Geometry- and Appearance-based Reasoning of Construction Progress Monitoring

Kevin Han\textsuperscript{1}, Joseph Degol\textsuperscript{2}, and Mani Golparvar-Fard\textsuperscript{3}

\textsuperscript{1}Assistant Professor, Dept. of Civil, Const., and Envi. Engineer., North Carolina State Univ., Campus Box 7908, 2501 Stinson Dr., Raleigh, NC Email: kevin_han@ncsu.edu
\textsuperscript{2}PhD Candidate, Dept. of Computer Science, Univ. of Illinois, Urbana-Champaign, Urbana, IL
\textsuperscript{3}Associate Professor, Dept. of Civil and Envi. Engineer. and Dept. of Computer Science, Univ. of Illinois, Urbana-Champaign, Urbana, IL

ABSTRACT

Although adherence to project schedules and budgets is most highly valued by project owners, more than 53\% of typical construction projects are behind schedule and more than 66\% suffer from cost overruns, partly due to inability to accurately capture construction progress. To address these challenges, this paper presents new geometry- and appearance-based reasoning methods for detecting construction progress, which has the potential to provide more frequent progress measures using visual data that are already being collected by general contractors. The initial step of geometry-based filtering detects the state of construction of Building Information Modeling (BIM) elements (e.g. in-progress, completed). The next step of appearance-based reasoning captures operation-level activities by recognizing different material types. Two methods have been investigated for the latter step: a texture-based reasoning for image-based 3D point clouds and color-based reasoning for laser scanned point clouds. This paper presents two case studies for each reasoning approach for validating the proposed methods. The results demonstrate the effectiveness and practical significances of the proposed methods.

Keywords: Progress monitoring, BIM, images, point cloud, laser scan, 3D reconstruction, material classification
INTRODUCTION

Adherence to project schedules and budgets is the most highly valued performance metric by project owners (Bevan and Steve 2016). Despite its significance, more than 53% of typical construction projects are behind schedule and more than 66% do not meet their budget requirements (Bevan and Steve 2016). Some of the major factors that lead to poor performance on jobsites include 1) inconsistency among contractors, subcontractors and owners in terms of how much a construction project is faring at any given date, 2) flawed performance management due to lack of frequent reporting of actual performance to project teams, and 3) planners’ missed connections to most up-to-date construction progress information (Changal et al. 2015).

Moreover, construction sites are dynamic environments filled with a wide range of dynamic objects (e.g., workers and equipment). In addition, project management teams have to deal with multiple parties (i.e., owners, themselves, and many trades) constantly updating construction documents and schedules. These challenges are likely the cause of stagnant construction productivity compared to other industries (e.g. manufacturing almost doubled productivity over the past decade (Changal et al. 2015)).

The following subsection describes gaps-in-knowledge in project controls studied by the research community.

Practical and Theoretical Gaps in Project Controls

Over the past decade, the Last Planner System (LPS) (Kim and Ballard 2010) has emerged as a production control theory that can reduce waste during execution of planning through better coordination. However, the recent observations from a large number of construction projects with LPS have revealed that sustaining commitment to the goals of LPS for a long period of time is difficult (Ballard and Tommelein 2012). The recent case studies in (O’Brien et al. 2008) also show that many companies still emphasize control related to global project aims and fulfillment of contracts rather than production control.
Wide gap between long-term & short-term planning (lookahead vs. weekly work planning):

Implementing LPS improves the reliability of short-term planning; however, without prioritizing tasks based on the downstream demand, LPS can not effectively achieve a continuous flow of information (Hamzeh and Bergstrom 2010; Sacks et al. 2013). Addressing this gap in performance requires having continuous feedback on the most updated state of the tasks and various ongoing work packages on the site. The main issue is the cycle time for receiving feedback, which is typically the cycle of weekly work planning sessions (i.e. one or two weeks). This time period is too long to avoid waste, especially for tasks where their constraints are removed only a few days prior to their execution (Sacks et al. 2010b). Managing and responding to the high-level of details in production plans are needed daily, if not hourly (Dave et al. 2014; Brodetskaia et al. 2013).

Limited means of collecting, analyzing, and communicating status information:

The weekly work plans (WWP) do not have prior provisions for systematic status assessment (Sacks et al. 2010b), which according to the Construction Industry Institute (CII), National Research Council (NRC), National Institute of Science and Technology (NIST), and American Society of Civil Engineers (ASCE), is a key component to continuous improvement (CII 2010; NRC 2009; NIST 2011; Li et al. 2011).

Lack of effective methods for collecting status of the work advancement

Situation awareness is key to prompt and effective onsite decision making. To achieve an enhanced awareness on the status of the work advancements, the limitations in methods need to be addressed. In today’s best practices, most queries of the as-built conditions are done by traveling between a site and trailers to access paper-based documents (Kamat and Akula 2011; Bae et al. 2012), or by searching through smartphones or tablets which requires specific three-dimensional (3D) plan views to be manually generated for each task (Chen and Kamara 2011; Bae et al. 2012). This process is time-consuming given thousands of elements on a site (Sacks et al. 2013; Eastman et al. 2011). Analyzing performance based on experience is also often prone to errors (Turkan et al. 2013; Golparvar-Fard et al. 2013; Bosché 2012).
Lack of Support for Bringing “Power to the Edge” on Jobsites.

Since LPS plans are updated weekly, it is difficult to know “who is working on what task in what location” on a daily or hourly basis. There is also a major time lag between encountering an issue on site and when supervisors are informed (Garcia-Lopez and Fischer 2014). Thus, supervisors typically make decisions based on outdated information (Garcia-Lopez and Fischer 2014; Sacks et al. 2013). The inability to have two-way communication on task scope, methods, and resources also delays approval processes and leads to waste. Bringing Power to the Edge (Alberts and Hayes 2005)—empowering the individuals who actually do the work—requires enhanced communication and removal of constraints for quick and effective onsite decision making. While commercial mobile apps (i.e., PlanGrid and Autodesk360) are powerful in decentralizing work tracking and shortening time for accessing information, there are still gaps in knowledge on how site feedback can be captured and integrated with a Building Information Modeling (BIM)-based tracking system on a daily basis (Dave et al. 2014).

Lack of Methods for Intuitive Visualization of Project Information.

Despite the benefits of face-to-face discussions in toolbox (daily huddle) meetings, anecdotal observations from the recent implementation of LPS show that Last Planners successfully receive information on success and failure of their tasks only about 73% of the time during the performance review meetings and 60% of workers are not informed about their status (Salem et al. 2005). Moreover, Salem et al. (2005) report inconsistencies in remembering issues that are discussed during these toolbox meetings. For instance, 42% to 100% of planners remembered issues from these meetings, while the range was 17% to 86% for workers. Although there is growing recognition among researchers that visual analytics and visualization tools can improve communication rates in and out of the meetings, little is done on formalizing, developing, and validating BIM-based methods to benchmark, analyze, and communicate work status and other relevant project information in near real-time to both on- and offsite users.

OBJECTIVES AND CONTRIBUTIONS

To address the abovementioned gaps in knowledge, researchers have worked on developing
frameworks and tools that enable frequent data collection and progress deviation analysis (detailed in the Background section). These works aim at achieving a continuous flow of project information by analyzing visual data, which will enable a smooth flow of production.

Figure 1 illustrates how project control tools that visualize as-built and as-planned project information can achieve smooth flow of production. Leveraging these emerging sources of information can enable instantaneous project controls through automated and near real-time assessments of work-in-progress.

Achieving this goal can also support root-cause assessment on plan failures, facilitate information flows, and ultimately improve the reliability of weekly work planning. In particular, it can bridge the current knowledge gap and lead to creation of methods for 1) *project-level monitoring* (by providing a mechanism for better understanding how a project compares with others in terms of cost, schedule, and labor hours) and 2) *enhanced communication* (by providing real-time project information, improving onsite decision-making and work-sequencing, and fostering collaborative partnerships).

The proposed vision-based progress monitoring method will support project management teams by creating Integrated Project Models (IPM) as shown in Figure 1. The main contributions of the proposed method are 1) geometry- and appearance-based reasoning of progress detection and 2) two alternative approaches for image-based and point cloud-based (i.e., laser scanned) methods. An additional contribution is efficient processing (fast computation time) of large point clouds for detecting BIM elements.

**BACKGROUND**

**Related Work**

With an ever increasing number of visual data available on construction sites due to advances in computer vision and 3D imaging technologies, there have been dramatic advances in model-based construction progress detection leveraging as-built modeling techniques. Some of these techniques use image-based point clouds. Some other techniques use laser scanned point clouds.
**Image-based Point Clouds**

Siebert et al. (2014) used a camera-equipped Unmanned Aerial Vehicles (UAVs) to capture images of earthwork projects for creating 3D maps of the terrain. These surveyed point clouds can be used for measuring progress. Similarly, Golparvar-Fard et al. (2009, 2011) created point clouds from unordered sets of images. They aligned these point clouds with BIMs and compared geometries of as-built and as-planned models to reason about progress deviation. To deal with limited visibility and occlusions that were the challenges observed in these papers, Han and Golparvar-Fard (2015) proposed an appearance-based method that reasons about progress by recognizing textures of materials on construction images that were aligned with BIMs. The images were aligned with BIMs automatically after the image-based point clouds were aligned with BIMs.

**Laser Scanned Point Clouds**

Turkan et al. (2012, 2013) used surface-based recognition to detect building elements from scanned point clouds for automated progress detection and then improved the accuracy of progress tracking using the earned value analysis. Bosché et al. (2013) proposed a Scan-vs-BIM object recognition framework for tracking the built status of Mechanical, Electrical, and Plumbing (MEP) works. Similarly, Kim et al. (2013) compared 4D BIM with detected building elements from laser scanned point clouds to measure construction progress. These laser scanned methods are based on geometry recognition and generally provide more accurate and denser point clouds of the structures of interest than the image-based methods. However, the image-based methods provide multiple viewpoints and, therefore, wider viewpoints and occlusions.

**Point of Departure: Visual Analytics & Model-based Tracking Methods**

Over the past decade, several opportunities have emerged that can support work tracking:

1. The benefits of BIM (Young et al. 2009; Eastman et al. 2011) and its synergy with lean construction principles is well established (Sacks et al. 2010a; Sacks et al. 2013). BIM – augmented with production performance metrics– can serve as a great basis for representing as-planned performance and actual work deviations.
2. The number of images taken at construction sites to document work-in-progress has exponentially grown (Han and Golparvar-Fard 2015). It is now common to have at least a few hundred images taken on a jobsite on a daily basis. These images are either collected on the ground by construction personnel via consumer-grade cameras or by companies that offer professional photography services to construction projects; or, most recently, from above via camera-equipped UAV. The rapid advancement in camera, sensing, aeronautics and battery technologies have all contributed to UAVs becoming affordable, reliable, and easy to operate on construction sites. These camera-equipped UAVs can document work-in-progress by taking hundreds to thousands of overlapping images from various viewpoints in a short amount time (Ham et al. 2016; Han and Golparvar-Fard 2017).

3. The advancement in cloud computing and pervasiveness of smart devices on jobsites provides a great platform to connect onsite personnel to virtual models. A recent report (Constructech 2014) shows that 80% of U.S. contractors used commodity smartphones and tablets on their construction sites in 2014. Such platforms can be used for facilitating push and pull of information from an integrated information model.

4. The latest empirical observations from more than 100 projects (Alarcón et al. 2008) reveal the probability to reach high Percentage-Plan-Complete (PPC) values in lean projects can be duplicated by using information and communication tools (39% probability to reach PPC of 80% compared to 21% for projects without information and communication tools).

Leveraging these emerging sources of information and communication tools creates a unique opportunity for developing new methods to facilitate the implementation of lean construction principles and lighten the extra “burdens” imposed on project participants on collecting, analyzing, and communicating project status.

In an attempt to leverage these emerging opportunities to address gaps-in-knowledge, Golparvar-Fard et al. (2009, 2011) and Han and Golparvar-Fard (2015) propose computer vision-based progress monitoring methods that leverage visual data and BIM. Golparvar-Fard et al. (2009, 2011) compare the physical presence of as-built models (point clouds) to as-planned models (BIM).
These studies reveal challenges with occlusions and limited visibility. To deal with these challenges, Han and Golparvar-Fard (2015) propose a method that reasons about construction progress based on detected appearances (material textures in images) of BIM elements. The proposed geometry- and appearance-based reasoning method combines the advantages of Golparvar-Fard et al. (2009, 2011) and Han and Golparvar-Fard (2015).

The appearance-based method of Han and Golparvar-Fard (2015) is designed for image-based as-built models. The method outputs images aligned with BIM and uses BIM to segment and extract image patches to be classified. Therefore, it is not suitable for laser scanned point clouds. Back-projection of 3D points to image planes and an algorithm that fills holes between points could be a possible solution for using laser scanned point clouds. However, these back-projected images may lose texture and may have different color ranges from the typical red, green and blue (RGB) images. These challenges need to be investigated before implementing the appearance-based method for point clouds without associated images, which is the focus of this paper.

One of the two proposed appearance-based reasoning methods in this paper attempts to address this issue with point clouds without images (e.g., laser scanned). This method is based on a simple statistical model with less computational complexity compared to that of the image-based method. This method is designed for immediate practical use and, therefore, designed for fast computation time. The other image-based method is built on Golparvar et al. (2009) and Han and Golparvar-Fard (2015) in an effort to bring advantages of geometry-based and appearance-based detection together for improved performance. The following section details these two approaches.

METHOD

Figure 2 presents a process model of the proposed method. Highlighted boxes in colors other than gray indicate new contributions that were built on the authors’ previous methods (Golparvar-Fard et al. 2009; Han and Golparvar-Fard 2015). The output of the process model visualized along with an IPM will support WWP and Coordination during construction (see Figure 1).

Generating 3D as-built models is the initial process. As shown in Figure 2, inputs to the proposed method are images taken by commodity cameras and/or 3D point clouds captured by laser
scanners. The image-based 3D reconstruction process consists of structure-from-motion (SfM) for sparse reconstruction (Wu 2016) and multi-view stereo (MVS) for dense reconstruction (Goesele et al. 2007). Images are inputs to this pipeline of SfM-to-MVS and camera poses (intrinsic and extrinsic camera parameters) and point clouds are outputs. On the other hand, 3D laser scanners typically used in construction sites (e.g., time-of-flight terrestrial laser scanners) are used to generate 3D point clouds with commercial available software. Corresponding features (e.g., corners) between point clouds and BIMs are manually picked and similarity transformations are applied (by solving least squares problems of absolute orientation (Horn 1987)), to register the as-built and as-planned models and create IPMs.

**Geometry-based Filtering**

After the preparation of an IPM, the next step is geometry-based filtering. This is a simple occupancy check that examines whether or not there are points occupied by BIM elements. Due to registration errors, a threshold $\theta_{reg}$ with varying values was tested. Minimum and maximum coordinates of the entire BIM ($\text{min}_B$ and $\text{max}_B$) were used as an initial filtering. This process helps to remove points that are not part of the structure of interest. This process also reduces the size of image-based point clouds substantially and reduces computation times for the subsequent steps. According to the authors’ experience, images taken by UAVs tend to capture background objects that are not the structure of interest and therefore their point clouds consist of many non-relevant points. Thus, image-based point clouds that are typically captured from within the construction site boundary benefit more from this step compared to laser scanned point clouds that are captured from within the building footprint.

Then, the element-level filtering by minimum and maximum coordinates of BIM elements ($\text{min}_{Bim_i}$ and $\text{max}_{Bim_i}$) is performed. During this process, space distribution of a point cloud within each BIM element boundary is computed for filtering out false positives (i.e., there are some points within the boundary but they are not part of any BIM elements). To maximize efficiency and minimize computation time, vectorized computation and minimal computational complexity are critical factors. Therefore, a simple normal distribution with a standard deviation ($\sigma_{pc_i}$) is implemented
(see line 8 in Figure 3). This approach, as shown in Figure 3, checks the minimum number of
points within each BIM element ($\theta_{nPc}$) and also checks densities to avoid false negatives.

Geometry-based filtering detects BIM elements. The next step, appearance-based reasoning,
classifies material classes of the detected BIM elements (third column in Figure 2). Two different
approaches are described in the following subsections: color-based reasoning for point clouds
without images (e.g., laser scanned) and texture-based reasoning for image-based point clouds
with images that are aligned with BIMs.

**Color-based Reasoning**

For point clouds without images, color ranges of BIM elements are compared against pre-
collected material patches. The color ranges are created based on a normal distribution. The
averages and standard deviations of the pre-collected material patches ($\text{avg}_{Mat,j}$ and $\sigma_{Mat,j}$) are used
as training data. The reasoning process is based on the average color value of the points within
each BIM element ($\text{avg}_{BIM,i}$) falling into the range of the chosen threshold $\theta_{mat}$ (see Figure 4). This
statistical model is based on training data (pre-collected material patches) and features (colors in
this case). The overall process is presented by Figure 4.

The filter processes are carefully structured by logical variables and operations to maximize
efficiency. As previously stated, one of the goals was to study and propose a possible practical
solution that can be implemented immediately. Therefore, maximizing efficiency and minimizing
computation time were very important factors unlike a more sophisticated machine learning
algorithm that is designed for image-based material recognition presented in the following section.

**Texture-based Reasoning**

For point clouds with images, a learning approach is used for material recognition. The initial
step is patch extraction using camera parameters and BIM as was done by Han and Golparvar-
Fard (2015)’s approach (see Figure 5). For each image $c$ that is used to create 3D point clouds,
$N$ image patches per BIM element $FACE^i_c$ are extracted. In this paper, $w_{BIM}$ is used to assign
more weight to the expected material type, taking advantage of using BIM as a priori knowledge
(line 11 in Figure 5). The next step of material classification follows a similar approach to that of
Cimpoi et al. (2015) and DeGol et al. (2016) for material classification of the patches. In particular, a combination of Fisher Vectors (Perronnin et al. 2010) and Convolutional Neural Network (CNN) (Krizhevsky et al. 2012) features are input to a Support Vector Machine (SVM) for learning.

Fisher Vectors are created by first extracting dense SIFT (Lowe 1999) features from each patch. In training, the dense scale-invariant feature transform (SIFT) features are reduced to a dimensionality of 80 by Principal Components Analysis (PCA) before being clustered into 256 modes with a Gaussian Mixture Model (GMM). The Fisher Vectors are then mean and covariance deviations from the GMM modes ($\ell^2$ normalized and sign square-rooted). Convolutional Neural Network features are created using the pre-trained VGG-M network of Krizhevsky et al. (2012). The features are extracted from the last convolutional layer of the network rather than the fully connected layers.

Classification is then performed using a one vs. all SVM scheme. This scheme has been shown to achieve exemplary results for 2D texture recognition (Hayman et al. 2004; Cimpoi et al. 2014; Cimpoi et al. 2015; Degol et al. 2016). A $\chi^2$ kernel is used with the SVM. The Fisher Vector and CNN features are normalized individually before being concatenated for learning.

**EXPERIMENT SETUP**

Two different types of as-built data were prepared to test texture- and color-based reasonings: image-based and laser scanned point clouds, respectively, from two construction projects. Image-based 3D reconstruction was used to create an as-built point cloud of a hotel project (denoted as HP). A laser scanner was used for the same purpose on a biomedical building project (denoted BP). Hypothetical WWP schedules and their corresponding BIMs were generated, and the goal was to simulate progress monitoring for the given weeks. Table 1 summarizes this data preparation for the image-based and laser scanned approaches.

**Global Filtering**

As can be seen in Table 1, there are millions of points associated with each point cloud. Processing these points can be time-consuming. Typically, each point consists of six numbers (X, Y, and Z coordinates and RGB values), excluding normal values (three numbers in X, Y, and Z...
directions) that the proposed method does not use. The first step (lines 1-3) in Figure 3 removes all background points that can be significant. Figure 6 and Table 2 show how an image-based point cloud can have a large percentage of unwanted points (background).

The main cause, in the case of HP, was the use of a UAV for data capture. Due to safety concerns related to cranes, a UAV operator had to fly the UAV at high altitudes. Thus, many images had background buildings and roads that surround HP. The initial filtering process removed these objects. On the other hand, the laser scanned point cloud had a much smaller percentage reduction because the laser scanner was stationed within the building footprint. It had limited viewpoints compared to that of the UAV.

Element-level Filtering

The next step is filtering by each BIM element (lines 4-19 in Figure 3). The texture- and color-based reasoning happen within this step. This step performs the same filtering based on logical operations but this time at element-level. In other words, points per element are extracted. As mentioned in Section 4, varying values of threshold $\theta_{\text{reg}}$ were tested. Figure 7 shows two examples of points per element - a large concrete slab and column. The elements that are detected by these two filterings (one on the entire model and the other on each element) are input to the reasoning methods. This reduces computation times on “non-existing” elements during the reasoning process.

Figure 8 shows the effect of $\theta_{\text{reg}}$ on the second filtering process. Due to registration errors and the real structure (e.g., formwork in Figure 8b) having larger volume/area than the BIM elements, filtering purely based on the size of the BIM elements may filter out elements of interest (e.g., partly missing formwork in Figure 8a). Therefore, $\theta_{\text{reg}}$ needs to account for these cases and should be greater than 0. As seen in Figure 8a & b (HP) and 8c & d (BP), varying $\theta_{\text{reg}}$ has a significant impact on some of the elements - i.e., formwork of the core walls for HP and steel girders on the second floor of BP.
Training Data for Color- and Texture-based Reasoning

For color-based reasoning, patches of surfaces of different materials were extracted. Some of the extracted patches are presented in Figure 9. These patches are used as a training set that sets a statistical boundary for classifying material types.

For texture-based reasoning, the Construction Material Library (CML) initially collected by Dimitrov, Han, and Golparvar-Fard (2014, 2015) was used as the training dataset. CML consists of more than 3,000 images that are categorized into 20 construction material classes.

RESULTS

This section provides detailed analyses of all thresholds/factors discussed in the Method section.

Geometry- and Color-based Reasoning: BP

The first study was conducted by varying values of $\theta_{\text{reg}}$ (see Figure 6 and 7 for the effect of varying $\theta_{\text{reg}}$). As summarized by Figure 10, increasing $\theta_{\text{reg}}$ did not enhance the accuracy. Instead, it increased the number of false positives. This is due to the accurate registration between the laser scanned point cloud and BIM. By increasing $\theta_{\text{reg}}$, the chance of including points from unwanted objects increases (see Figure 11). $\theta_{nPc}$ of [1000:1,000,000] were tested and yielded the same results except when $\theta_{nPc}$ equals 1000. Since a large $\theta_{nPc}$ can cause false positives, a value that is lower than 1,000,000 but larger than 1000 ($\theta_{nPc} = 10,000$) was selected for the following studies.

Similarly, varying values of $\theta_{\text{space}}$ affects detection of BIM elements. The accuracies shown in Figure 12 refer to the accuracies of BIM element detection. As can be seen from Figure 12, increasing $\theta_{\text{space}}$ (wider distribution) increases the accuracies until a certain point ($\theta_{\text{space}} = 32$ in this case). $\theta_{\text{space}} = 8$ and $\theta_{\text{space}} = 16$ yielded the highest accuracy of 91.9%.

This is expected performance for the wider distribution of points because a small cluster of points will not be counted as an inlier (see Figure 11 for a case where a small cluster of points create a densely populated point in a part of a BIM element but not distributed throughout this BIM element). Not detecting these elements is a key indicator that this method is robust. Therefore,
counting numbers of true negatives (counting these elements as existing elements) is important and captured.

These results show that this geometry-based reasoning is very effective in detecting over 90% of BIM elements and can be used for progress monitoring as it is robust to true negatives (not detecting elements when it should not). Lastly, varying values of $\theta_{\text{mat}}$ effects the performance of material classification. Figure 13 shows the accuracies of material classification at operation-level. As can be seen in Figure 13, the highest accuracy is 67.57%. The main challenges are changes in lighting conditions and colors saved from the laser scanners. For instance, 61 scans were captured with some taken in the morning and some taken in the evening. As can be seen in Figure 9, the first four from the left are all concrete surfaces and data captured in the evening shows higher blue color values (B of RGB).

**Geometry- and Texture-based Reasoning: HP**

The main goal of this section is to validate the applicability of integrating geometry-based reasoning with the texture-based reasoning. Figure 14 presents the impact of varying values of $\theta_{\text{reg}}$ on detection of BIM elements. All cases except when $\theta_{\text{reg}}$ equals to zero successfully detected all BIM elements. However, there were more false positives as $\theta_{\text{reg}}$ is increased because of non-relevant points being captured by larger boundaries of BIM elements.

HP has a larger misalignment compared to that of BP (see Figure 15). For this reason, increasing $\theta_{\text{reg}}$ improves the accuracy of BIM element detection at the expense of increasing the number of false positives, similar to BP (see Figure 14). Moreover, the types of formwork used at this site are not included in the current CML. Therefore, there was one element that was not classified correctly due to having the number of patches capturing blue meshes outnumbering the number of patches capturing wooden formwork (see Figure 16). As can be seen in Figure 16, the formwork has a large area of blue meshes that are not part of CML. HP is a vertical construction project with many occlusions resulting in many patches that do not capture textures of interest (i.e., construction materials). For this reason, the effect of varying $w_{\text{BIM}}$ was significant. As seen in Figure 17, the accuracy increased from 65% to 91% when $w_{\text{BIM}}$ was increased from 1 to 2.5 (e.g., the number
of concrete patches for a concrete element is multiplied by $w_{BIM}$.

**Computation Time**

One of the main contributions of this paper is the efficient processing of point clouds (for geometry-based filtering and color-based reasoning). A computer with a 3.60 GHz CPU and 64 GB of RAM is used. Table 3 summarizes computation times for the proposed method. For processing more than a hundred million points, the proposed method takes around 30 seconds including both filtering and reasoning steps. This result and the accuracy of BIM element detection shows effective performance and possible use in practice (e.g., quick identification of existence). The processing time for the image-based approach (i.e., 3D reconstruction and texture-based reasoning) is not investigated - 3D reconstruction, training CML, and classification took hours to run with a powerful Graphics Processing Unit (GPU)-enabled server.

**Discussion on the Collected Data**

One of the objectives of this paper is to show that the proposed methods can utilize visual data that many construction companies already have. Thus, to test texture- and color-based reasoning, two construction projects with aerial images and laser scans that were already collected as part of their project control practices were chosen. Comparison of the two methods on the same dataset was not part of the scope. However, if they were tested on the same dataset, they would yield comparable results to the presented results - the texture-based method having higher accuracies than those of the color-based method. This is because the features used in both methods are specific to material types (e.g., color and texture of concrete) rather than being specific to projects.

In this paper, hypothetical WWPs and corresponding 4D BIMs were prepared to present as-planned conditions. However, in actual practice, WWPs and 4D BIMs may not be up-to-date (e.g., a quick change order that was not reflected in a WWP and a BIM). In this case, the as-built condition (e.g., dimension and location of a wall) may be significantly different from the as-planned condition. The proposed geometry-based filtering process would not be able to capture the progress correctly in this case. However, it can still "signal" the project management team and draw their attention to where the discrepancy is happening. They can either update the BIM or overwrite...
the progress. When a discrepancy is small (e.g., non-design problems: \(\text{discrepancy + registration error} < \theta_{\text{reg}}\)), the geometry filtering can still capture the as-built condition and compare that with the as-planned condition.

CONCLUSIONS AND FUTURE WORK

The proposed progress monitoring method has the following contributions: 1) combining geometry- and appearance-based reasoning methods and 2) providing an efficient and fast solution that can be used in practice. The geometry-based reasoning detects the existence of BIM elements without differentiating operation-level activities (e.g., formwork vs concrete). The appearance-based reasoning recognizes different material types and, therefore, can detect operation-level progress. Over 90% of the BIM elements in the two case studies were detected by the geometry-based detection. About 68% and 90% accuracies were achieved by color-based and texture-based reasoning, respectively.

The proposed method can be used with non-image-based datasets, such as laser scanned point clouds. However, enhancing training datasets (2D and 3D patches for image-based and point cloud based reasoning) and reducing computation time for texture-based reasoning need further investigation. The current datasets do not have enough samples for various construction materials.

Another remaining challenge is preparing proper model breakdown structures (MBS) and 4D BIMs. The companies that share HP and BP did not have proper MBSs and 4D BIMs. Given 3D BIMs and construction schedules, the authors created 4D BIMs of HP and BP. The authors also carefully inspected and removed any discrepancies found in the BIMs. This is one of the practical huddles that needs to be addressed to automate the proposed progress monitoring method that uses 4D BIMs.

DATA AVAILABILITY STATEMENT

Data generated or analyzed during the study is available from the corresponding author by request.
ACKNOWLEDGEMENTS

The authors would like to thank industry partners for providing access to their job sites and all undergraduate students who were involved in web development and data collection. This work is funded in part by the National Science Foundation (NSF) grant CMMI-1360562 and CMMI-1446765, the Department of Defense (DoD) National Defense Science and Engineering Graduate Fellowship (NDSEG), and the National Center for Supercomputing Applications (NCSA)’s Institute for Advanced Computing Applications and Technologies Fellows program. The authors gratefully acknowledge the support of NVIDIA Corporation with the donation of the Tesla K40 GPUs used for this research. Any opinions, findings, conclusions or recommendations presented in this paper are those of the authors and do not reflect the views of NSF, DoD, NCSA, NVIDIA, or the individual acknowledged above.
REFERENCES


CII, ed. (2010). IR252.2a – Guide to Activity Analysis. Construction Industry Institute, Austin, TX.
USA.


schedule and 3d sensing technologies.” *Automation in Construction*, 22(0), 414–421.


Table 1. Summary of data preparation

<table>
<thead>
<tr>
<th>Project</th>
<th># of Images</th>
<th># of Scans</th>
<th>Numb. of 3D points</th>
<th># of BIM Elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>HP</td>
<td>532</td>
<td>N/A</td>
<td>6,390,085</td>
<td>69</td>
</tr>
<tr>
<td>BP</td>
<td>N/A</td>
<td>61</td>
<td>148,622,647</td>
<td>40</td>
</tr>
</tbody>
</table>
Table 2. Point reduction from the initial geometry-based filtering

<table>
<thead>
<tr>
<th></th>
<th>HP</th>
<th>BP</th>
</tr>
</thead>
<tbody>
<tr>
<td># of points before initial filtering</td>
<td>6,390,085</td>
<td>148,622,647</td>
</tr>
<tr>
<td># of points after initial filtering</td>
<td>1,348,148</td>
<td>125,539,249</td>
</tr>
<tr>
<td>% reduction</td>
<td><strong>79.9</strong></td>
<td><strong>15.5</strong></td>
</tr>
</tbody>
</table>
Table 3. Computation time. * includes color-based reasoning

<table>
<thead>
<tr>
<th></th>
<th>HP</th>
<th>BP</th>
</tr>
</thead>
<tbody>
<tr>
<td># of points processed</td>
<td>6,390,085</td>
<td>148,622,647</td>
</tr>
<tr>
<td>Initial Filtering (sec)</td>
<td>0.22</td>
<td>4.75</td>
</tr>
<tr>
<td>Element-level Filtering (sec)</td>
<td>1.64</td>
<td>25.97 *</td>
</tr>
<tr>
<td><strong>Total (sec)</strong></td>
<td><strong>1.86</strong></td>
<td><strong>30.72</strong></td>
</tr>
</tbody>
</table>
Start

- Prepare Master Schedule
- Reverse Phase Scheduling
- Commit to Weekly Work Plan
- Look-Ahead Schedule
- Coordinate Weekly Work Plan
- Constraint Analysis

Integrated Project Model (Plan + As-built)

- Monitor Work Progress
- Progress Problems / Deviations?
  - Yes, Performance Problem
  - No, Confirm Task Complete

- Report Completed Task
- Yes, Performance Problem

- Inspect Task
- QA/QC Issues?
  - No, Confirm Task Complete
  - Yes, Pull Information

Lead and Conduct Work As-Planned

Master & Reverse Phase Scheduling
Weekly Work Planning and Coordination
Daily Work Execution, and Performance Monitoring and Reporting

End
Input: Coordinates of a BIM element $BIM_i \in BIM$ and a point cloud $pc$
Output: Filtered point cloud $pc_i$ and/or A material class per BIM element $c_i \in C$

1. $min_{BIM} = min_{BIM} - \theta_{reg}$;
2. $max_{BIM} = max_{BIM} + \theta_{reg}$;
3. Filter out points outside: $min_{BIM} < pc < max_{BIM}$;
4. foreach BIM element $BIM_i$ do
   5. $min_{BIM_i} = min_{BIM_i} - \theta_{reg}$;
   6. $max_{BIM_i} = max_{BIM_i} + \theta_{reg}$;
   7. Filter out points outside: $pc_i = min_{BIM_i} < pc < max_{BIM_i}$;
   8. if $max_{BIM_i} - min_{BIM_i} \sigma_{pc_i} < \theta_{space}$ and count $pc_i > \theta_{nPc}$ then
   9. if Laser scanned point cloud then
      10. Run color-based reasoning (Figure 4);
      11. Return $c_i$;
   else
      13. Run texture-based reasoning (Figure 5);
      14. Return $c_i$;
   end
  else
   16. $c_i = null$;
end
Input: Color information by material types: $Mat_j^c \in MAT^C$ and BIM elements $BIM_j^c \in BIM^C$; Material class $Mat_j \in MAT$

Output: Material class per BIM element: $c_i$

1. foreach $Mat_j \in MAT$ do
2.     if $avg_{Mat_j^c} + \sigma_{Mat_j^c} \theta_{mat} < avg_{BIM_j^c} < avg_{Mat_j^c} + \sigma_{Mat_j^c} \theta_{mat}$ then
3.         $c_i = Mat_j$;
4.     else
5.         $c_i = null$;
6.     end
7. end
**Input:** Back-projected faces $FACE_c^i$ of element $E^i$ for all images $c$ in $C$;
$N$: number of image patches per $FACE_c^i$;
$\delta$: size of the image patch;
$\eta$: max number of iterations; and
$w_{BIM}$: weight assigned to the expected material type

**Output:** Observed material $M^i$ for each element $E^i$

```plaintext
foreach element $E^i$ do
  $m_{expected} = \text{Parse 4D BIM and read the expected material type}$
  foreach image $c$ in $C$ do
    while $\text{iter} < \eta \text{ or } n < N$ do
      Randomly extract a sample patch within $FACE_c^i$
      if Succeed extracting a sample patch then
        Classify material and return the category ($m$) with max score
      end
      $\text{iter}++$
    end
  end
  $f_m^i = f_{m_{expected}} \times w_{BIM}$
  $M^i \leftarrow \arg\max_m f_m$: return material with max frequency of observation
end
```