (2017) 466–473

Procedia Computer Science

www.elsevier.com/locate/procedia

7th International Conference on Advances in Computing & Communications, ICACC-2017, 22-24 August 2017, Cochin, India

A Global Image Descriptor Based Navigation System for Indoor Environment

Athira Sudarsanan V.[a], Michael George[a*], Babita Roslind Jose[a], Jimson Mathew[b]

[a]*School Of Engineering, Cochin University of Science And Technology, Kochi, India 682022*

[b]*Department of Computer Science and Engineering, Indian Institute of Technology Patna, India 801103*

Abstract

We present a vision based indoor navigation system based on the GIST global image descriptor. The use of lightweight GIST descriptor enables faster computation and low memory requirement. Our system takes care of the key tasks of creating map and localization of user. This work also focuses on finding destination and displaying route from current location to destination. A topological map is created from keyframes that are extracted from a walkthrough video of the indoor environment, using the GIST descriptor. The L2 norm of the GIST descriptor is used for both the keyframe extraction as well as localization.

Peer-review under responsibility of the scientific committee of the 7th International Conference on Advances in Computing & Communications.

Keywords: *Keyframe;Indoor navigation;GIST;Routing;Localization*

1. Introduction

Navigation is the task of finding a path from one place to another place. There are three purposes for navigation: to find current location, to determine the destination and how to reach destination from current location. The navigation system can be classified into two main categories: outdoor navigation and indoor navigation. For outdoor navigation, people currently rely on GPS (Global Positioning System). But in an indoor environment GPS system fails because of non-reception of GPS signals in indoor setting. Although techniques for indoor navigation [1], [2] based on Infrared (IR), WiFi etc. have been proposed, they have not yet attained the precision of GPS. In this work, we concentrate on vision based technique which works on low-fidelity images which would be more apt for practical usage. Vision based navigation systems can be classified as metric map and topological map [3].

The system we present could find potential use [4], [5], [6], [7] in a variety of application areas like: accessibility aids for the visually impaired people, shopping malls including hypermarket, hospitals, hotels, airports, railway stations, indoor robotics, tourism etc. The rest of the work is divided into sections on system architecture, results and

* Corresponding author. Tel.: +91-9447539225; *E-mail address:* michaelgeorge@cusat.ac.in

1877-0509 © 2017 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the scientific committee of the 7th International Conference on Advances in Computing & Communications
10.1016/j.procs.2017.09.086

conclusion.

2. System architecture

The entire task of the system is divided into five as shown in Fig. 1. First the developer will insert the walkthrough video into the system. From this video, the system will extract the important frames (keyframes) and construct the topological map. Then the map will be stored in the database of the system. Whenever the user communicates with the system it will take data stored in the database and perform localization along with routing towards destination. In this system, an appearance based image descriptor called GIST [8], [9] is used to perform both localization and map construction. Other existing feature point descriptors like SIFT produce errors if too few or too many features are detected. This situation can be occur in different environments; for example if the light condition has been changed, the algorithm will detect key points from unwanted areas like shadows projected from the roof. Under the same condition, GIST performs well because it calculates descriptor from frequency domain. In addition to this, SIFT require more memory and large computation time since it detects every important feature from a particular image.

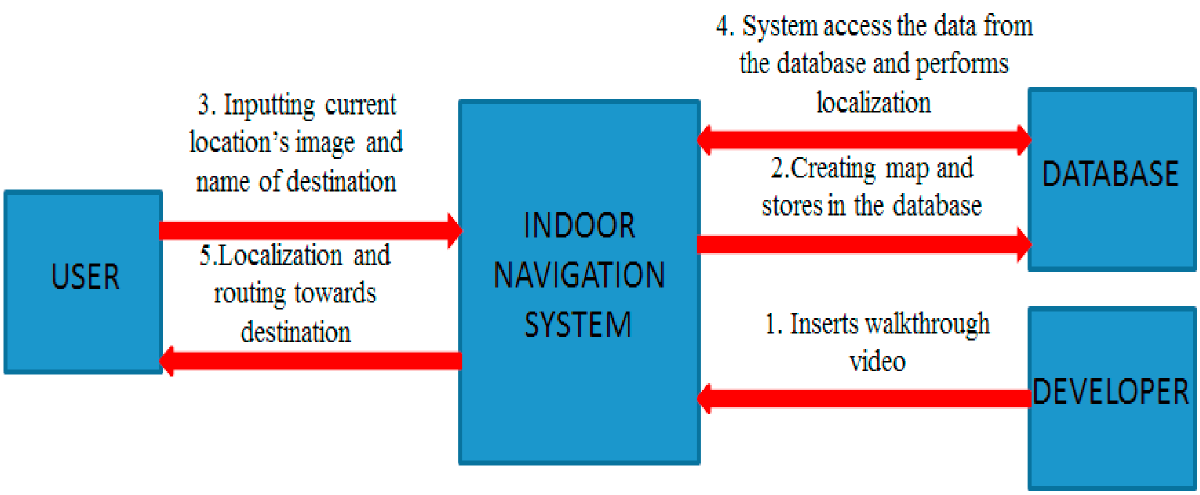

Fig. 1. Overview of indoor navigation system

2.1 Keyframe extraction

The main idea of this part is to capture the important frames called keyframes from the given indoor walkthrough video. For that the system will first extract the entire frames from the walkthrough video and find the corresponding GIST descriptor of each frame. Then the system will find the L2 norm between two of these descriptors. Whenever this L2 norm become greater than a particular threshold, the frame will be treated as a keyframe. The concept is visualized in Fig. 2. Here 'th' is the threshold value. The GIST descriptors of frames 1, 2, 3,...N, N+1, N+2 are gd_1, gd_2,gd_3,…gd_N,gd_{N+1},gd_{N+2} respectively. Frame 1 being the first frame is selected as a keyframe. The frames 2, 3,........N-1 have the GIST distance less than the threshold (th) when compared to frame 1 (previous keyframe). The frame N is found to have the GIST distance greater than the threshold, and hence selected as the next keyframe.