Poster: WhereAmI: Image-based Positioning in Dense Urban Areas

Quoc Duy Vo, Klaus Mueller, Pradipta De
Dept of Computer Science, SUNY Korea
{rayvo, mueller, pradipta.de}@sunykorea.ac.kr

Categories and Subject Descriptors: H.3.4 [**Systems and Software**]: Distrbuted Systems

Keywords: Mobile, Locationing, OCR, Android

1. INTRODUCTION

Real time computer vision techniques, like OCR, on mobile platforms is opening new opportunities [4]. Image based locationing has traditionally focused on identifying well-known landmarks, like Eiffel Tower, in an image captured using a smartphone. However consider a user lost in the sprawling bazaars of a big city where a well-known landmark may be difficult to spot.

We propose a technique where texts appearing in names of shops, billboards, and road signs act as pseudo-markers to identify the location. Collection of different signs, which are keywords for the area, is often unique. For example, in Figure 1, the presence of "Indian Home Industries" and "Jain Gift House" together can be unique to an area. The texts in the image clicked by the user are extracted from the image to create a location signature. The signature is matched against a repository of signatures.

The repository of signatures is generated using geo-tagged images from online sources, like Flickr and Pinterest. Cities with CCTV or webcams, such as in Seoul or Singapore, can also feed images of different areas with geo-tag. The database of well-known signatures are populated by analyzing these pictures.

Unlike previous work on image based matching [1, 2], we are using *hints* embedded in the image. The cost of querying the location is minimized as few text strings need to be uploaded. Response time to the query will be faster as text matching may yield response quicker than image matching.

2. APPROACH

The design of the system, called *WhereAmI*, requires a frontend logic on the mobile phone to run OCR on image, and a backend service for matching signatures. The concept is described in Figure 1.

To extract text from an image, we are using the Tess4J OCR library [3]. While detecting the texts, we also assign weights to texts based on their relative size in the image. The weights help in guiding the search among large set of signatures. The bag of keywords is uploaded to a hosted
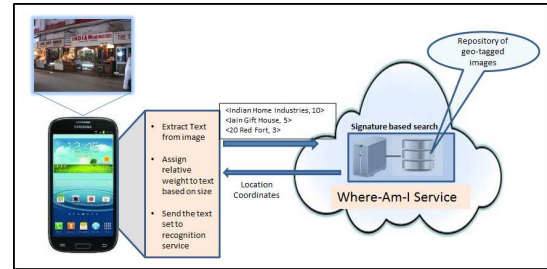

Figure 1: WhereAmI App and Service

WhereAmI service. WhereAmI service checks the repository of signatures, which are generated a priori by analyzing images with geo-tags. The effectiveness of the application hinges on fast matching of the correlated keywords.

There can be scenarios where the image does not contain enough distinguishing keywords. In this case, the user is given a feedback to snap another picture by guiding the user to point left or right.

Currently the work is exploring open source OCR tools on mobile platforms. The backend service will implement fast algorithms for matching against a database of signatures.

Acknowledgment

3. REFERENCES

[1] M. Cristani, A. Perina, U. Castellani, and V. Murino. Geo-located image categorization and location recognition. *Pattern Recognition and Image Analysis*, 19(2):245–252, Jun 2009.

[2] A. Mulloni, D. Wagner, I. Barakonyi, and D. Schmalstieg. Indoor positioning and navigation with camera phones. *IEEE Pervasive Computing*, 8(2), Apr 2009.

[3] Tess4j. http://tess4j.sourceforge.net/.

[4] G. Wang, Y. Xiong, J. Yun, and J. R. Cavallaro. Accelerating computer vision algorithms using OpenCL on the mobile GPU - a case study. In *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2013.