Neurocomputing 173 (2016) 83–92

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Where am I in the dark: Exploring active transfer learning on the use of indoor localization based on thermal imaging

Guoyu Lu [a,*], Yan Yan [b], Li Ren [a], Philip Saponaro [a], Nicu Sebe [b], Chandra Kambhamettu [a]

[a] Department of Computer and Information Sciences, University of Delaware, USA
[b] Department of Information Engineering and Computer Science, University of Trento, Italy

ARTICLE INFO

Article history:
Received 20 March 2015
Received in revised form
9 June 2015
Accepted 5 July 2015
Available online 10 August 2015

Keywords:
Image-based localization
Active transfer learning
Thermal imaging

ABSTRACT

Indoor localization is one of the key problems in robotics research. Most current localization systems use cellular base stations and Wifi signals, whose localization accuracy is largely dependent on the signal strength and is sensitive to environmental changes. With the development of camera-based technologies, image-based localization may be employed in an indoor environment where the GPS signal is weak. Most of the existing image-based localization systems are based on color images captured by cameras, but this is only feasible in environments with adequate lighting conditions. In this paper, we introduce an image-based localization system based on thermal imaging to make the system independent of light sources, which are especially useful during emergencies such as a sudden power outage in a building. As thermal images are not obtained as easily as color images, we apply active transfer learning to enrich the thermal image classification learning, where normal RGB images are treated as the source domain, and thermal images are the target domain. The application of active transfer learning avoids random target training sample selection and chooses the most informative samples in the learning process. Through the proposed active transfer learning, the query thermal images can be accurately used to indicate the location. Experiments show that our system can be efficiently deployed to perform indoor localization in a dark environment.

1. Introduction

The purpose of indoor localization [1] is to help the users navigate in a large and complicated environment. For example, in an airport, travelers want to receive prompt navigation information in order to board their plane. Currently, most indoor localization systems are based on GPS. The GPS signals are scattered by the roofs and walls; hence, the strength of signals is attenuated, which potentially negatively affects the localization accuracy. Another technology is based on pre-deployed beacons and requires a high beacon distribution density. Furthermore, the radio signal based indoor localization systems can usually locate a person while failing to tell the user orientation information.

Although image-based localization is mainly used in outdoor environments to make up the deficiency of weak GPS signals, recent research has studied indoor localization as well. The basic idea of indoor image-based localization is matching a query image with all the images in a database to find the nearest neighbor image in a descriptor space. Structure-from-Motion (SfM) techniques enable the 3D model to be utilized in the localization system. Based on the matching between 2D images and a 3D reconstruction model, the 3D coordinate of the camera can be returned together with the pose estimation information. SfM techniques do not require the camera to be calibrated as in stereo reconstruction; thus images used for SfM reconstruction are easier to obtain. Due to the same reason, updating the SfM reconstruction model is easier than other techniques. A user only needs to capture an image and then the localization system matches the image against the 3D model in the whole localization process. Meanwhile, the 3D model can perform the functionality of a 3D map that helps users better understand the building structure for visit planning purposes. A 3D model, however, usually contains millions of descriptors, and searching through the whole descriptor space for correspondences would potentially consume a lot of time, making the system less practical.

During the Structure-from-Motion process, every image used in the reconstruction is assigned with pose estimation information, including position and orientation. Instead of searching through the whole point cloud of a 3D model, we search the nearest neighbor of the query image among the images used for the 3D SfM reconstruction and assign the pose information of the returned image to the query image. Then we can assign 3D coordinates based on the 2D image retrieved, making use of the advantages of both the 3D model and the 2D images.

* Corresponding author.
E-mail address: luguoyu@udel.edu (G. Lu).

http://dx.doi.org/10.1016/j.neucom.2015.07.106
0925-2312/

Current image-based indoor localization systems make the assumption that light resources are always sufficiently present. The reality, however, is that the lighting conditions may vary dramatically from place to place. In many emergent situations, such as a power outage of a building, there can be little to no lighting. This makes the localization and navigation tasks challenging for both robots and humans. In dealing with this problem, we propose to use thermal-infrared imagery for indoor image-based localization. Imagery from a long wave infrared camera is based on the temperature of the object and is not dependent on the lighting. Since indoor buildings are composed by different objects with different materials, e.g. glasses and wooden tables, the surface temperature of the indoor objects also varies. By recognizing the object shapes, thermal imagery is an ideal choice to perform localization tasks in a dark environment.

Unlike common RGB color images, capturing thermal images require more effort. To overcome some of the limitations of the thermal cameras, we have to invest more time into focusing and framing a scene. This process is labor intensive, which means that far fewer samples were taken. Our thermal camera has to be plugged into a PC with certain ports, so the hardware is a limitation for the ease of data collection. It is very expensive to collect sufficient images to perform accurate localization. To solve this problem, we propose performing transfer learning between color images as the source domain and thermal images as the target domain. Transfer learning based image localization aims to leverage the useful information from visible images to thermal images. During training, there are usually fewer target training samples than the source samples. By leveraging the source task, the learning of the target task is enhanced. In the traditional transfer learning, the target samples are selected randomly. However, not all the target samples are equally informative. Active transfer learning selects the most informative target samples to train the model, which provides a higher localization accuracy by avoiding learning the model through randomly selected samples. We captured color and thermal images in different locations and train the location classification model based on the active transfer learning algorithm. Here, each location is treated as a landmark group for classification. The whole process is illustrated in Fig. 1.

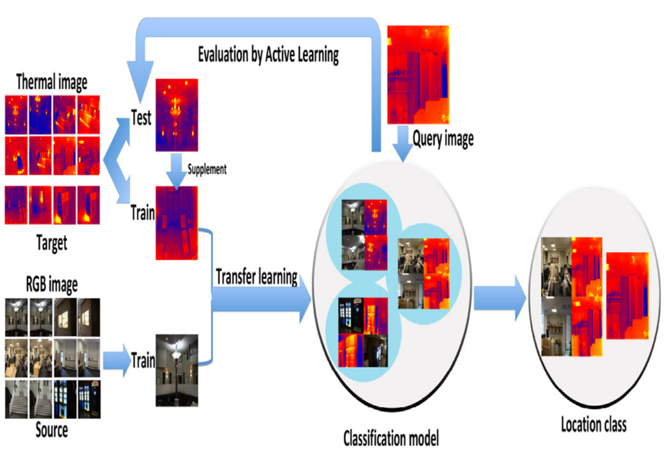

Fig. 1. Our image-based indoor localization system based on active transfer learning: the target domain is the thermal images and the source domain is RGB images. The target domain is separated into a training set and testing set. Through transfer learning, the source domain training set is adapted into a target domain to train a classification model for the thermal images. Using the trained classification model, the most informative target sample in the test set is added into the training set through active learning. Utilizing the enriched training set, the system learns a new classification model. After the target training set reaches a certain number of samples, the classification model is finalized to classify the query images into one of the location classes.

To summarize, the contributions of our paper are as follows: (1) We present a framework for solving the indoor image-based localization problem in a dark environment; (2) We apply active transfer learning on indoor image-based localization problem in order to adapt the training set to be most informative. (3) Finally, we jointly use the common RGB image and thermal images together to learn a better model for indoor localization based on thermal images.

The rest of the paper is organized as the following: Section 2 provides the related work on image-based localization, the usage of thermal image in computer vision problems and the transfer learning; Section 3 introduces our image localization system based on the transfer learning between RGB images and thermal images; Section 4 presents the proposed active transfer learning method; Section 5 provides our image localization experiments result; Section 6 discusses the failure case of our method; Section 7 concludes the paper.

2. Related work

Image-based Localization: Image-based localization [2] is widely applied to localization problems, especially in weak GPS signal areas. This paper calculates the pose of a query image by utilizing a database of building facades and associates a 3D-coordinate system with images in the database. Schindler et al. [3] selected the vocabulary [4] using informative features to improve image-retrieval performance on a large street-side image database. Xiao et al. [5] further improved localization accuracy by using geometric verification with a bag-of-words method.

3D SfM models [6,49] are used for image-based localization problems in enhancing the accuracy. Li et al. [7] used mutual visibility information for 3D-to-2D matching. Sattler et al. [1] directly matched descriptors of 2D images to descriptors of 3D model. Irschara et al. [8] proposed to retrieve images containing the most descriptors matching the 3D points. The proposed localization framework achieves a high image-registration rate by accelerating the matching process. Gronat et al. [9] changed the place recognition as a classification problem based on a classifier trained by geo-tagged images. Lu et al. [42,48] improved the localization accuracy and memory efficiency through local feature processing.

Indoor Localization: For indoor image-based localization, many different techniques have been tried. Ravi et al. [10] matched a query image to a database using color histograms. Kosecka et al. [11] detected edges of room images, generating edge histograms for each image for matching. Liu et al. [12] used a Transfer Regression Model to localize. Kawaji et al. [13] applied online transfer learning in humanoid robots for object recognition. Wannous et al. [14] proposed an automatic indexing method of the content stream of a camera mounted on the shoulder based on the presence in specific 3D places related to instrumental activities to detect the activity related places. The system in [15] involved RANSAC and applied ASIFT features to perform affine invariant image matching. Yu et al. [16] proposed combining the color, person detection, face recognition, and non-background information for localization. Lu et al. [50] separated the entire view into several directions and learned a multi-view localization system to predict location and orientation.

Transfer Learning: Transfer learning deals with applying knowledge learned from some existing tasks to new domains which share some commonalities. Sun et al. [17] used Wifi based indoor localization with transfer learning when variation in signal distributions causes the old localization model to be inaccurate. Jiang et al. [18] presented a cross-domain SVM algorithm which adapts previously learned support vectors from one domain to facilitate