Automated computer vision-based detection of components of under-construction indoor partitions

Hesam Hamledari *, Brenda McCabe, Shakiba Davari

Department of Civil Engineering, University of Toronto, Toronto, ON M5S 1A4, Canada

ARTICLE INFO

ABSTRACT

This paper presents a computer vision-based algorithm that automatically detects the components of an interior partition and infers its current state using 2D digital images. The algorithm relies on four integrated shape and color-based modules, which detect studs, insulation, electrical outlets, and three states for drywall sheets (installed, plastered, and painted). Based on the results of the four modules, images are classified into five states. The proposed method was validated using three image databases of indoor construction sites captured by a quadcopter (a type of unmanned aerial vehicle), a smartphone, and collected from publically available sources on the internet. The method's high accuracy rates, its fast performance, and applicability to different contexts such as automated robotic inspection are indicative of its promising performance. The visual detection results can potentially provide situational awareness for construction trades, provide future progress tracking systems with information on actual state, and help leverage the use of image processing at indoor sites.

1. Introduction

Constant monitoring of work at construction sites is crucial for reducing cost and schedule overruns and enhancing quality control, documentation, and communication [1,2]. When the work moves indoors, the need for situation awareness and monitoring increases because of the many trades involved and the impact of indoor construction on the overall project. Indoor partitions cover a significant portion of indoor construction and the delays associated with them can result in costly consequences and re-scheduling of the project [3]. The continuous data capture and tracking of work on partitions can benefit many trades including site managers, framers, insulation installers, electricians, drywall installers, plasterers, painters, and laborers.

To infer the current state of work for a partition, information is needed about the installation of its components such as studs, insulation blankets, electrical outlets, and drywall sheets. Superintendents and site managers need to manually inspect indoor sites to gather information about each object, which makes inspections labor intensive, expensive, time-consuming, and inefficient [4]. In recent years, there has been a great momentum toward utilization of digital images for automated data collection at construction sites using computer vision techniques. These visual resources are easy to capture, and they provide a less computationally intensive and more cost-effective alternative to other reality capture technologies [5]. The use of digital images offers a robust

means of detecting small changes in the appearance of partitions such as plastering and painting of drywall sheets that are not detectable using other tools.

While most studies have focused on the use of computer vision techniques at outdoor sites, the indoor applications face many challenges including changing viewpoints, highly cluttered scenes, occlusions, and diverse illumination conditions [2,6]. The achromatic characteristics of objects such as studs, electrical outlets, and their small and narrow sizes adds to this complexity. Furthermore, different states of work on a drywall sheet (installed, plastered, and painted) are very similar in terms of appearance, and they do not possess distinctive color and texture features. Hence, there is a need for solutions that can automate both the visual detection of components of interior partitions and the inference of a partition's current state. The visual detection of these components can help document and identify the current state of work, provide many trades with situation awareness, and potentially provide future vision-based progress tracking systems with information about actual state of work.

This paper proposes a computer vision-based method for automated detection of components of indoor under-construction partitions using 2D digital images. The method introduces four vision-based modules that automatically detect studs, insulation, electrical outlets, and different states of drywall sheets. The modules were also integrated to detect a partition's current state based on the recognition of its components. The method was validated using multiple image databases captured by a smart phone, an unmanned aerial vehicle (UAV), and collected from publically available sources on the internet, and at different illumination conditions. The recognition of both components and their

* Corresponding author.
  E-mail addresses: h.hamledari@mail.utoronto.ca (H. Hamledari),
brenda.mccabe@utoronto.ca (B. McCabe), sdavari@cs.toronto.edu (S. Davari).

current state are considered in the evaluation process. The performance of the method is also discussed with respect to image size and challenging indoor scenes. These analyses help ensure the applicability of the proposed methods to different contexts.

2. Background

In recent years, there has been a dramatic increase in the number of digital photos that are captured daily at a construction site [1]. The easy, economical, and fast access to these images initiated a stream of research on data collection and monitoring of construction sites using different forms of these media. Video, still, and time-lapse images enable project managers to monitor construction sites with less effort [7], improve the communication between stakeholders engaged in a project, and better document the progress [8]. Furthermore, development and introduction of computer vision techniques for automating the extraction of project-related information from digital images opened up many opportunities to leverage the use of visual resources. As a result, there is a large body of research conducted on the automated vision-based methods in the context of the construction industry. In general, image-based techniques can be categorized into 3D reconstruction-based solutions and image processing techniques.

Image-based 3D reconstruction solutions generate point clouds using thousands of overlapping images. These studies have addressed research problems such as generation of 3D and 4D as-built models [5, 9,10], and automated monitoring and visualization of construction progress [1,3,11–14]. One of the pioneering works was introduced as 4D augmented reality (D4AR) [4,15], where 3D as-planned models were superimposed over unordered daily photographs to visualize deviations from the schedule [4]. A structure from motion (SfM) technique was employed to automatically register camera viewpoints in an existing 3D model, which made significant strides toward automated progress monitoring. Even though the augmented reality-based methods have shown great promise, their application for indoor construction sites has not been adequately advanced. Due to their object-based approach, they cannot efficiently detect deviations in objects if they are not modeled in BIMs (e.g., operation-level works).

Several promising studies have been conducted [12,16] to solve some of these problems, including an appearance-based approach [16] in which materials were extracted from patches of 2D images, and the operation-level progress was detected using frequency diagrams of materials in the images. This method, although accurate and robust, has several drawbacks. It has a high computational cost, requires a comprehensive library of materials for training classifiers, and has not been specifically applied indoors where the material selection is significantly greater.

Image processing studies extract information directly from one to few digital images and have proven cost-effective and efficient for automated recognition and tracking of resources [17] such as workers [18–21] and equipment [18,21–24], classification of materials [12,25–27], productivity analysis [24,28,29], recognition of structural elements [30–32], and condition assessment [33].

While most solutions have been developed for outdoor use, vision-based techniques have started to be investigated for indoor sites [2]. The augmented reality-based method first presented as D4AR [4] was tailored for indoor progress monitoring by introducing an object-based 3D walk-through model [14]. This work improved the visualization of progress for indoor construction sites, provided construction managers with a realistic view of progress, and was the first step toward the application of augmented reality-based methods for indoors. This tool required the user to manually enter spatial-temporal information including the time, location and viewpoint for each photograph.

The indoor application of image processing techniques has also been studied for detection of project-related objects [30,34–37], materials [25–27], and state of progress [3,13,14]. One algorithm automatically detects the concrete columns in images using an integrated shape and color-based approach [30]. Others investigated the recognition of objects such as bricks [34], windows [35,37], and doors [36] in images. To automate quality inspections at indoor sites, a vision-based algorithm was developed to analyze tile alignment [38]. Image processing has also proven effective for detection of defects in tiling work [39]. As part of a pioneering study toward image-based progress tracking of indoor finishing work [3], a cascade algorithm was developed to identify the state of work for drywall construction [13]. This study clearly demonstrated the promising use of image processing for indoor progress tracking. However, the algorithm lacked consideration of other states of work, requires setting of threshold parameters for each state, and it classified images that were known to be in the finishing stage.

There are numerous challenges associated with the application of vision-based methods to indoor situations such as frequent changes in viewpoint, occlusion, highly cluttered scenes, extreme lighting conditions, and illumination patterns [6,13]. On the other hand, fixed cameras cannot be used indoors due to the obstruction of line of sight as walls are erected [8], so the input images needed for the operation of computer vision-based methods at indoor sites are typically captured manually and at different viewpoints at indoor sites, posing a challenge for visual detection methods.

Automated data collection at construction sites have also been studied using radio-based technologies and laser scanning. Tracking objects using radio frequency identification (RFID) technology [40–43] requires the installation of tags on objects, which will later be scanned using a reader that is located within a certain range. The use of UWB provides a wider range of coverage and has recently been studied for indoor use [44,45]. These radio-based technologies have shown promise for tracking materials and components in the dynamic environment of construction sites [46,47], and they have been recently integrated with virtual models [48,49]. The radio-based solutions, however, suffer from labor-intensive installation, scanning, and maintenance, and they cannot efficiently indicate progress of partially completed or operation-level tasks [42].

Laser scanning [10,40,50–53] involves merging 3D point clouds generated by the scanners into one as-built model. That model is then aligned to and compared with as-planned BIMs or CAD models to detect deviations. The integrated application of laser scanning and image-based 3D reconstruction techniques has also been the subject of some studies [10,54]. Laser scanners have proven to offer more accurate point clouds compared to the ones generated using image-based solutions [55]. The limitation of laser scanning-based methods include its lower accuracy and data loss at spatial discontinuities [42] caused by the mixed pixel phenomenon. In addition, laser scanners are challenged to generate accurate point clouds for reflective materials [4,42], and they are currently unable to provide semantic information for 3D models [4].

3. Methods

The system presented herein (Fig. 1) aims to automatically detect a partition's current state using 2D images. The system relies on four vision-based modules, designed to detect the components of interior partitions: steel studs, batt insulation, electrical outlets, and three states for drywall sheets. These components are associated with the primary stages of work on interior partitions, and their visual recognition in images can provide essential information for progress tracking systems and benefit various trades such as site managers, framers, electricians, insulation installers, drywall installers, plasterers, painters, and laborers.

The input images are passed through the system without a priori information as to the existing objects in the scenes. The components are visually detected, and the image is classified into one of the five possible states corresponding to framing (A), insulation (B), installed drywall (C), plastered drywall (D), and painted partition (E).