Exploring the Potential of Indoor Mapping and Localization Using Internet Photos

Jiang Dong, Yu Xiao, Antti Ylä-Jääski
Aalto University, Finland
firstname.lastname@aalto.fi

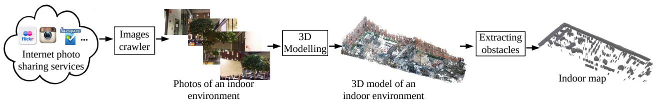
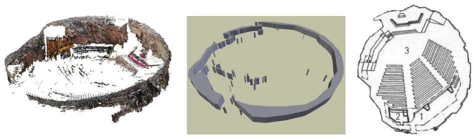


Figure 1: The process of indoor mapping.



(a) (b) (c)

Figure 2: (a) 3D model of the Rock Church. (b) Indoor map of the Rock Church generated from the 3D model. (c) Indoor map of the Rock Church retrieved from website.

1. INTRODUCTION

Millions of photos are being uploaded to Internet photo sharing sites like Flickr and Instagram every day. This brings new opportunities of image-based applications and services. In this poster, we explore the potential of using Internet photo collections for building 3D models of indoor environment and providing indoor mapping and localization using these models.

2. INDOOR MAPPING AND LOCALIZATION

We developed a prototype to generate maps of different indoor environments, as shown in Figure 1. Internet photos of indoor environments are used as the input of the system. Based on the APIs provided by photo sharing sites, photos of indoor environments can be downloaded by their key words. 3D models can be built in the form of 3D point clouds from Internet photos using Structure-from-Motion (SfM) techniques [2]. The obstacles are extracted from the point cloud by flattening it after removing the points at the ground level. Subsequently, the shapes of obstacles can be extracted by generating non-convex polygons out of unordered 2D points [3]. An indoor map is generated by extruding all shapes towards the up-facing axis and placing them on a ground plane.

As an example[1], we downloaded 1,513 photos of the Rock Church in Helsinki from Flickr in August 2014. Figure 2a shows the generated 3D model by using VisualSfM [1], a state-of-the-art SfM tool. Figure 2b shows the indoor map extracted from the 3D model. Figure 2c shows the floor plan retrieved from the website. In general, the generated indoor map clearly shows the outline of the church. However, the stage and chairs are missing in the generated indoor map when compared with the one retrieved from website.

[1]Due to the page limit, more technical details are available in our research report: https://wiki.aalto.fi/display/crowdsensing/Research+Report

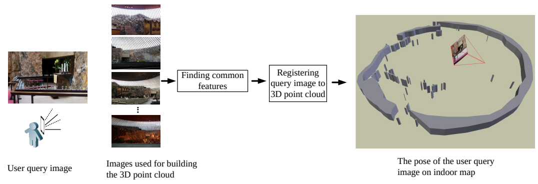


Figure 3: The process of indoor localization.

The location of the user can be calculated by registering the user query image to the 3D point cloud based on SfM. As shown in Fig. 3, the feature matching will be conducted between the query image and each of the images used for building the point cloud. If enough matches can be found, the features provided by the query image will be added to the point cloud, and the camera pose of the query image will be returned, which can then be used as the user's position and facing direction.

3. DISCUSSION

Our studies show that Internet photo collections can serve the purpose of indoor mapping and localization in the case of the Rock Church, which has relatively simple layout. However, there are still several technical challenges to be solved: (1) Lacking focal length information of photos may cause biased result in the process of mapping and localization; (2) Photos available on Internet photo sharing sites do not cover every corner of the place in question; (3) Model updating is needed since indoor scenes such as interior decoration of a shopping mall may change seasonally; (4) Building 3D models from large amount of photos using SfM is computationally expensive.

4. REFERENCES

[1] Visualsfm : A visual structure from motion system. http://ccwu.me/vsfm/.

[2] S. Agarwal, Y. Furukawa, N. Snavely, et al. Building rome in a day. *Commun. ACM*, Oct. 2011.

[3] M. Duckham, L. Kulik, M. Worboys, and A. Galton. Efficient generation of simple polygons for characterizing the shape of a set of points in the plane. *Pattern Recogn.*, 41(10):3224–3236, Oct. 2008.