IMPROVING CONSTRUCTION QUALITY BY COMBINING THREE-DIMENSIONAL MODELS AND VISUAL RECOGNITION SOFTWARE

William Croll, University of Southern California Viterbi School of Engineering

Abstract

The construction industry is primarily concerned with increasing project quality while simultaneously maintaining and increasing the efficiency with which that quality is attained. It is proposed that both of these objectives can be achieved using a combination of existing visual analysis and three-dimensional modeling tools to identify, in real time, inconsistencies between design plans and physical structures. There are multiple avenues for achieving this, each with advantages and drawbacks. While it is impossible to objectively comment on the effectiveness of such a system without real-world testing, the existing components used to assemble it provide a framework for a theoretically effective system. The potential for cost savings and increased efficiency make for a marketable system with prospectively wide industry appeal.

Introduction

One of the primary sources of disputes in modern construction projects is deviation from the project plans, whether real or perceived (Berger, 2011). This difficulty arises from the most widely used method of error-checking: on-site inspection by construction managers and supervisors (Hendrickson, 2008). This method is by its nature inefficient and potentially inaccurate. To remedy this, it is proposed that the use of visual recognition and analysis software be expanded in the construction industry to automatically identify inconsistencies between plans and physical structures in real time. The necessary technology is available and, if implemented properly, has the potential to increase both the quality and efficiency of work on construction projects of any scale.

Motivation

The current construction management process emphasizes planning in order to minimize error, but construction sites are complex dynamic environments that give rise to a myriad of difficulties in the execution of these plans (Picard, 2004). Furthermore, the potential impact of seemingly minor errors increases dramatically as projects become more complex and larger in scale. Ramifications of small deviations from established plans may not become apparent for days or months after they occur, making them extremely difficult to correct if they are not quickly identified (Berger, 2011). The current methods of quality assurance place the burden of catching these sometimes minute inaccuracies squarely on managers and supervisors. Taking into account these individuals’ additional responsibilities with regard to productivity monitoring and site organization, it becomes immediately apparent that the established system is not an optimal use of their time or energy.

This represents a significant problem when coupled with current industry trends. The construction sector is experiencing continuous increases in project complexity (Gidado, 1996). This progression only increases the need for a viable method of rapidly detecting deviations from established building plans. With each beam welded into place and every layer added to a structure, errors become costlier to correct, whether through design modifications or partial demolition and reconstruction. Rework due to unplanned changes can incur costs ranging from ten to fifteen percent of the total contract value (Senaratne & Sexton, 2008). It is with this in mind that numerous advances in project planning have been made, but these efforts are purely preventative and are primarily concerned with reducing the potential for error. Regardless of the degree of planning, that potential exists and necessitates an effective means of cross-referencing structures with the plans on which they are based.

Compounding the difficulties imposed by increased project complexity is the persistent compression of modern construction schedules. The consequential increase in the value of time is illuminating the time-intensive nature of current analysis techniques and increasing the demand for effortless data interpretation (Gong & Caldas, 2010). All of these factors have created an environment in the construction industry that is ripe for the increased use of visual recognition and analysis techniques.

Visual Recognition

Computerized visual recognition is a technology that has benefited greatly from the rapid increase of video quality and computing power. In principle, it is an automated method of identifying the content of an image, usually by decomposing it into basic elements and cross-referencing those with a database of known images (Berg, 2005). This process is done efficiently and intuitively by the human brain, but until recently programmers have struggled to translate it to computer code. As the...
technology has matured, visual recognition has been applied to a variety of industries, most notably security in the form of facial recognition databases.

To anyone with even a cursory knowledge of the construction industry, the potential advantages of such a technology are staggering. Of particular relevance is recent research that focuses on the matching of similar but non-identical shapes based on alignment and relationships rather than objective fixed-perspective geometry (Berg, 2005). Once perfected, this technique will allow for image recognition from multiple angles. This and other modifications can increase the utility of visual recognition technology to construction projects, and the industry is already warming to the concept.

Industry Status

Use of Visual Recognition

This rapidly maturing technology has been effective in some construction firms, with an initial focus placed largely on operations monitoring. In a 2010 study, researchers attempted to create a system for automatically translating videos of construction sites into productivity information (Gong & Caldas). The preliminary results showed a distinct potential for the integration of automatic visual analysis into the construction field for the purpose of productivity measurement. Expanding the utility of existing video recording equipment represents a effort to maximize resource usage, which is in keeping with the overarching philosophy of industrial engineering.

Other promising results include the use of visual recognition for tool tracking. Construction companies should maintain a balance between the costs of losing equipment and the cost of tracking it (Goedert, Foster, Jewell, & Bartek, 2009). The current industry standard for tracking both equipment and materials is a combination of Global Positioning Systems (GPS), Wireless Fidelity (WiFi), Radio Frequency Identification (RFID), and Bluetooth. However, these technologies require the attachment of some sort of tag or chip to the equipment and require significant expenditures of both time and funds to implement (Brilakis, Cordova, & Clark, 2008). Furthermore, these systems are not a cost-effective solution to equipment loss. Visual recognition software could potentially fulfill this function, as researchers have developed a program capable of identifying different pieces of equipment in a controlled environment with a success rate in excess of ninety percent (Arif et al., 2010). This and similar programs are able to detect the presence or absence of components, as well as their relative positions. These and other experiments demonstrate the impact of automated visual recognition on the construction industry.

Photogrammetry

One of the disadvantages of image analysis is that it attempts to define a three-dimensional system using data that is technically only two-dimensional. To remedy this, some have proposed the use of a process called photogrammetry. Essentially, photogrammetry is the digital process of calculating three-dimensional coordinates from two-dimensional images (Memon, Majid, & Mustaffar, 2006). Most current applications of photogrammetry deal exclusively with static images taken from large distances, usually aerial or even satellite photos, but research by the The Bureau of Taiwan High Speed Rail has shown the potential this technology carries when applied at closer range (Liao, 2000). Figure 1 demonstrates the level of detail attainable through multiple-angle photogrammetry.

![Figure 1: 3-D model constructed using two close-range images (Liao, 2000).](Image)

This is another early-stage technology with potential to change the industry standard. It is currently used for progress and productivity evaluation and is most often analyzed manually by a human technician or observer (Memon et al., 2006). Recent increases in computing power have this technique poised to join visual recognition in the category of automated systems, but photogrammetry and visual recognition have largely been put to use independently.

Building Information Modeling

A widely discussed topic of late has been Building Information Modeling (BIM). A key characteristic of most BIM systems is the extensive use of highly detailed three-dimensional renderings of construction projects, mostly in the planning phases (United States General Services Administration [USGSA], 2010). Translating two-dimensional specifications into three-dimensional models affords designers the opportunity to preemptively detect design flaws and interferences, among other advantages. BIM is rapidly becoming an industry standard, with over half of the U.S. construction industry employing it in some form (McGraw-Hill Construction, 2009). The bid-based
nature of the construction industry is essentially forcing this widespread adoption as firms try to remain competitive. The effort required to create a three-dimensional model is considerable, and failing to extract every valuable piece of information from such a model amounts to leaving money on the table.

**Proposed System**

The preceding information reveals three trends in the construction industry: increased use of computer models, necessity for automation, and a desire for increases in both quality and efficiency. To that end, it is proposed that visual recognition software be combined with BIM models and photogrammetry programs to identify inconsistencies between construction plans and physical buildings in real time. As shown in Figure 2, the photogrammetry process generates the necessary data for interface with three-dimensional BIM models.

![Figure 2: Steps of photogrammetry (Memon et al., 2006)](image)

These steps are currently performed entirely by technicians. In order to automate the process, the target locations and points need to instead be generated by the computer. The addition of a visual recognition program makes that possible. There is room for variation with regard to methods of implementation, but any potential system would need to share a specific set of traits.

**General Characteristics**

Visual recognition is a processor-intensive task, but adding the real time creation of three-dimensional models necessitates a powerful computer system. In order to achieve this, any system would need a central processing center to receive input from visual sensors and disseminate results to supervisors. Portable computers simply will not provide the necessary power for such detailed analysis.

Additionally, assembling models from single-angle images is not accurate enough, meaning any system would need the ability to observe building elements from multiple angles (Liao, 2000). This could be achieved through static or mobile cameras, although both solutions carry some disadvantages.

Finally, the system is useless if it is unable to effectively communicate the results of its analysis to someone on the construction site. To this end, a Graphical User Interface (GUI) would need to be developed. In order to save on programming costs, such an interface would likely be based on existing three-dimensional modeling programs.

The system will therefore have three main components: a sensor network, a data center, and a user interface. Different design variants can be generated by altering the design of each of these major components. However, many of these design choices will be restricted due to the minimum required levels of performance.

**Necessary Capabilities**

If this technique of automated analysis is to be reliable, it should meet certain required benchmarks for accuracy. This is achievable in most design configurations so long as sufficiently high-resolution digital imaging is used, as demonstrated by the research of the Vexcel Corporation (2003). Their studies show promising levels of accuracy from digital photographs, but their technology still requires manual designation of components by a technician.

In order to remove a technician from the process, the visual recognition algorithm used must be able to differentiate between components at a sufficiently high rate of accuracy. Although mostly used in more controlled environments like manufacturing plants, current generation programs for object identification already boast accuracy rates of over ninety percent (Arif et al., 2010). These programs represent the minimum level of accuracy required for the proposed monitoring system.

Finally, images of appropriate angle and clarity will need to be obtained continuously or at least on a regularly recurrent basis. Methods of achieving this requirement are addressed in subsequent sections, but, in general, each component should be visible from at least two different perspectives (Liao, 2000). This is necessary for the same reason that humans have two eyes. When one is covered, depth perception and special awareness are severely diminished. A computer program attempting to analyze an object from a single angle is similarly hamstrung.

**Difficulties Moving Forward**

Construction sites are by nature chaotic and cluttered visual environments. The required level of accuracy for object verification set out in the previous section becomes much more difficult to achieve when additional elements are added to a camera’s visual field (Berg, 2005). Some method of image filtering must be employed to maintain the required accuracy levels, either through mobile cameras or relevancy ranking algorithms. The latter
necessitates more powerful data servers and increased programming costs, but carries with it a higher potential for universal application.

Furthermore, there are sometimes small deviations from plans that must be made in order to respond to unforeseen real-world conditions (Berger, 2011). While a manual override could largely correct this complication, a truly automated system would require the addition of screening criteria to identify those changes that are not significant and can be ignored. Setting the tolerance levels for such a screening program will require extensive research and will be largely variable and situation-based. Both of these issues will need to be resolved before a real-world system based on this plan can be developed and deployed.

**Possible Methods of Implementation**

There is some design leeway in this plan’s implementation, but not in all areas. The data processing center, located either on-site or at a central facility, will have requirements determined primarily by maximum project size and the processing requirements of the various analysis algorithms employed. The required software has largely been developed already, at least conceptually (Lukins & Trucu, 2007). Applicable visual recognition and photogrammetry algorithms have been published, and the bulk of the required programming will deal with interfacing those processes with each other and with BIM models. The primary component with the most design flexibility is the visual sensor or sensors. There are several possible methods for achieving the necessary levels of image quality and perspective, some of which are detailed and assessed below.

**Fixed Camera Network**

The simplest and most cost-effective solution to the issue of obtaining images of construction projects from multiple angles is a network of fixed cameras. These types of systems are already routinely installed in construction sites for equipment and productivity monitoring (Gong & Caldas, 2010). Because they capture images from a fixed perspective, they are naturally well-suited to the photogrammetry process illustrated in Figure 2. They reduce the financial impact of this project and can be easily wired to transmit data to the central processor.

This option is not, however, without its disadvantages. Because the cameras are placed at fixed points, their layout must be carefully planned in order to achieve the necessary overlapping of visual fields. This in turn requires that the system be custom-built for every unique construction site, making its deployment more of a service than a product and introducing a recurring cost to clients. Unfortunately, the complexity and variety of construction projects will likely necessitate some degree of customization in any system of this variety.

Fixed cameras are a satisfactory option for designing and testing the software components, but are not a viable long-term option. The first incarnation of this system will likely use them, but moving forward a more flexible system of gathering visual input will be required.

**Cable-Controlled Camera System**

A more flexible option, literally and figuratively, is a single camera mounted on a cable-traverse system like the one originally patented by Skycam (Figure 3).

The system allows a camera to move freely through three-dimensional space while maintaining image stability (Skycam.com). If integrated into the proposed analysis system, this computer controlled camera could move freely about the construction site and capture required images from multiple angles. The traverse system’s control program offers a simple method of tracking camera perspective and orientation, while its high-resolution camera is more than capable of capturing images of requisite resolution. It overcomes many of the challenges associated with the static-camera network approach, but poses some difficulties of its own.

Skycam must be mounted at high altitudes and in several places, meaning this approach would be particularly useful in dense urban environments with nearby buildings to anchor the cables from. Poles might also be raised around the site’s perimeter to serve as anchor points. Additionally, the available viewing angles would become increasing limited as the structure under construction increases in height, restricting the camera’s utility and freedom of movement. Also, the large cranes often used in large-scale construction would likely interfere with the cables from time to time.

This type of system could be most readily applied to smaller construction projects with minimal overall changes in elevation. The need for several support cable anchor points introduces custom site-based design difficulties similar to those encountered with the network of static cameras approach. As a result, this is not an optimal solution but could be implemented to investigate the possibilities of using a mobile camera.
Drone-Mounted Camera

Should the research into mobile camera usage prove positive, one interesting universal solution is the use of drone-mounted cameras. Recent innovations in unmanned aerial vehicles, specifically vertical take-off and landing microdrones has made them extremely well suited to surveying and mapping, and they have even been used for photogrammetry (Bonne, 2007). Such a drone could be operated by a technician or even a pre-programmed route and monitor the entire site autonomously, although this approach would be programming intensive and would require significant early-stage planning.

Figure 4: Draganflyer X6 Drone (draganfly.com)

Using a drone (Figure 4) can overcome many of the difficulties associated with this project. The drone can reposition a camera to remove clutter from the visual field and can take images from virtually any perspective. There are some limitations, especially the industry standard twenty-minute flight time (Bonne, 2007). However, because construction projects progress relatively slowly, monitoring flights would only be necessary at regular intervals, allowing the drone time to recharge between flights. Using multiple drones could also remedy this issue. Also, the drones are likely to only be useful in calm weather conditions. Incidentally, image recognition technology also becomes less effective in cloudy scenarios, meaning the system as a whole would likely not function properly in unfavorable climates (Berg, 2005). In spite of these limitations, drones are a versatile and viable platform for mobile image capture.

Video Glasses

One of the most mobile and versatile elements already common in construction sites is a human being. Supervisors, who are already likely wearing safety glasses and making regular rounds to monitor progress, could be issued goggles or glasses with integrated cameras. These devices have been so thoroughly developed and reduced in cost that basic versions are sometimes sold as children’s toys. The instability caused by walking and incidental movement reduces the quality and clarity of the recorded images, but this is not an insurmountable obstacle. Image stabilization and high frame rates can largely remedy these issues. Unfortunately, these higher specifications will also increase hardware costs and the images will not be as stable as those obtained from other mobile cameras.

Combined System

Due to their various limitations, none of the proposed monitoring solutions will likely achieve the desired level of functionality on their own. The configuration with the best potential for success is a combination of a static camera network and either a cable-, drone-, or human-mounted camera. The feeds from the static cameras would be regularly and automatically analyzed for occlusion and visual obstruction. If such a scenario were detected, the mobile camera could then be dispatched to obtain an appropriate image of the obstructed component.

Potential Impact

Real-Time Error Identification

The central objective of this proposal and associated research has been the reduction of construction errors. If successful, the proposed system is capable of largely removing the problem of undetected plan deviations due to insufficient oversight. This would also remove the high costs of repairing those errors and reduce contract overruns by as much as fifteen percent of total contract value (Senaratne & Sexton, 2008). When the fact that this system would mostly represent a one-time cost to construction firms is taken into account, it becomes a potential financial boon for construction firms, especially those bidding on fixed-price contracts.

Change Extrapolation and Plan Modification

A secondary benefit of this system could grow from its ability to compare the real-time images to a BIM model. The analysis program could be written in such a way that all detected plan deviations are extrapolated forward through the construction process. The computer simulates the continued construction of the building, looking for interferences. For example, an elbow joint in a pipe might be oriented upward instead of downward. Seeing this, the computer builds the virtual structure around this misplaced pipe. If the new pipe is occupying what might otherwise have been hollow wall space, then it might not need to be moved. The system would then notify the construction supervisor that, while an error has occurred, it does not give rise to any unplanned interferences. This type of filtering would have dual benefits. First, it would allow irrelevant errors to go uncorrected, saving both time and materials that would otherwise be devoted to repairs.
Additionally, it would reduce the occurrence of bureaucratic change orders, which substantially slow the construction process (Berger, 2011). These simulations allow managers to make more informed decisions and consequently increase overall productivity.

**Increased Efficiency**

All of the benefits described thus far result in overall boosts in efficiency. Projected forward, these increases allow companies to offer more accurate and competitive bids while simultaneously increasing their turnover rates and obtaining more contracts per year. It will also allow supervisors to devote more of their time to productivity monitoring and site organization, potentially resulting in even greater boosts in efficiency.

**Augmented Reality**

One of the most interesting potential benefits of this system would be dependent on the use of video goggles for some degree of image capture. The BIM model could then be superimposed on a supervisor’s vision using heads-up display technology embedded in the glasses. This would also simplify the process of surveying, digitally marking reference points that are currently denoted with manual measurements and marker flags (Gidado, 1996). This offers untold potential for increases in efficiency and could revolutionize the way in which we view construction sites. However, such a system would be complicated and require much more research into several technologies to be practical.

**Conclusions**

**Feasibility**

This research indicates that a system with the proposed specifications and features is, with sufficient experimentation and programming, feasible. The exact hardware requirements and system configuration cannot be determined without further work, but from a theoretical standpoint, all of the required technological elements already exist. Development would largely center on creating interfaces between and automating the different computer processes involved.

**Marketability**

Two key factors make this system highly marketable to construction firms: the competitive nature of the industry drives firms to seek out techniques and practices that allow them to lower their costs and offer more appealing bids. Since current bidding processes usually attempt to account for overruns, this system’s ability to substantially reduce these costs makes it very appealing. Additionally, the fact that it could be sold as a wholesale system and would not cause firms to incur significant future costs improves the system’s marketability. A small number of studies indicating significant cost-saving potential could lead to rapid and wide-scale adoption mirroring the levels currently being experienced by BIM systems.

**Future Work**

Moving forward, there is a great deal of work to be done in realizing this method of quality assurance. Visual recognition needs to be improved and adapted to include greater degrees of relevancy considerations. Additionally, research needs to be conducted to determine the overall system requirements and the best method for obtaining visual data. Finally, the system must be built and tested to make any realistic claims regarding functionality and effectiveness. If the success rates even approach those currently experienced by the constituent base technologies, then this plan carries a great deal of potential.

To truly validate this plan, a limited but functional system must be constructed. In their research, Lukins and Trucco were able to identify seven identical columns based on their relative positions (2007). Similarly, the prototype system will detect only the major structural elements like support beams and foundations. Since the rest of the structure is built around these key components, their positions are most crucial and the most easily monitored. Finer detail will be possible once the basic algorithms have been designed and an image database can be built.

The changes brought on by the integration of these technologies can be likened to the difference between free-hand sketching and coloring. Simply filling in the lines is much easier than drawing the lines themselves. When a color escapes its designated boundary, you know immediately and have the opportunity to correct it. This system offers that same type of feedback.

**References**

3D-4D building information modeling. (n.d.). Retrieved from United States General Services Administration website: [http://www.gsa.gov/portal/content/105075](http://www.gsa.gov/portal/content/105075)


**Biographical Sketch**

William Croll is currently a senior majoring in mechanical engineering at the University of Southern California Viterbi School of Engineering. He worked in the university’s Mork Family Department of Material Science as well as with Cytec Engineered Materials researching and testing current-generation composite materials. He is currently involved as the lead product-designer in a school-supply design startup with a student from USC’s Greiff Center for Entrepreneurship. His future plans include working as a product design engineer and eventually transitioning to business and management. He can be reached by email at croll@usc.edu and welcomes the opportunity to discuss this proposal and associated research.