Intuitive Appliance Identification using Image Matching in Smart Buildings

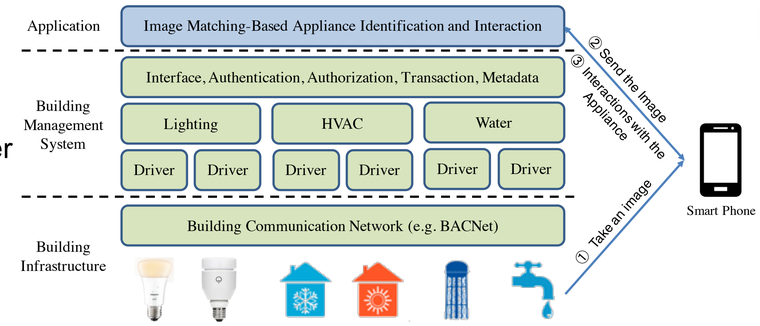
Kaifei Chen*, John Kolb*, Jonathan Fürst†, Dezhi Hong‡, Randy H. Katz*
*University of California at Berkeley, †IT University of Copenhagen, ‡University of Virginia
{kaifei, jkolb}@berkeley.edu, jonf@itu.dk, hong@virginia.edu, randykatz@berkeley.edu

Motivation

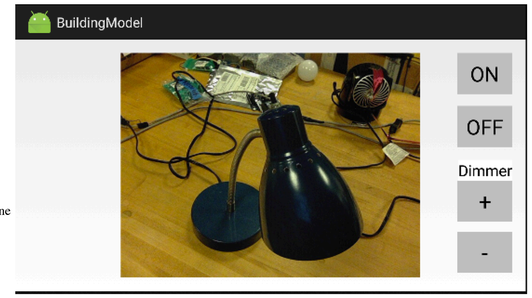
- The number of IoT appliances is growing
- Identifying which appliances to interact with through software becomes harder and more tedious for users
- Existing approaches are not intuitive
  - Some require users to describe the appliance in cumbersome ways (e.g. SQL query)
  - Some require the deployment of additional infrastructure (e.g., infrared transceivers)

Our Proposal

- We explore an intuitive way to identify appliances: ***what you see is what you interact with***:
  1. User takes a picture of a smart appliance
  2. The mobile app sends images to the server
  3. Server matches the images onto the 3D model and returns the control interface
  4. The mobile app displays the control interface for user to interact with the smart appliances through the BMS



Using Building Management System (BMS)



Mobile Application User Interface

System Overview

- Our system consists of two phases
  1. **Modeling phase**
     - The building manager collects and annotates a visual 3D building model
     - Image registration uses canonical Structure from Motion approaches
  2. **Matching phase**
     - Users identify and interact with appliances by matching the image from a smart phone to the 3D model
     - The new image is compared with its closest image in the 3D model

Modeling

Modeling Phase Overview

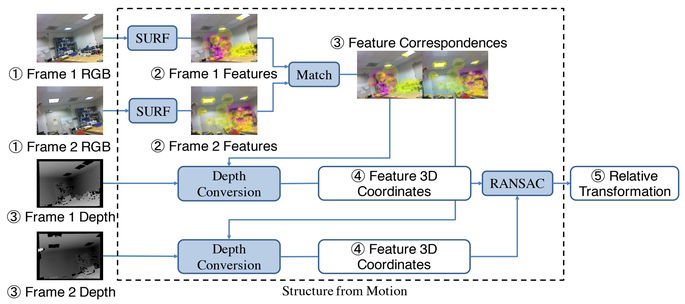
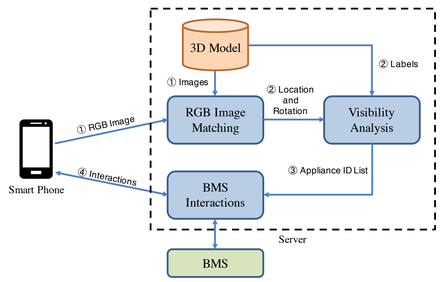
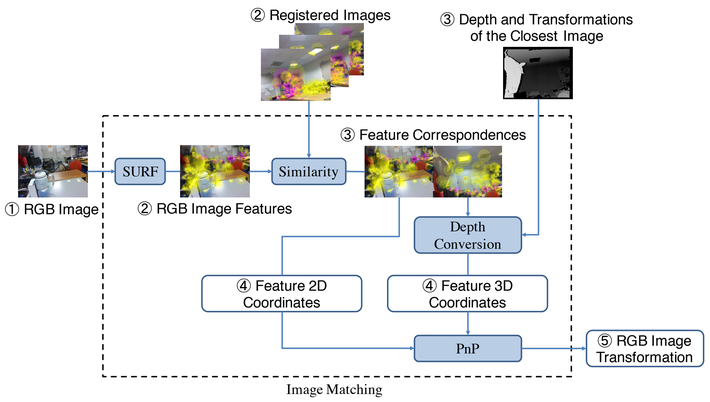


Image Registration Algorithm in the Structure from Motion (SfM) Component.

Matching



Matching Phase Overview



RGB Image Matching Algorithm

Implementation and Evaluation

- We implemented our client on Android and our server by extending the Real-Time Appearance-Based Mapping (RTAB-Map) [Labbe et al. 2014]
- We evaluate it on two datasets
  - RTAB-Map multi-session dataset collected by a robot
    - 304 images for modeling, 45 images for matching
    - 44 out of 45 images are matched
    - 100% precision and recall
  - An in-situ deployment in our lab collected by a human
    - 1075 images for modeling, 76 images for matching
    - 54 out of 76 images are matched
    - 63% have recall >0.8, and 78% have precision >0.8

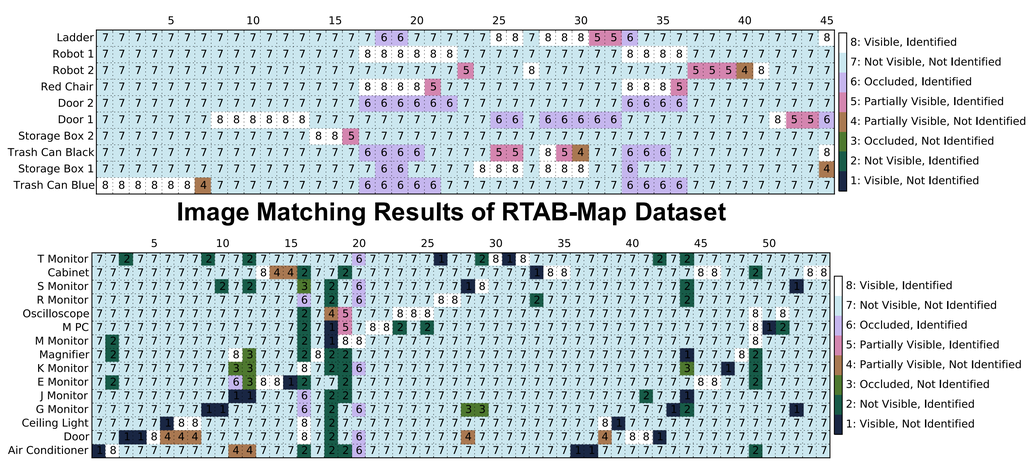
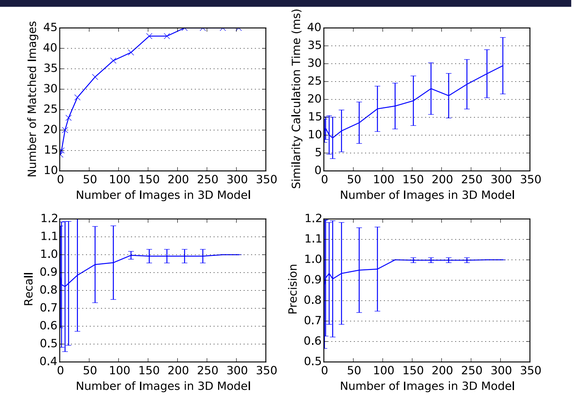


Image Matching Results of RTAB-Map Dataset

Image Matching Results of Lab Dataset



Number of matched images (top left), similarity calculation time (top right), recall (bottom left), and precision (bottom right) with different numbers of images in the 3D model