AISIM 2005 – 1st Annual Inter-university Symposium on Infrastructure Management
Submitted on: 2005-07-07
Word count: 4,750 (including references) + 9 figures and tables

3D CAD DRAWING AS A PRIORI KNOWLEDGE FOR MACHINE VISION IN CONSTRUCTION

Frederic Bosche¹, Carl T. Haas², Carlos H. Caldas³

¹ Frederic Bosche, Ph.D. Candidate, Department of Civil Engineering, University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada; email: bosche.frederic@mail.utexas.edu
² Carl T. Haas, Professor, Department of Civil Engineering, University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada; email: haas@mail.utexas.edu
³ Carlos H. Caldas, Assistant Professor, Department of Civil Engineering, University of Texas, Austin, TX 78712; PH (1) 471-6014; Fax (1) 471-3191; email: caldas@mail.utexas.edu
ABSTRACT

The process of infrastructure management includes: (i) monitoring and evaluation, (ii) planning (iii) design, (iv) construction, and (v) operations and maintenance, all linked through a common management information system. Several of these activities require sensing the infrastructure and its environment. Advanced sensing approaches are therefore more and more viewed as important tools for assessing performance, improving speed and quality of decision making processes, as well as increasing safety. Driven by new technologies, different applications have already been investigated and some of them have already been adopted by companies. They include: as-built status assessment using LADAR/LIDAR technology, compliance checking using a single-axis laser and primitives-fitting algorithms, as well as work environment modeling using a single-axis laser and convex-hull fitting algorithms. LADARs/LIDARS applications are characterized by significant data processing durations (a few days) and single-axis laser based applications are characterized by acquisition times of a few minutes per object.

However, as previously explained, the construction industry needs to take real-time decisions, and for this reason it needs real-time information. Specifically, heavy construction faces safety and productivity issues directly related to the difficulties surrounding the operations of heavy equipment. Real-time automated obstacle avoidance and path planning in-board calculators would definitely improve these factors and more generally the overall performance.

Real-time 3D modeling may become feasible with a new technology called “flash LADAR”. Like a digital video camera it acquires streams of images with frequency up to 30 Hz, but, instead of brightness, each pixel of the 160x124 array stores a range value. Algorithms are now being developed to use these capacities with the aim of revolutionizing machine vision.

Unfortunately, the flash LADAR images are very noisy, much more than brightness images in equivalent environments. This implies that, even with the best efforts for cleaning images, the accuracy for detecting objects and evaluating their sizes and speeds is low and possibly still insufficient for safe obstacle avoidance applications. This issue could be overcome by applying complex pattern recognition algorithms, but they are usually too slow to be applied on real-time applications.

This paper proposes an approach to improving the quality and speed of obstacle detection algorithms by involving some a priori knowledge. Although such knowledge is not available in most pattern recognition problems, it is in the case of construction. 3D CAD provides site layout information that could be used by real-time machine vision programs.

With this objective in mind, the issues of the integration of 3D CAD engines’ and flash LADARs’ data must be addressed. Indeed, design processes provide well defined information including perfectly parallel, perpendicular, flat, etc. forms (strong forms) like pipes, beams, columns and floors, whereas weak and non-parametric forms are produced from existing infrastructure conditions. There are then two major challenges: (1) defining a way to evaluate the two different types of data in a common model, and (2) developing methods for comparing one to the other in order to create reliable environment models in a real-time manner.

The paper presents the investigation currently conducted in the Civil Engineering department at the University of Waterloo. The issues described above are discussed and solutions suggested. Finally, some early results are presented demonstrating the feasibility of this proposed approach and justifying the continuous effort in this direction.
INTRODUCTION

The process of managing infrastructure includes: (i) monitoring and evaluation, (ii) planning, (iii) design, (iv) construction, and (v) operations and maintenance, all linked through a common management information system. The first, fourth and fifth stages of this process require for their proper application an accurate and precise sensing and modeling of the actual infrastructure and its surrounding environment. The purposes are multiple: infrastructure state assessment, work progress monitoring, real-time environment sensing for safety assurance, and ultimately, faster, better and cheaper decision-making processes.

Infrastructure, and more generally, construction and asset sensing and modeling are still most of the time performed manually. For instance, almost 2% of construction work is assigned to manual quality control and tracking of work package completion (1). This is very tedious, timely and costly. Then, safety is the first concern on every construction or O&M operation. Nonetheless, according to OSHA (Occupational Safety and Health Administration), the construction industry continues to report the largest number of fatal work injuries of any industry. A significant number of equipment-human or equipment-objects accidents result from missing safety features installed on currently operated heavy equipment (2). The infrastructure construction sector uses a lot of heavy equipment. Therefore it probably shows even higher injury and fatality levels than other sectors. Developing and using advanced management technologies are therefore necessities that managers must now understand.

The construction research laboratory of the University of Texas at Austin is now investigating a recently introduced range sensing technology, commonly called Flash LADAR. Compared to the LADAR technology, a flash LADAR has a lower accuracy but gains significantly in terms of acquisition time with frequencies up to 10 Hz. This technology appears to be very promising.

While, in safety applications (in construction and O&M), the sensed data is the only one required to perform safety processes, in infrastructure monitoring and evaluation this data must be checked against normality. Normality is often referred to the CAD drawings used for the construction. Monitoring and evaluating infrastructure, like any other construction type, requires checking sensed information (obtained manually or with the use of advanced technologies) against CAD data. This is a major issue with respect to data integration and interoperability. With the explosion of computer and IT usage, interoperability has definitely become a new area of major concern for managers. Intense research is conducted in this area, but it is still in its first steps. In the CAD-sensed data case, some research is already being conducted, notably in Augmented Reality (AR). AR doesn’t require any compliance checking or any data comparison process, only data superimposition. In the present case, CAD and sensed data must be compared in order to either evaluate and monitor the state of the infrastructure asset, or update CAD as-built drawings.

The present paper investigates the feasibility and then proposes a method for integrating and comparing CAD and sensed data for real-time applications such as obstacle avoidance for equipment operations. These two types of data are very different. The problem is therefore to assess the possibility to develop a common model to gather and compare them.

The paper first describes state-of-the-art sensing technologies that can be used in infrastructure management applications. In the following two sections, a framework for integrating and analyzing sensed and 3D CAD data is introduced and demonstrated with preliminary experimental results. Finally, conclusions and future work are presented.
STATE-OF-THE-ART SENSING TECHNOLOGIES

There are several alternative technologies that can be used for sensing the environment: digital (video) cameras, sonars, and more recently range cameras. These approaches are all discussed below but the emphasis is placed upon laser-based technologies.

First, it is interesting to note that a construction environment can be defined as a combination of regular and irregular forms, more specifically of strong, weak, and non-parametric forms (3). Strong forms include parametric prisms, cylinders, blocks, etc., that represent for example, pipes, beams, columns and floors. Then, weak forms have been defined by Hirschberg (4) as parametric objects with “rubberbanding” capabilities. An example is a rectangle which may become an irregular four sided polygon to fit a distorted door frame. Finally, non-parametric forms include wire nets that may represent contour data, polylines that may represent cracks, and occupancy arrays or octrees that may represent amorphous volumes or grossly deformed objects. Interestingly, Most of the recently developed approaches for construction site sensing focus on only one of these types of forms at a time. While each of them is developed for a specific application, none of them, taken separately, can provide good precision and accuracy for real-time sensing of the construction site.

Traditional Approaches: Digital Cameras or Sonars

The most obvious approach to machine vision is to copy Nature. Like many other animals, humans can locate themselves in the three dimensional world thanks to their two eyes performing a stereovision. Stereovision permits to perceive depth without acquiring depth information. In robot vision this can be copied by mounting two cameras next to each other (e.g. (5)). Similar approaches exist and may use only one but omni-directional camera (6). The stereovision approach is mostly used for developing robot navigation in unknown and unstructured areas (5).

Stereovision and omni-directional vision rely on digital (video) cameras, technologies that permit the acquisition of data from an entire scene in a real-time manner. Many data processing methods were developed, rapidly increasing the capacities of using this technology. These methods include from edge detections techniques (7, 8) to advanced calibration techniques (9, 10).

The advantage of the digital (video) camera technology is that it provides similar information to what Humans see. It is therefore easier to understand the acquired data. It is also now very affordable. However, in order to obtain range values, the data must be further processed since digital cameras only provide brightness and focus information.

Another traditional approach for obtaining range scanning information uses the sonar technology. This technology, first investigated in 1916 in England by the Anti-Submarine Detection Investigation Committee (ASDIC), is widely used in the robotics community. Range values are obtained by calculating the time of flight of the echo of sound. Examples of applications in robotics can be found in (11, 12). The advantage of the sonar technology is that, like cameras, it is very simple and cheap. However, more interestingly, it directly provides range distance and can even be used to get speed information by studying the Doppler Effect. It has been studied for several decades so that data processing techniques are well advanced. A limitation is that it doesn’t flash entire scenes at once, but only parts of it, limiting the speed of...
data acquisition. Also, it has low precision at significant distances. For these reasons, sonars are usually used in unconstrained areas.

**Laser–based Approaches**

Existing range scanning techniques for 3D modeling can be categorized into *dense point cloud* and *sparse point cloud* approaches. A sparse point cloud approach enables magnitudes more rapid modeling than dense point cloud approaches. Both approaches have their own merits depending on the application, given the necessary trade-off between quality and speed.

*Single axis laser scanner*

A first approach uses a single axis laser scanner and is based on the sparse point cloud approach. This enables to directly acquire range information. Although this approach reduces the amount of data that need to be processed (an issue specific of LADARs - see below), it requires the user’s input during the acquisition step. This limitation prevents the method from being used in real-time applications. However special operations that require, for instance, the modeling of pre-fabricated materials can definitely benefit from this approach. Research on this approach has been conducted at the University of Texas at Austin, and more details can be found in (13-15).

*LADAR*

A more recent approach for modeling construction sites and work in progress is the use of LAser Detection And Ranging device (LADAR). This technology resembles a digital camera except that instead of brightness, it stores range (depth) information, providing 2½D scans of scenes (although they are usually referred as 3D models).

The LADAR technology shows great promise due to its high accuracy. Nonetheless, its major limitation is the time necessary to process the dense point clouds generated by the LADAR. A picture may require a few days to be processed. Human input is also required to further identify objects by associating regions of interest to a priori known objects (from CAD drawings for instance). An advantage is that all strong, weak and non-parametric forms can be modeled with this method. However, similarly to the laser scanner – based approach, it requires the scanned scene to be static.

While many applications have been suggested and some of them successfully implemented, the long acquisition and processing time, along with the need for the user’s input, prevent this method from being used in real-time applications on construction sites.

When developing machine vision systems for obstacle avoidance applications, other technologies must therefore be considered.

*Flash LADAR*

The following approach uses a flash LADAR. This new technology can be described as a digital camera providing arrays of range distances instead of brightness. Like LADAR, this technology is therefore used for a dense point cloud approach. A flash LADAR acquires data from an array of laser beams reflecting on the surrounding obstacles. Distances are calculated with the time-of-flight method. But, contrary to a regular LADAR that acquires data by scanning “pixel-by-
pixel” the surrounding environment, a flash LADAR acquires the data for all the pixels simultaneously, and can acquire scene pictures with frequencies up to 30 Hz. This is the reason for its name: flash LADAR. More exactly, the National Institute of Standards and Technology (16) defines the term flash as:

“A generic term for a LADAR system comprised of a broad field illumination source (commonly a laser, but for close proximity it can be a bank of LEDs) and an FPA detector, such that the range image is completely acquired simultaneously in one burst. Although in some applications scanners can achieve real-time frame rates, only flash LADARs can achieve very high frame rates. In general, because of the limited number of pixels available on an FPA (currently about 256x256 is the maximum in the laboratory) flash LADARs are unable to achieve the pixel density of a scanner. However, hybrids are being developed in which flash LADARs themselves become the instrument that is mounted e.g. in a pan-tilt platform or beam steering mechanism”.

Flash LADAR constitutes state-of-the-art laser technology and is subject of intense research. FIGURE below shows a flash LADAR developed in collaboration by the Swiss Center for Electronics and Micro-technology (18) [http://www.csem.ch] and the Swiss Federal Institute of Technology (EPFL) (18-20).

FIGURE 1

The characteristics shown in Table 1 are very interesting since they fill the lacks of the previously presented approaches. Interestingly, there is no need for a user’s input during the acquisition process and algorithms can be developed to automatically process the data. The removal of the user out of the process and the high acquisition frequency of this type of devices enable the development of real-time site sensing for machine vision and obstacle avoidance applications. Site sensing can now be performed at high frequency, and consequently, dynamic objects can be detected.

Two approaches to site modeling using a flash LADAR are simultaneously considered in a work conducted by a research team at the University of Texas at Austin (21). On one side, data is analyzed and bounding boxes are fitted to segmented objects (Figure 2). On the other side, the data is filtered and reduced by using the occupancy grid approach. While the first approach preserves the original data for the processing and allows for object segmentation, the second enables some filtering and a reduction of the total amount of data for faster processing. The two methods can actually be combined with the filtering and data reduction steps performed before segmentation.

TABLE 1

FIGURE 2
Summary of Features, Capabilities and Limitations of Existing Systems

Table 2 summarizes the characteristics of the different methods presented in this section. It must be understood that they all have their own merits and limitations, and all have specific situations for which they are more beneficial.

In the case of real-time environment sensing for obstacle avoidance application, it appears that the Flash LADAR technology may provide the most adapted trade-off between the acquisition and processing speeds and the quality of the range information it provides.

TABLE 2

FRAMEWORK FOR INTEGRATING AND ANALYZING FLASH LADAR AND 3D CAD INFORMATION

With the Flash LADAR as the technology for real-time environment sensing, the authors are developing a framework for the integration of real-time sensed data from Flash LADAR and a priori knowledge from 3D CAD drawings. The pursued goal is real-time applications such as obstacle avoidance and path planning for equipment operations.

Design processes provide well defined information including perfectly parallel, perpendicular, flat, etc. forms (strong forms (3)) like pipes, beams, columns and floors, whereas weak and non-parametric forms are produced from sensed existing infrastructure conditions. These two types of forms, very different in nature, are also usually formatted differently. Thus, integrating sensed and 3D CAD data presents two major challenges: (1) defining a way to integrate the two different types of data in a common global model, and (2) developing methods for comparing one to the other in order to create reliable environment models in a real-time manner.

Data Integration

The first issue is that the two data types are dissimilar. While infrastructure drawings are developed on CAD drawings, sensed data are acquired using software like Labview or Matlab. The result is that the two data have completely different formats. A common format must therefore be identified in order to enable their comparison.

Preliminary investigations directed the team to the STL format. STL stands for Stereolithography. This format was developed by the Spanish company 3D Systems for rapid prototyping application. This format, available on all major CAD engines exports each volume of a drawing into forms similar to bounding boxes. Each object is exported as a series of $n$ facets and $n+1$ vertices. Figure 3 illustrates facets and vertices and Figure 4 provides an example of a pipe-spool transformed into its STL format.

FIGURE 3

FIGURE 4
The advantage of the STL format is that, as mentioned previously, objects are modeled almost like bounding boxes which are the way objects are modeled when acquired with the Flash LADAR. Bounding boxes are also defined as series of facets and vertices. The only difference is that bounding boxes cannot be hollow. An illustration of this difference is that the STL format can be used to model a torus while the bounding box approach, with the same input data, would model a disc instead.

Preliminary experiments have been conducted to verify the feasibility of this method. Among others, the team used a model of one canopy constituting the Expodach (Expo-roof) of the Hanover International Exhibition of 2000. Details of the structure and the entire project can be found in (22). This complex structure (see Figure 5) conceived and designed by Thomas Herzog constitutes a severe experimental test. The CAD drawing originally developed on Bentley Microstation was exported in the STL format and imported on Matlab using a code developed by the team. The CAD and Matlab models are respectively shown in Figure 5 and Figure 6. The characteristics of this experiment are the following. The STL file of the canopy (containing about 8200 facets) is 1.5 MB. This file can was imported on Matlab in about five minutes from the CAD engines. Detailed observations show that the model is very accurate and precise. Reducing the level of details when exporting the CAD model into the STL file would decrease the file size and consequently the processing time without seriously impacting the accuracy of the output model. The import time is nonetheless satisfactory since, in the case of real-time applications the information provided by CAD drawings wouldn’t have to be updated at high frequency like the sensed data.

FIGURE 5

FIGURE 6

Therefore, this initial experiment shows that 3D CAD drawings can effectively be exported in formats that can be used on real-time processing software like Matlab. The STL format can be used to compare 3D CAD drawings with bounding boxes resulting from the processed data acquired with a Flash LADAR.

**Data Comparison**

The second challenge to this data integration problem relates to the comparison of the two models now imported in a common model. If the integration process is performed successfully, the CAD and the sensed information should somewhat overlap. Most likely, they will never fully overlap. Then, the task of building the most faithful model consists of defining the most probable shape for each object. An illustration of an overlap is given in Figure 7. The strong form obtained from the CAD drawing is similar but doesn’t completely overlap the non-parametric form obtained from the sensed data.

The characterization of the optimal object form is a complex task that requires intense research effort. Multiple approaches can be investigated, and each approach may be suited for a specific application.
Also, only one example of the expectable overlapping situations has been presented here. Other situations may occur including:

- The CAD and sensed objects don’t overlap
- The CAD and sensed objects partly overlap
- The CAD and sensed objects overlap but have different sizes (or shapes).

For each of these cases, the comparison method must propose an optimal solution that doesn’t reduce the quality of the final model.

FIGURE 7

CONCLUSION AND FUTURE WORK

This paper presents the need the asset management industry has for interoperability technologies. Among others, it is pointed at the fact that most site sensing systems only rely on instant information. It is therefore suggested the use 3D CAD data as a priori knowledge for building or construction environment sensing. A framework for the development of such a technology is introduced. Overall, the literacy and experimental results demonstrate the feasibility of integrating in real-time a priori known (3D CAD) and sensed (from Flash LADAR) information in a common model.

Challenges mainly lay in the data comparison part. With the definition of a common data format, they constitute the major investigation and development focus of the research currently conducted at the Civil Engineering department of the University of Waterloo.

REFERENCE

15  C. Kim. Spatial information acquisition and its use for infrastructure operation and maintenance. Dissertation, University of Texas at Austin, 2004
18  CSEM. SwissRanger SR-2: datasheet. Place. Published, 2004
19  R. Lange. 3D time-of-flight distance measurement with custom solid-state image sensors in CMOS/CCD technology. PhD Dissertation, Electrical Engineering and Computer Science, University of Siegen, 2000
20  O. Gut. Work on the SwissRanger 2 calibration. Geodetic Metrology, ETH Zurich, 2004
23  I. 3D Systems *Stereolithography interface specification*. 1989
LIST OF TABLES

TABLE 1  Technical Characteristics of the Flash LADAR
TABLE 2  Summary of the Characteristics of the Different Sensing and Modeling Methods

LIST OF FIGURES

FIGURE 1  The time-of-flight Swiss Ranger sensor from CSEM.
FIGURE 2  Example of a scene modeled using the convex-hull modeling method: (1) raw scanned data (2) smoothed data (3) segmentation and object detection (4) convex-hull encompassing the detected object.
FIGURE 3  STL format displays volumes with facets defined by three vertices and one normal.
FIGURE 4  Example of object (pipe spool) modeled with the STL format.
FIGURE 5  (1) and (2): Pictures of the Expodach at the International Exhibition of Hanover in 2000, (3) and (4) Wireframe and phong displays of a canopy of the Expodach.
FIGURE 6  Display of the imported STL-formatted file of the structure shown in figure 7.
FIGURE 7  First example of large overlap between an a priori known object and a sensed one.
### TABLE 1  Technical Characteristics of the Flash LADAR

<table>
<thead>
<tr>
<th>Specification</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of pixels (ROI option)</td>
<td>Up to 124 x 160</td>
</tr>
<tr>
<td>Depth resolution</td>
<td>Down to 5 mm</td>
</tr>
<tr>
<td>Wavelength of illumination</td>
<td>870 mm</td>
</tr>
<tr>
<td>Illumination power</td>
<td>800 mW optical</td>
</tr>
<tr>
<td>Maximum range</td>
<td>7.5 m</td>
</tr>
<tr>
<td>Frame rate</td>
<td>Up to 30 fps</td>
</tr>
<tr>
<td>Diagonal field of view (HxW)</td>
<td>+/- 30° (+/- 21° x +/- 23°)</td>
</tr>
<tr>
<td>Interface</td>
<td>USB 2.0</td>
</tr>
<tr>
<td>Connector</td>
<td>Mini USB Type-B</td>
</tr>
<tr>
<td>Power supply</td>
<td>+ 12 V / 1.5 A DC</td>
</tr>
<tr>
<td>Power consumption</td>
<td>18 W max</td>
</tr>
<tr>
<td>Lens</td>
<td>f = 8 mm, F/# = 1.4, M 12x0.5</td>
</tr>
<tr>
<td>Dimensions (24)</td>
<td>135(W) x 45(H) x 32(D)</td>
</tr>
<tr>
<td>Weight</td>
<td>0.2 kg</td>
</tr>
</tbody>
</table>
### TABLE 2  Summary of the Characteristics of the Different Sensing and Modeling Methods

<table>
<thead>
<tr>
<th>Acquisition approach</th>
<th>Quality of range information</th>
<th>Need for user’s input during acquisition</th>
<th>Need for user’s input during data processing</th>
<th>Speed of acquisition</th>
<th>Speed of data processing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digital (video) camera</td>
<td>Dense point cloud</td>
<td>Medium (only if stereovision)</td>
<td>No</td>
<td>No</td>
<td>Real-time</td>
</tr>
<tr>
<td>Sonar</td>
<td>Dense point cloud</td>
<td>Fair</td>
<td>No</td>
<td>No</td>
<td>Real-time</td>
</tr>
<tr>
<td>Laser scanner</td>
<td>Sparse point cloud</td>
<td>Excellent</td>
<td>Yes</td>
<td>No</td>
<td>Medium</td>
</tr>
<tr>
<td>LADAR</td>
<td>Dense point cloud</td>
<td>Excellent</td>
<td>No</td>
<td>Yes</td>
<td>Medium</td>
</tr>
<tr>
<td>Flash LADAR</td>
<td>Dense point cloud</td>
<td>Good</td>
<td>No</td>
<td>No</td>
<td>Real-time</td>
</tr>
</tbody>
</table>
FIGURE 1  The time-of-flight Swiss Ranger sensor from CSEM.
FIGURE 2  Example of a scene modeled using the convex-hull modeling method: (1) raw scanned data (2) smoothed data (3) segmentation and object detection (4) convex-hull encompassing the detected object.
FIGURE 3  STL format displays volumes with facets defined by three vertices and one normal.
FIGURE 4  Example of object (pipe spool) modeled with the STL format.
FIGURE 5 (1) and (2) : Pictures of the Expodach at the International Exhibition of Hanover in 2000, (3) and (4) Wireframe and phong displays of a canopy of the Expodach
FIGURE 6  Display of the imported STL-formatted file of the structure shown in figure 7.
FIGURE 7 First example of large overlap between an a priori known object and a sensed one.