

Geometry- and Appearance-based Reasoning of Construction Progress Monitoring

Kevin Han¹, Joseph Degol², and Mani Golparvar-Fard³

¹Assistant 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

²PhD Candidate, Dept. of Computer Science, Univ. of Illinois, Urbana-Champaign, Urbana, IL

³Associate 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

25 INTRODUCTION

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

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

40 The following subsection describes gaps-in-knowledge in project controls studied by the re-
41 search community.

42 **Practical and Theoretical Gaps in Project Controls**

43 Over the past decade, the Last Planner System (LPS) (Kim and Ballard 2010) has emerged
44 as a production control theory that can reduce waste during execution of planning through better
45 coordination. However, the recent observations from a large number of construction projects with
46 LPS have revealed that sustaining commitment to the goals of LPS for a long period of time is
47 difficult (Ballard and Tommelein 2012). The recent case studies in (O'Brien et al. 2008) also
48 show that many companies still emphasize control related to global project aims and fulfillment of
49 contracts rather than production control.

50 **Wide gap between long-term & short-term planning (lookahead vs. weekly work planning):**

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

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

61 The weekly work plans (WWP) do not have *prior* provisions for systematic status assessment
62 (Sacks et al. 2010b), which according to the Construction Industry Institute (CII), National Re-
63 search Council (NRC), National Institute of Science and Technology (NIST), and American Soci-
64 ety of Civil Engineers (ASCE), is a key component to continuous improvement (CII 2010; NRC
65 2009; NIST 2011; Li et al. 2011).

66 *Lack of effective methods for collecting status of the work advancement*

67 Situation awareness is key to prompt and effective onsite decision making. To achieve an
68 enhanced awareness on the status of the work advancements, the limitations in methods need to be
69 addressed. In today's best practices, most queries of the as-built conditions are done by traveling
70 between a site and trailers to access paper-based documents (Kamat and Akula 2011; Bae et al.
71 2012), or by searching through smartphones or tablets which requires specific three-dimensional
72 (3D) plan views to be manually generated for each task (Chen and Kamara 2011; Bae et al. 2012).
73 This process is time-consuming given thousands of elements on a site (Sacks et al. 2013; Eastman
74 et al. 2011). Analyzing performance based on experience is also often prone to errors (Turkan et al.
75 2013; Golparvar-Fard et al. 2013; Bosché 2012).

76 *Lack of Support for Bringing “Power to the Edge” on Jobsites.*

77 Since LPS plans are updated weekly, it is difficult to know “*who* is working on *what* task in *what*
78 location” on a daily or hourly basis. There is also a major time lag between encountering an issue
79 on site and when supervisors are informed (Garcia-Lopez and Fischer 2014). Thus, supervisors
80 typically make decisions based on outdated information (Garcia-Lopez and Fischer 2014; Sacks
81 et al. 2013). The inability to have two-way communication on task scope, methods, and resources
82 also delays approval processes and leads to waste. Bringing Power to the Edge (Alberts and Hayes
83 2005)—empowering the individuals who actually do the work—requires enhanced communication
84 and removal of constraints for quick and effective *onsite* decision making. While commercial
85 mobile apps (i.e., PlanGrid and Autodesk360) are powerful in decentralizing work tracking and
86 shortening time for accessing information, there are still gaps in knowledge on how site feedback
87 can be captured and integrated with a Building Information Modeling (BIM)-based tracking system
88 on a daily basis (Dave et al. 2014).

89 *Lack of Methods for Intuitive Visualization of Project Information.*

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

101 **OBJECTIVES AND CONTRIBUTIONS**

102 To address the abovementioned gaps in knowledge, researchers have worked on developing

103 frameworks and tools that enable frequent data collection and progress deviation analysis (detailed
104 in the Background section). These works aim at achieving a continuous flow of project information
105 by analyzing visual data, which will enable a smooth flow of production.

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

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

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

123 **BACKGROUND**

124 **Related Work**

125 With an ever increasing number of visual data available on construction sites due to advances in
126 computer vision and 3D imaging technologies, there have been dramatic advances in model-based
127 construction progress detection leveraging as-built modeling techniques. Some of these techniques
128 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.

155 2. The number of images taken at construction sites to document work-in-progress has expo-
156 nentially grown (Han and Golparvar-Fard 2015). It is now common to have at least a few hundred
157 images taken on a jobsite on a daily basis. These images are either collected on the ground by con-
158 struction personnel via consumer-grade cameras or by companies that offer professional photogra-
159 phy services to construction projects; or, most recently, from above via camera-equipped UAV. The
160 rapid advancement in camera, sensing, aeronautics and battery technologies have all contributed
161 to UAVs becoming affordable, reliable, and easy to operate on construction sites. These camera-
162 equipped UAVs can document work-in-progress by taking hundreds to thousands of overlapping
163 images from various viewpoints in a short amount time (Ham et al. 2016; Han and Golparvar-Fard
164 2017).

165 3. The advancement in cloud computing and pervasiveness of smart devices on jobsites pro-
166 vides a great platform to connect onsite personnel to virtual models. A recent report (Constructech
167 2014) shows that 80% of U.S. contractors used commodity smartphones and tablets on their con-
168 struction sites in 2014. Such platforms can be used for facilitating push and pull of information
169 from an integrated information model.

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

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

178 In an attempt to leverage these emerging opportunities to address gaps-in-knowledge, Golparvar-
179 Fard et al. (2009, 2011) and Han and Golparvar-Fard (2015) propose computer vision-based
180 progress monitoring methods that leverage visual data and BIM. Golparvar-Fard et al. (2009,
181 2011) compare the physical presence of as-built models (point clouds) to as-planned models (BIM).

182 These studies reveal challenges with occlusions and limited visibility. To deal with these chal-
183 lenges, Han and Golparvar-Fard (2015) propose a method that reasons about construction progress
184 based on detected appearances (material textures in images) of BIM elements. The proposed
185 geometry- and appearance-based reasoning method combines the advantages of Golparvar-Fard
186 et al. (2009, 2011) and Han and Golparvar-Fard (2015).

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

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

202 **METHOD**

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

207 Generating 3D as-built models is the initial process. As shown in Figure 2, inputs to the pro-
208 posed method are images taken by commodity cameras and/or 3D point clouds captured by laser

209 scanners. The image-based 3D reconstruction process consists of structure-from-motion (SfM) for
210 sparse reconstruction (Wu 2016) and multi-view stereo (MVS) for dense reconstruction (Goesele
211 et al. 2007). Images are inputs to this pipeline of SfM-to-MVS and camera poses (intrinsic and
212 extrinsic camera parameters) and point clouds are outputs. On the other hand, 3D laser scanners
213 typically used in construction sites (e.g., time-of-flight terrestrial laser scanners) are used to gen-
214 erate 3D point clouds with commercial available software. Corresponding features (e.g., corners)
215 between point clouds and BIMs are manually picked and similarity transformations are applied
216 (by solving least squares problems of absolute orientation (Horn 1987)), to register the as-built and
217 as-planned models and create IPMs.

218 **Geometry-based Filtering**

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

230 Then, the element-level filtering by minimum and maximum coordinates of BIM elements
231 (min_{BIM_i} and max_{BIM_i}) is performed. During this process, space distribution of a point cloud within
232 each BIM element boundary is computed for filtering out false positives (i.e., there are some points
233 within the boundary but they are not part of any BIM elements). To maximize efficiency and mini-
234 mize computation time, vectorized computation and minimal computational complexity are critical
235 factors. Therefore, a simple normal distribution with a standard deviation (σ_{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 (θ_{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 (avg_{Mat_j} and σ_{Mat_j}) are used as training data. The reasoning process is based on the average color value of the points within each BIM element (avg_{BIM_i}) falling into the range of the chosen threshold θ_{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 $FACE_c^i$ 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

263 Cimpoi et al. (2015) and DeGol et al. (2016) for material classification of the patches. In partic-
264 ular, a combination of Fisher Vectors (Perronnin et al. 2010) and Convolutional Neural Network
265 (CNN) (Krizhevsky et al. 2012) features are input to a Support Vector Machine (SVM) for learning.

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

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

278 **EXPERIMENT SETUP**

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

286 **Global Filtering**

287 As can be seen in Table 1, there are millions of points associated with each point cloud. Pro-
288 cessing these points can be time-consuming. Typically, each point consists of six numbers (X,
289 Y, and Z coordinates and RGB values), excluding normal values (three numbers in X, Y, and Z

290 directions) that the proposed method does not use. The first step (lines 1-3) in Figure 3 removes all
291 background points that can be significant. Figure 6 and Table 2 show how an image-based point
292 cloud can have a large percentage of unwanted points (background).

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

299 **Element-level Filtering**

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

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

315 **Training Data for Color- and Texture-based Reasoning**

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

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

322 **RESULTS**

323 This section provides detailed analyses of all thresholds/factors discussed in the Method sec-
324 tion.

325 **Geometry- and Color-based Reasoning: BP**

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

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

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

341 counting numbers of true negatives (counting these elements as existing elements) is important and
342 captured.

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

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

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

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

368 of concrete patches for a concrete element is multiplied by w_{BIM}).

369 **Computation Time**

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

379 **Discussion on the Collected Data**

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

388 In this paper, hypothetical WVPs and corresponding 4D BIMs were prepared to present as-
389 planned conditions. However, in actual practice, WVPs and 4D BIMs may not be up-to date
390 (e.g., a quick change order that was not reflected in a WVP and a BIM). In this case, the as-
391 built condition (e.g., dimension and location of a wall) may be significantly different from the as-
392 planned condition. The proposed geometry-based filtering process would not be able to capture the
393 progress correctly in this case. However, it can still "signal" the project management team and draw
394 their attention to where the discrepancy is happening. They can either update the BIM or overwrite

395 the progress. When a discrepancy is small (e.g., non-design problems: *discrepancy + registration*
396 *error* $< \theta_{reg}$), the geometry filtering can still capture the as-built condition and compare that with
397 the as-planned condition.

398 **CONCLUSIONS AND FUTURE WORK**

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

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

411 Another remaining challenge is preparing proper model breakdown structures (MBS) and 4D
412 BIMs. The companies that share HP and BP did not have proper MBSs and 4D BIMs. Given 3D
413 BIMs and construction schedules, the authors created 4D BIMs of HP and BP. The authors also
414 carefully inspected and removed any discrepancies found in the BIMs. This is one of the practical
415 huddles that needs to be addressed to automate the proposed progress monitoring method that uses
416 4D BIMs.

417 **DATA AVAILABILITY STATEMENT**

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

420 **ACKNOWLEDGEMENTS**

421 The authors would like to thank industry partners for providing access to their job sites and
422 all undergraduate students who were involved in web development and data collection. This work
423 is funded in part by the National Science Foundation (NSF) grant CMMI-1360562 and CMMI-
424 1446765, the Department of Defense (DoD) National Defense Science and Engineering Graduate
425 Fellowship (NDSEG), and the National Center for Supercomputing Applications (NCSA)'s In-
426 stitute for Advanced Computing Applications and Technologies Fellows program. The authors
427 gratefully acknowledge the support of NVIDIA Corporation with the donation of the Tesla K40
428 GPUs used for this research. Any opinions, findings, conclusions or recommendations presented
429 in this paper are those of the authors and do not reflect the views of NSF, DoD, NCSA, NVIDIA,
430 or the individual acknowledged above.

REFERENCES

- Alarcón, L. F., Diethelm, S., Rojo, O., and Calderón, R. (2008). "Assessing the impacts of implementing lean construction." *Revista ingeniería de construcción*, 23(1), 26–33.
- Alberts, D. S. and Hayes, R. E. (2005). *Power to the Edge eBook*. DoD Command and Control Research Program, <http://www.dodccrp.org/files/Alberts_Power.pdf>.
- Bae, H., Golparvar-Fard, M., and White, J. (2012). "Enhanced hd⁴ar (hybrid 4-dimensional augmented reality) for ubiquitous context-aware aec/fm applications." 253–262.
- Ballard, G. and Tommelein, I. (2012). "Lean management methods for complex projects." *Engineering Project Organization Journal*, 2(1-2), 85–96.
- Bevan, M. and Steve, J. (2016). "LCI National Webinar: Preliminary findings from national project performance research." Lean Construction Institute, <http://www.leanconstruction.org/media/img/safetyweek/FINAL-2016-05-26_LCI_National_Webinar-Owner_Project_Performance_Study_Overview_and_CallToAction.pdf>.
- Bosché, F. (2012). "Plane-based registration of construction laser scans with 3d/4d building models." *Advanced Engineering Informatics*, 26(1), 90–102.
- Bosché, F., Guillemet, A., Turkan, Y., Haas, C., and Haas, R. (2013). "Tracking the built status of mep works: Assessing the value of a scan-vs-bim system." *Journal of Computing in Civil Engineering*, (ja).
- Brodetskaia, I., Sacks, R., and Shapira, A. (2013). "Stabilizing production flow of interior and finishing works with reentrant flow in building construction." *Journal of Construction Engineering and Management*, 139(6), 665–674.
- Changal, S., Mohammad, A., and van Nieuwland, M. (2015). "The construction productivity imperative: How to build megaprojects better." *Insights & Publications, McKinsey&Company*.
- Chen, Y. and Kamara, J. (2011). "A framework for using mobile computing for information management on construction sites." *Elsevier Journal of Automation in Construction*, 20(7), 776 – 788.
- CII, ed. (2010). *IR252.2a – Guide to Activity Analysis*. Construction Industry Institute, Austin, TX,

458 USA.

459 Cimpoi, M., Maji, S., Kokkinos, I., Mohamed, S., and Vedaldi, A. (2014). “Describing textures in
460 the wild.” *Conf., Computer Vision and Pattern Recognition* (June).

461 Cimpoi, M., Maji, S., and Vedaldi, A. (2015). “Deep filter banks for texture recognition and seg-
462 mentation.” *Conf., Computer Vision and Pattern Recognition*.

463 Constructech (2014). “Resources/about magazines (June).

464 Dave, B., Kubler, S., Främling, K., and Koskela, L. (2014). “Addressing information flow in lean
465 production management and control in construction.” *Proceedings of The 22nd International
466 Group for Lean Construction (IGLC22)*, 2, 581–592.

467 Degol, J., Golparvar-Fard, M., and Hoiem, D. (2016). “Geometry-informed material recognition.”
468 *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*.

469 Dimitrov, A. and Golparvar-Fard, M. (2014). “Vision-based material recognition for automated
470 monitoring of construction progress and generating building information modeling from un-
471 ordered site image collections.” *Advanced Engineering Informatics*, 28(1), 37 – 49.

472 Eastman, C., Teicholz, P., Sacks, R., and Liston, K. (2011). *BIM handbook: A guide to building
473 information modeling for owners, managers, designers, engineers and contractors*. Wiley. com.

474 Garcia-Lopez, N. P. and Fischer, M. (2014). “A system to track work progress at construction
475 jobsites.” *Proceedings of the 2014 Industrial and Systems Engineering Research Conference*.

476 Goesele, M., Snavely, N., Curless, B., Hoppe, H., and Seitz, S. M. (2007). “Multi-view stereo for
477 community photo collections.” *2007 IEEE 11th International Conference on Computer Vision*,
478 1–8 (Oct).

479 Golparvar-Fard, M., Bohn, J., Teizer, J., Savarese, S., and Peña Mora, F. (2011). “Evaluation
480 of image-based modeling and laser scanning accuracy for emerging automated performance
481 monitoring techniques.” *Automation in Construction*, 20(8), 1143–1155.

482 Golparvar-Fard, M., Heydarian, A., and Niebles, J. C. (2013). “Vision-based action recognition of
483 earthmoving equipment using spatio-temporal features and support vector machine classifiers.”
484 *Advanced Engineering Informatics*, 27(4), 652–663.

485 Golparvar-Fard, M., Peña Mora, F., and Savarese, S. (2009). “Application of d4ar – a 4-dimensional
486 augmented reality model for automating construction progress monitoring data collection, pro-
487 cessing and communication.” *ITcon*, 14, 129–153.

488 Ham, Y., Han, K., Lin, J., and Golparvar-Fard, M. (2016). “Visual monitoring of civil infrastruc-
489 ture systems via camera-equipped unmanned aerial vehicles (uavs): a review of related works.”
490 *Visualization in Engineering*, 4(1), 1–8.

491 Hamzeh, F. and Bergstrom, E. (2010). “The lean transformation: A framework for successful
492 implementation of the last planner system in construction.” *International Proceedings of the*
493 *46th Annual Conference. Associated Schools of Construction*.

494 Han, K. and Golparvar-Fard, M. (2015). “Appearance-based material classification for monitor-
495 ing of operation-level construction progress using 4d bim and site photologs.” *Automation in*
496 *Construction*, 53(0), 44 – 57.

497 Han, K. and Golparvar-Fard, M. (2017). “Potential of big visual data and building information
498 modeling for construction performance analytics: An exploratory study.” *Automation in Con-*
499 *struction*, 73, 184 – 198.

500 Hayman, E., Caputo, B., Fritz, M., and Eklundh, J.-O. (2004). “On the significance of real-world
501 conditions for material classification.” *European Conference on Computer Vision*.

502 Horn, B. K. P. (1987). “Closed-form solution of absolute orientation using unit quaternions.” *Jour-*
503 *nal of the Optical Society of America A*, 4(4), 629–642.

504 Kamat, V. and Akula, M. (2011). “Integration of global positioning system and inertial naviga-
505 tion for ubiquitous context-aware engineering applications.” *Proc., National Science Foundation*
506 *Grantee Conference, Atlanta, GA*, 1–10.

507 Kim, C., Son, H., and Kim, C. (2013). “Automated construction progress measurement using a 4d
508 building information model and 3d data.” *Automation in Construction*, 31(0), 75 – 82.

509 Kim, Y. and Ballard, G. (2010). “Management thinking in the earned value method system and the
510 last planner system.” *Journal of Management in Engineering*, 26(4), 223–228.

511 Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). “Imagenet classification with deep convo-

512 lutional neural networks.” *Advances in Neural Information Processing Systems* 25.

513 Li, S., Becerik-Gerber, B., Jog, G., and Brilakis, I. (2011). “Civil and environmental engineering
514 challenges for data sensing and analysis.” *Computing in Civil Engineering (2011)*, 110–117.

515 Lowe, D. (1999). “Object recognition from local scale-invariant features.” *International Confer-
516 ence on Computer Vision*.

517 NIST (2011). “2011-2012 criteria for performance excellence.

518 NRC (2009). *Advancing the Competitiveness and Efficiency of the U.S. Construction Industry*.
519 Committee on Advancing the Competitiveness and Productivity of the U.S. Construction Indus-
520 try, National Research Council, National Academies Press.

521 O’Brien, W. J., Formoso, C. T., Ruben, V., and London, K. (2008). *Construction supply chain
522 management handbook*. CRC press.

523 Perronnin, F., Sánchez, J., and Mensink, T. (2010). “Improving the fisher kernel for large-scale
524 image classification.” *European Conference on Computer Vision*.

525 Sacks, R., Barak, R., Belaciano, B., Gurevich, U., and Pikas, E. (2013). “Kanbim workflow man-
526 agement system: Prototype implementation and field testing.” *Lean Construction Journal*, 9(1),
527 19–34.

528 Sacks, R., Koskela, L., Dave, B., and Owen, R. (2010a). “Interaction of lean and building informa-
529 tion modeling in construction.” *Journal of Construction Engineering and Management*, 136(9),
530 968–980.

531 Sacks, R., Radosavljevic, M., and Barak, R. (2010b). “Requirements for building information mod-
532 eling based lean production management systems for construction.” *Automation in Construction*,
533 19(5), 641 – 655 Building Information Modeling and Collaborative Working Environments.

534 Salem, O., Solomon, J., Genaidy, A., and Luegring, M. (2005). “Site implementation and assess-
535 ment of lean construction techniques.” *Lean Construction Journal*, 2(2), 1–21.

536 Siebert, S. and Teizer, J. (2014). “Mobile 3d mapping for surveying earthwork projects using an
537 unmanned aerial vehicle (uav) system.” *Automation in Construction*, 41(0), 1 – 14.

538 Turkan, Y., Bosché, F., Haas, C., and Haas, R. (2012). “Automated progress tracking using 4d

- 539 schedule and 3d sensing technologies.” *Automation in Construction*, 22(0), 414–421.
- 540 Turkan, Y., Bosché, F., Haas, C., and Haas, R. (2013). “Toward automated earned value tracking
541 using 3d imaging tools.” *Journal of Construction Engineering and Management*.
- 542 Wu, C. (2016). “Visualsfm: A visual structure from motion system.
- 543 Young, N., Jones, S., Bernstein, H. M., and Gudgel, J. (2009). “The business value of bim-getting
544 building information modeling to the bottom line.” *Bedford, MA: McGraw-Hill Construction*.

Table 1. Summary of data preparation

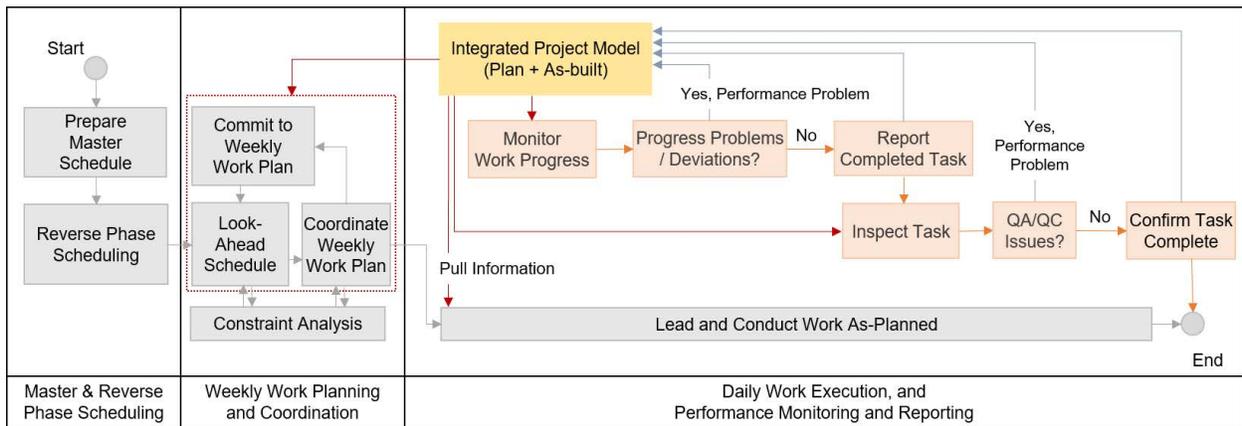
Project	# of Images	# of Scans	Numb. of 3D points	# of BIM Elements
HP	532	N/A	6,390,085	69
BP	N/A	61	148,622,647	40

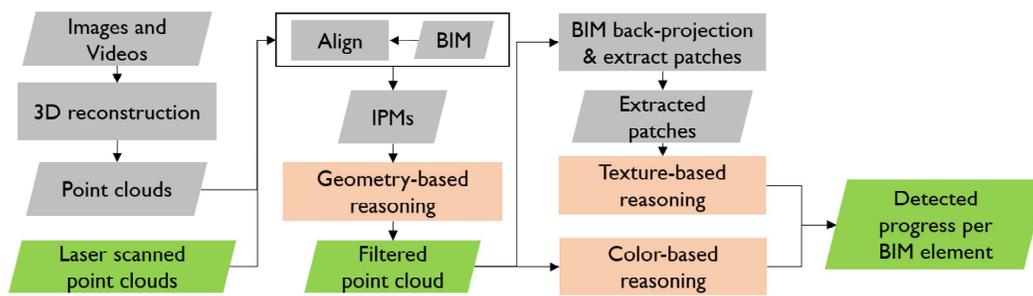
Table 2. Point reduction from the initial geometry-based filtering

	HP	BP
# of points before initial filtering	6,390,085	148,622,647
# of points after initial filtering	1,348,148	125,539,249
% reduction	79.9	15.5

Table 3. Computation time. * includes color-based reasoning

	HP	BP
# of points processed	6,390,085	148,622,647
Initial Filtering (sec)	0.22	4.75
Element-level Filtering (sec)	1.64	25.97 *
Total (sec)	1.86	30.72





Input: Coordinates of a BIM element $BIM_i \in \mathbb{BIM}$ and a point cloud pc

Output: Filtered point cloud pc_i and/or A material class per BIM element $c_i \in \mathbb{C}$

```
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  $c_i$ ;
12         else
13             Run texture-based reasoning (Figure 5);
14             Return  $c_i$ ;
15         end
16     else
17          $c_i = null;$ 
18     end
19 end
```

Input: Color information by material types: $Mat_j^C \in MAT^C$ and BIM elements $BIM_i^C \in BIM^C$; Material class $Mat_j \in MAT$
Output: Material class per BIM element: c_i

```
1 foreach  $Mat_j \in MAT$  do  
2   | if  $avg_{Mat_j^C} + \sigma_{Mat_j^C} \theta_{mat} < avg_{BIM_i^C} < avg_{Mat_j^C} + \sigma_{Mat_j^C} \theta_{mat}$  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 \mathbb{C} ;
 N : number of image patches per $FACE_c^i$;
 δ : size of the image patch;
 η : max number of iterations; and
 w_{BIM} : weight assigned to the expected material type

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

```

1 foreach element  $E^i$  do
2    $m_{expected}$  = Parse 4D BIM and read the expected material type foreach
   image  $c$  in  $\mathbb{C}$  do
3     while  $iter < \eta$  or  $n < N$  do
4       Randomly extract a sample patch within  $FACE_c^i$ 
5       if Succeed extracting a sample patch then
6         | Classify material and return the category ( $m$ ) with max score
7       end
8        $iter++$ 
9     end
10  end
11   $fm^i = fm_{expected} \times w_{BIM}$ 
12   $M^i \leftarrow \arg \max_m fm$ : return material with max frequency of observation
13 end

```

