Carnegie Mellon University
Civil and Environmental Engineering

Yujie Wei, Burcu Akinci
yujiew, bakinci@cmu.edu

Unstructured image-to-BIM registration in dynamic construction scenes for digital twin modeling

Last updated: Dec 2020

Abstract

Despite the growing use of Building Information Modeling (BIM) and as-built data capturing techniques, the discrepancies between as-designed model and as-built components still exist in real-world projects, especially for occluded MEP components (Figure 1). NIST (2008) reported that it costs the US 4.8 billion dollars each year to verify and validate discrepancies during the maintenance stage.

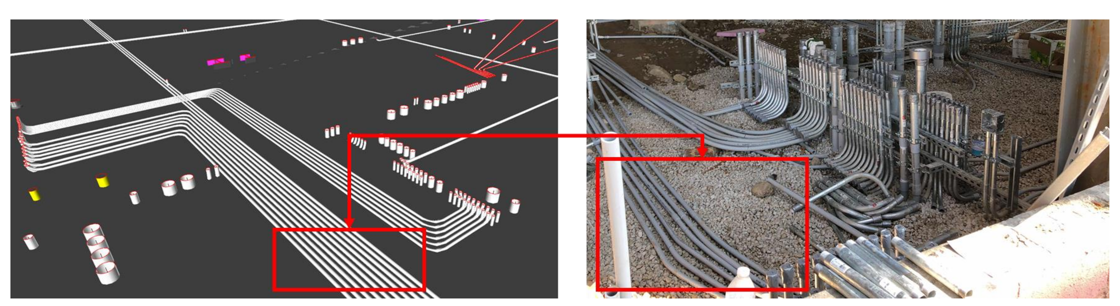


Figure 1. Discrepancies between as-built and as-designed models.

Existing approaches of mitigating such discrepancies heavily rely on manual documentation during construction and reality capturing at project closeout as shown in Figure 2. Manual documentation is error-prone and incapable of document actual geometries, while data collection at project closeout fails to captured components that have been occluded.

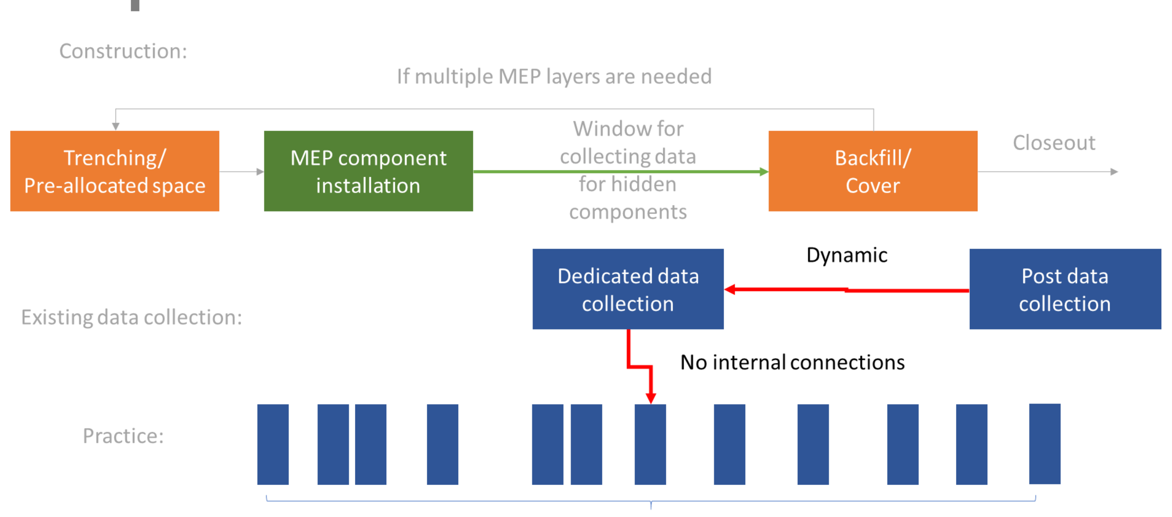


Figure 2. Existing workflows of documentation

An intuition of solving the problem above is to leverage unstructured data captured during construction (such as images) to update as-designed models. Therefore, we proposed a method to register unstructured images to as-designed models using semantic features.

Problem being Addressed

There are several engineering challenges that need to be addressed when associating unstructured images captured during construction to an as-designed model:

- **Visual difference.** The visual information presented on an image is quite different from the one shown in an as-designed model. For example, as-designed models usually lack the subtle texture that is necessary for most visual descriptor-based matching as shown in Figure 3.

Figure 3. Real-world images vs. BIM

- **Scope difference.** Unstructured images captured during construction often contain more objects, such as temporary structure, equipment, and human, that are usually not modeled in a BIM (Figure 4).

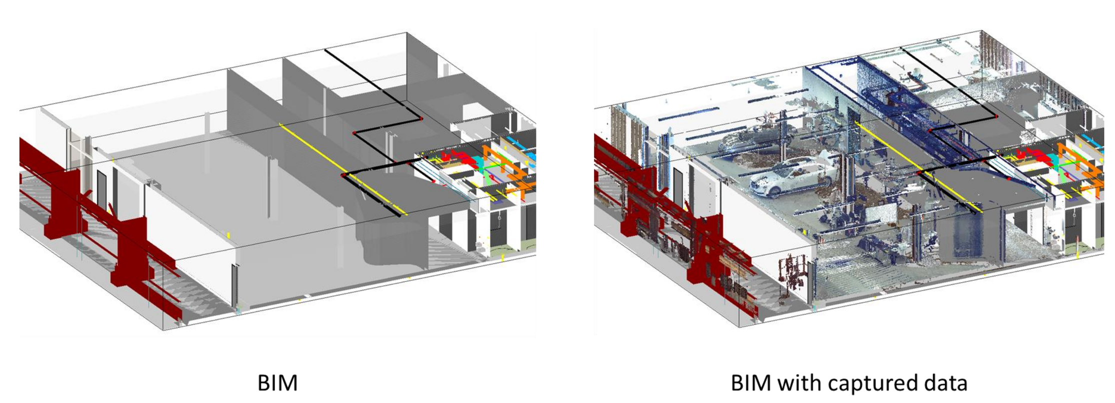


Figure 4. Scope differences

- **Uncertainty from Sparsity.** Unstructured data does not guarantee coverage in space and time, resulting in highly uncertain observations.

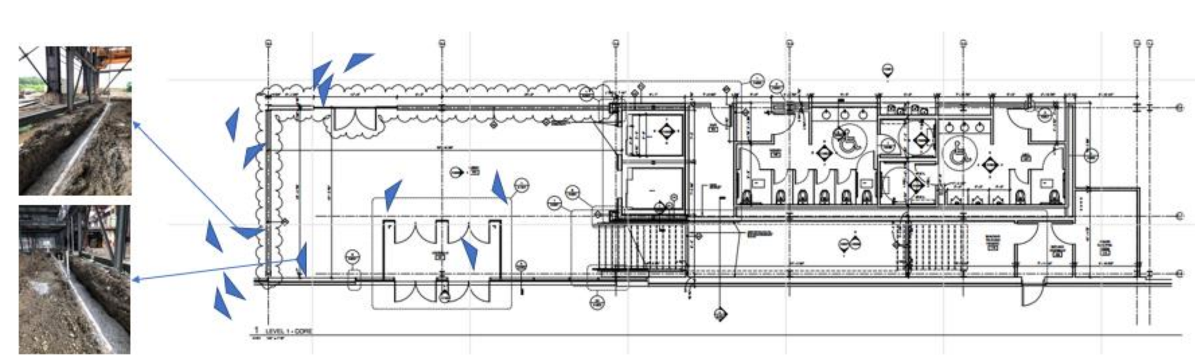
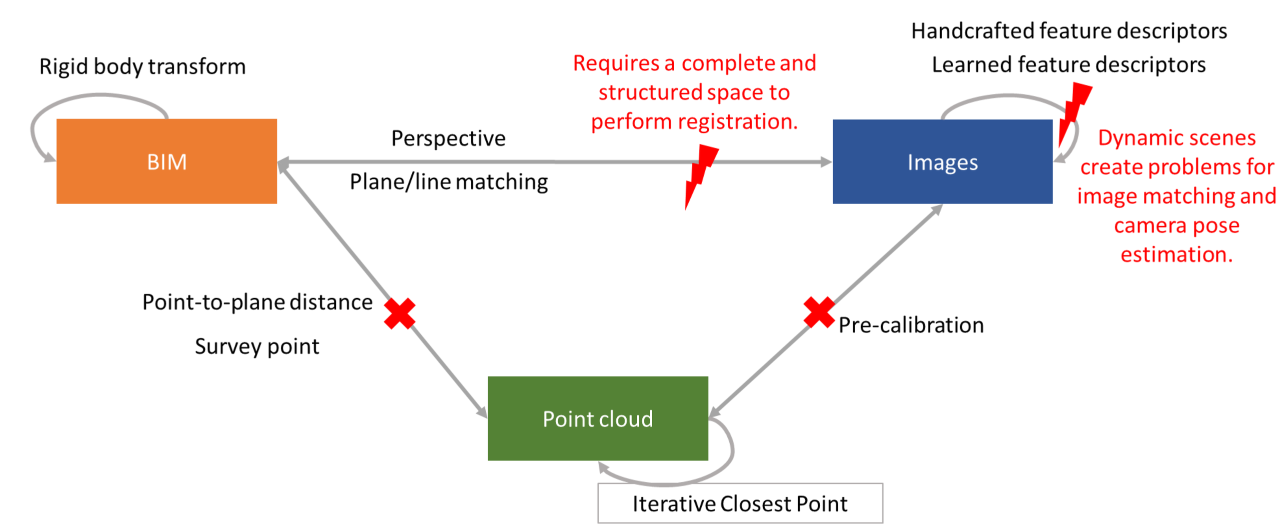


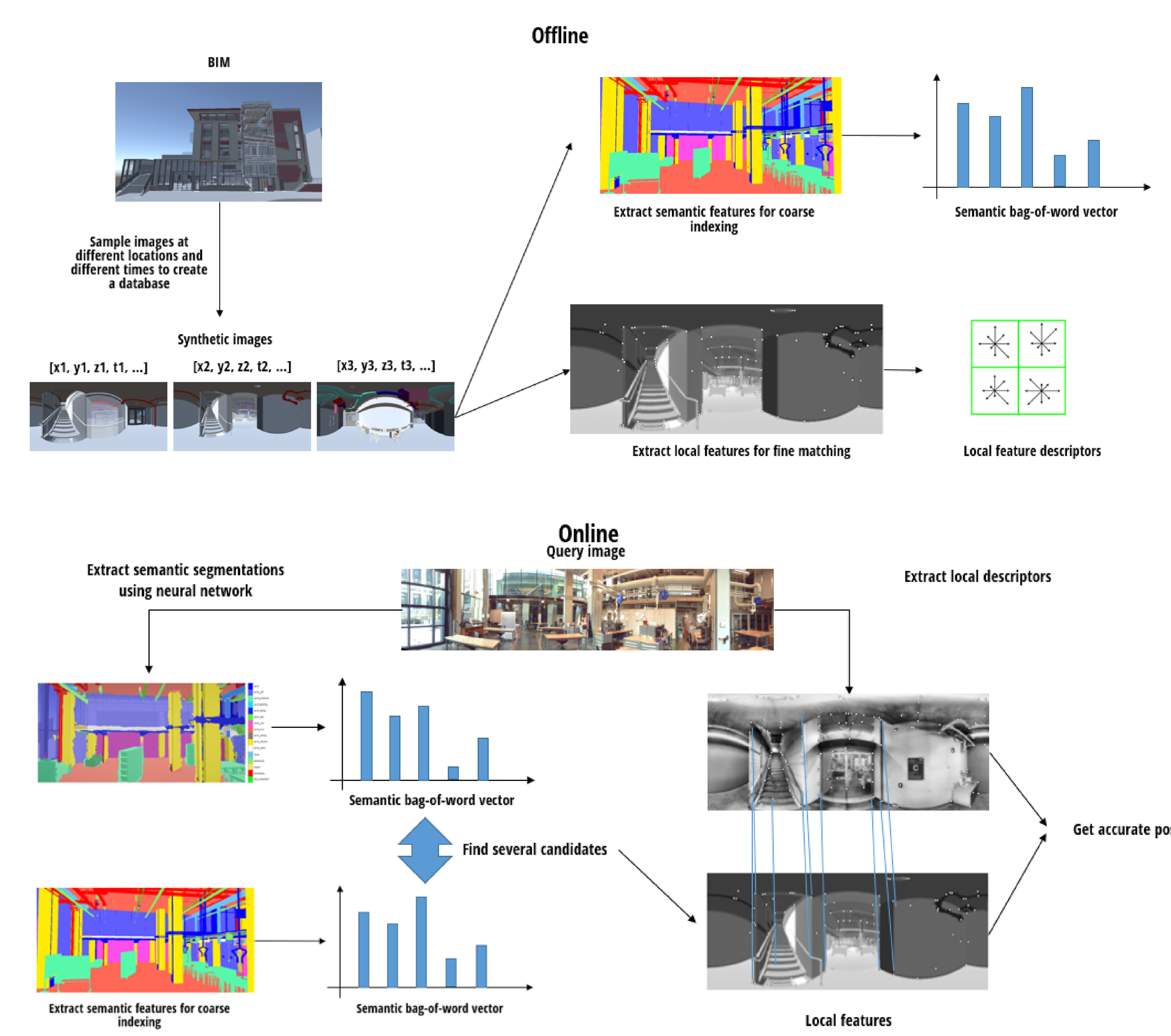
Figure 5. Distribution of unstructured images captured on 6/17/2018

Detour: Why not 3D reconstruction?

An intuition of solving the problem above is to leverage unstructured data captured during construction (such as images) to update as-designed models. Therefore, we proposed a method to register unstructured images to as-designed models using semantic features.



Workflow Overview



We can formally define the association problem as below: Let P be the camera projection matrix defined by a camera pose $P = K[R|t]$ where K is camera intrinsic, R is a rotation matrix, and t is a translation vector. The goal is to maximize the similarity between a synthetic view generated by putting a camera in a BIM and a real-world image.

Suppose $f(\cdot)$ is a function that returns some similarity measure given a pixel location, the problem can be described as:

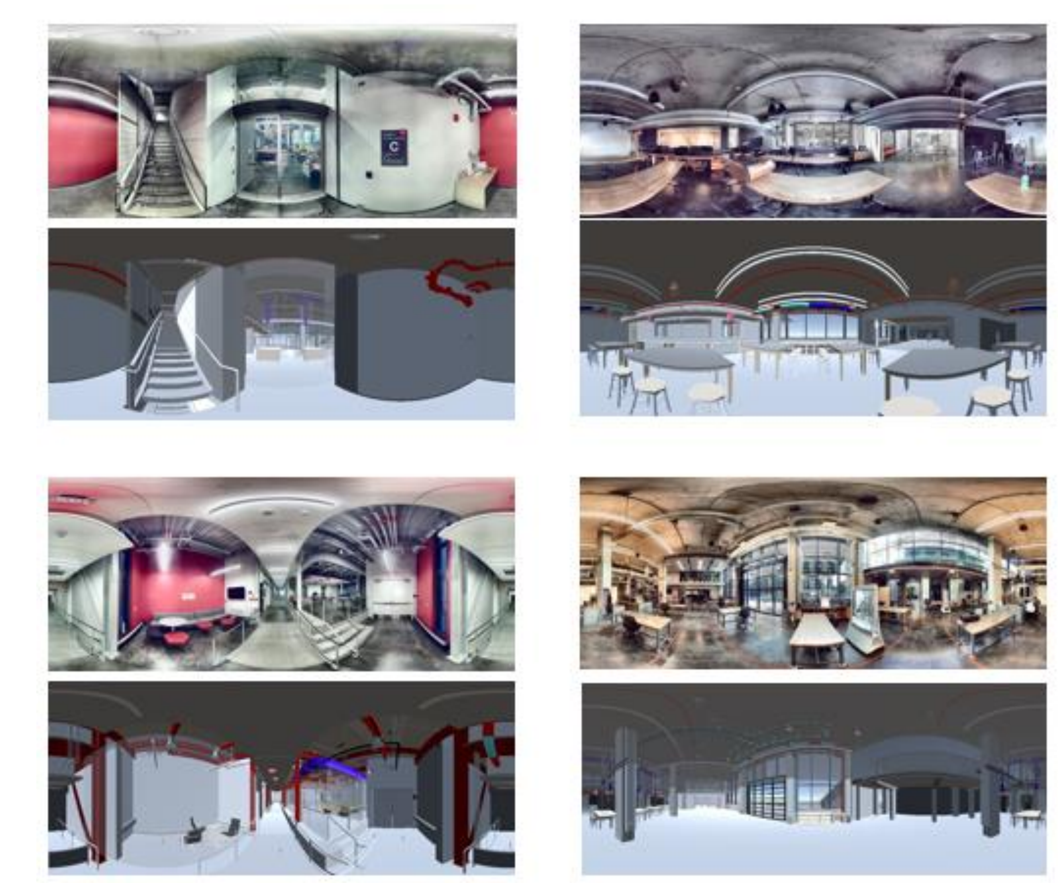
$$argmin_P \sum_{i=1}^{N} d(f(x_i), f((PX)_i))$$

In the optimization problem above, a key question is what $f(\cdot)$ should be used as a similarity measure. Here is a list of candidates:

- Visual local features
- Geometric features
- Semi-semantic features
- Semantic features

The proposed method leverages a hierarchical combination of semantic and visual features to support the registration.

Registration Results



Challenges and Next Step

- Structured learning for images with different scopes and intermediate states.
- Component-wise association.
- Represent uncertainties in a BIM.