Time-Quality Analysis of Spatial Data Processing for Bridge Management

Pingbo Tang[1] and Zhenglai Shen[1]

[1]Del E. Webb School of Construction, Arizona State University, 651 E. University Drive, Tempe, AZ 85287-0204; PH (480) 727-8105; FAX (480) 965-1769; email: tangpingbo@asu.edu, zshen8@asu.edu

ABSTRACT

Delays in delivering spatial information to support the decisions about maintenance of bridges introduce stoppages into the bridge management workflow. Previous studies show that spatial data collection and analysis are major components of these inspection programs. However, the time consuming spatial data processing cause days or weeks of delays in delivering the needed spatial information. Federal agencies, in many cases, either wait for the tedious spatial data processing, or use partial data processing results containing high uncertainties. Using the processing of 3D laser scanning point clouds as an example, this paper examines the technical feasibility and scientific challenges of quantifying the data processing time and the quality (e.g., accuracy) of derived information (e.g., minimum under-clearance of a bridge) useful for bridge management. Different decisions need different geometric attributes and relationships with various accuracy requirements. The optimal data processing strategy varies with the needed information quality and time limits. It is difficult to manually analyze the time-quality trade-offs of spatial data processing due to its ad-hoc nature and large number of possible parameter settings in data processing workflows. The authors proposed a computational framework that automatically records data processing histories of engineers along with the engineering needs. Analyzing those histories lead to insights into the trade-offs between required computational complexity/time and quality of delivered spatial information. This paper presents time-quality analysis results from two bridge inspection cases.

INTRODUCTION

Delays in delivering spatial information to support the decisions about maintenance of bridges introduce risks and stoppages into the bridge management workflow. According to American Society of Civil Engineers, U.S. needs around $1.1 trillion in addition to the current funding level to improve her infrastructure for preventing $3.1 trillion loss in GDP (ASCE 2013). Facing this budget shortfall, federal agencies invest significant amounts of resources into sensing technologies and inspection programs for achieving proactive resource allocation (Jaselskis, Gao, & Walters 2005; Tang & Akinci 2012a). Unfortunately, an effective use of these technologies is often hampered due to uncertainties. For instance, in Accelerated

Bridge Construction (Iowa Department of Transportation 2012), engineers are puzzled by the misalignments between bridge components, but they cannot use detailed geometric data from 3D laser scanners to conduct tolerance checks because this process may require days or even weeks (Yen, Akin, Ravani, & Lasky 2011).

Spatial data processing and analysis are important for bridge management. For example, engineers use 3D laser scanning for geometric assessments in transportation projects. Unfortunately, large investments have not yet enabled engineers meet the deadlines of delivering the spatial information for timely decision supports. For example, to prevent accidents caused by oversized truckloads, inspectors must conduct detailed 3D data analysis under bridges. However, this process can take days and expose travelers to risk of delayed posting of under-clearance (Fuchs et al. 2011; Tang 2009). On the other hand, it is also risky to act based on incomplete and less detailed information (Dedman 2008). It is thus critical to analyze the data processing time and quality of derived spatial information for effective data-driven decision-making.

Using 3D laser scanning for bridge inspection, this paper examines the technical feasibility and scientific challenges of quantifying the relationship between 3D data processing time and retrieved data quality in terms of Level-Of-Accuracy (LOA) and Level-Of-Detail (LOD) for supporting bridge management. Different decisions in bridge management need different geometric attributes and relationships with various accuracy requirements. The waiting time for the spatial information deliveries depends on both the information quality needs and the technical capabilities of algorithms. The difficulties root in how to capture and characterize such dependences. Manually documenting spatial data processing procedures and analyzing all possible data processing options is impractical. Therefore, the authors propose to develop an automated methodology for recording data processing histories of engineers along with their information needs. Analyzing those histories lead to insights into how delays occur. This paper presents delay analysis results from bridge inspection projects.

The workflow shown in Figure 1 demonstrates why existing studies have not yet been able to create a mapping between spatial data processing workflows and their expected performances in terms of Level-Of-Accuracy (LOA), Level-Of-Detail (LOD), and delay. Some studies have characterized spatial data processing workflows designed manually (Dai et al. 2013; Golparvar-Fard, Bohn, & Teizer 2011; Tang & Akinci 2012b). Since most data processing steps have multiple parameters and different algorithmic variants, manually conducting comprehensive assessments of exponentially large number of combinations of data processing options is unrealistic (Tang and Pradhan 2012). For example, in the workflow shown in Figure 1, multiple plane-fitting algorithms are possible implementations for the step of "Extract Planes on the Road Surface." Manual workflow performance analyses were only able to cover a small portion of all possible workflow instances (Tang & Akinci 2012a).

Several workflow mining techniques can automatically discover workflow structures from transaction or data processing histories (Kindler, Rubin, & Schäfer 2006; Medeiros, Aalst, & Weijters 2003). These methods are able to harvest network structure of workflows from data processing histories, and thus significantly improve the comprehensiveness and efficiency of workflow performance analysis (Freire et al.

2006; Medeiros et al. 2003). However, these methods focused on mining workflow structures, and have limitations in characterizing the computational complexity and data processing time of workflows. As detailed below, a methodology proposed by the authors will augment workflow-mining techniques with methods of capturing and characterizing workflow time-quality performance to address these limitations.

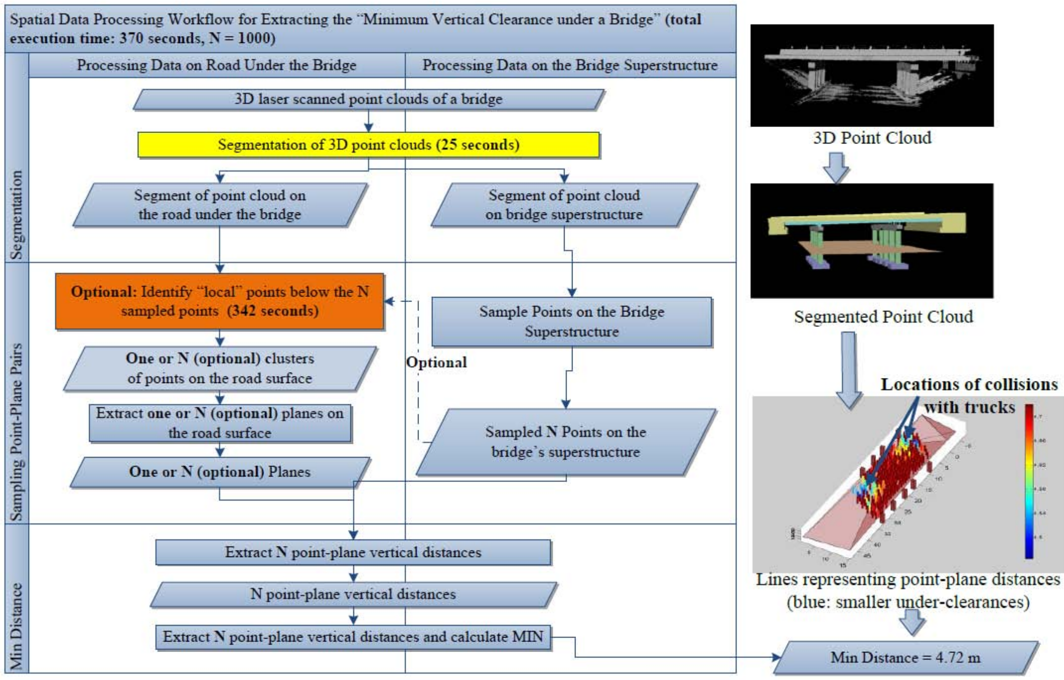


Figure 1. A spatial data processing workflow for extracting the minimum under-clearance of a bridge (yellow/orange boxes are time-consuming steps)

METHOD FOR ANALYZING TIME-QUALITY OF WORKFLOWS

To capture the spatial data processing histories of engineers and researchers, formal representations of spatial data processing workflows and mechanisms for recording data processing histories is necessary. Table 1 shows the formal representation of spatial data processing workflows and mechanisms for recording data processing histories. This computational framework has two major components: workflow representation and workflow recording. Each component includes several elements. Workflow representations formalize workflows so that they can be automatically recorded by computer algorithms and reused in future data processing. The workflow representation components includes five elements: *1) Query*; *2) Data*; *3) Operation*; *4) Workflow*; and *5) Execution History*. For instance, the element *Data* presents data items that can be inputs and outputs of data processing algorithms. Examples of such data include lines, point clouds, bridge superstructure objects and etc. Detailed definitions and examples of other elements of workflow representation are listed in Table 1. Workflow recording includes various methods for automatically recording various workflows during data processing. These workflow recording methods capture: *1) Input-Output*; *2) Configuration of data processing algorithms*; and *3) Performance*. The input-output information captures the structures of spatial data processing workflows, while the configuration elements capture the parameter settings of individual data processing steps in workflows. Similar as a construction schedule, a workflow is a directional network or called “Graph.” The execution time of a workflow will depend on not only individual data processing steps in the workflow, but also the overall structures depicting how the data processing steps linked to each other. Recording the performance of workflows along with workflow structure and algorithm configurations is important to help engineers to identify the “critical” data processing steps, and examine how the data processing steps in a workflow influence each other.

Table 1. Formal Representation of Spatial Data Processing Workflows and Mechanisms for Recording Data Processing Histories

	Elements	Definitions	Examples
Workflow Representation	Query	A specification of targeted geometric attributes of bridge infrastructure components or spatial relationship among components	Areas of cross-section of piers; Distances between the bridge’s superstructure and the highway below the bridge
	Data	Data items that can be inputs and outputs of data processing algorithms	Lines; Point Clouds; Bridge Superstructure Objects
	Operation	Objects encapsulating data processing algorithms in an Object-Oriented method	Recognize bridge superstructure from point clouds; Extracting lines from point clouds
	Workflow	A sequence of data processing operations interconnected by input-output relationships	A workflow taking point cloud of a bridge as inputs for generating the areas of cross-sections of all pairs
	Execution History	Execution histories and performances of operations and workflows	A plane extraction operation takes 0.1 seconds to run on a PC and have average surface error of 5 mm
Workflow Recording	Input-Output Recording	Mechanisms for recording input-output connections among operations that form workflows	Record the input-output relationships between operations in Figure 1
	Configuration Recording	Mechanisms for recording all possible values of data processing parameters and options of operations	From workflows having a “Plane Fitting” step, record all possible values of parameters of plane fitting algorithms
	Performance Recording	Mechanisms for recording the performance history	From workflows having a “Plane Fitting” step, record the execution time and quality of outputs of all “plane fitting” execution instances

TIME-QUALITY ANALYSIS

The total time needed for a bridge inspection roughly includes data collection time and data processing time. Some relevant studies are addressing the data collection aspect. This paper briefly reviews these relevant studies, while focusing on the data processing aspect. (Tang and Alaswad 2012) formed an analytical model to study the relationship between various scanning related parameters (scanning time, required resolution, etc.). Shen et.al (Tang, Shen, Kannan, & Cho 2013) studied the relationship between collected data quality and various data collection parameters. These studies aim at delivering 3D data with the needed qualities within time limits.

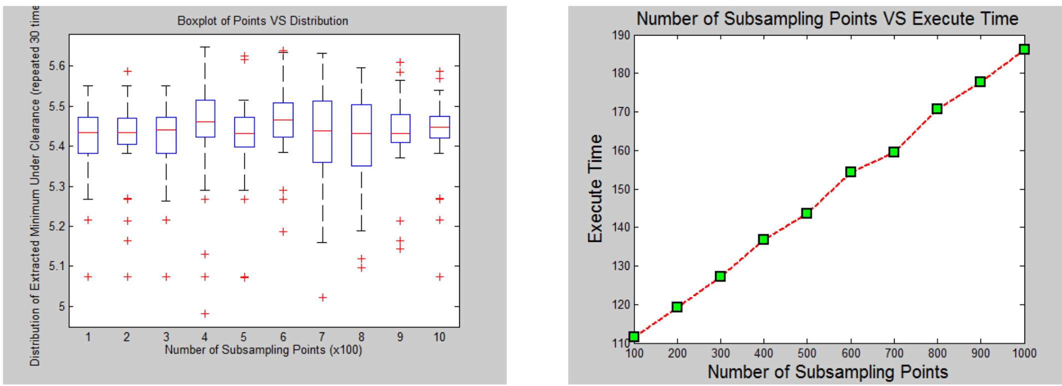
Assuming that the collected data quality can satisfy data quality requirements, does this ensure engineers or researchers can retrieve required information? Given enough data processing time (weeks, months), the quality of the required information will be satisfied with a high probability. The retrieved information, however, will be less valued, due to its delays. The data processing time will be determined by the workflow itself (configuration of algorithms) and the amount of data (inputs and outputs), which are recorded by workflow recording methods.

Based on the computational framework presented above, the researchers can capture data processing workflows and analyze how the amount of data and the configurations of workflows influence their performance. An optimal workflow configuration will achieve a balance between the quality of derived spatial information and the data processing time. Using the 3D laser scanning point cloud of a bridge shown in Figure 1, the authors conducted two detailed studies to show how to use the proposed computational framework to capture and analyze time-quality trade-offs of two typical workflows frequently conducted by bridge inspectors. These two workflows are: 1) a workflow for extracting the minimum under-clearance of a bridge, and 2) a workflow for extracting the average cross-section areas of bridge piers. The following section presents these the time-quality analysis of these two workflows.

TIME-QUALITY ANALYSIS OF TWO 3D DATA PROCESSING WORKFLOWS NEEDED IN BRIDGE MANAGEMENT

Case 1: Extracting Minimum Under-Clearance of a Bridge. This workflow is the one shown in Figure 1. It aims at identifying the minimum vertical distances between the bridge superstructure and the road below. The authors implemented this workflow based on the workflow representations presented above, and used the automatic timing functions of MATLAB to record the execution time of each step in this workflow. In this workflow, bridge inspectors often need to consider how many measurements would be sufficient for achieving a reliable estimation of the minimum vertical under-clearance. In addition, the number of measurements (sampling points under the bridge) can influence the data processing time. In this study, the authors characterize how the number of vertical distance measurements influences the data processing time and accuracy of the derived minimum under clearance, which is a type of “quality” of spatial information.

Figure 2 (a) shows relationship between the number of sampling points and the distribution of extracted minimum under-clearances, while Figure 2(b) represents the time needed to execute the workflow as shown in Figure 1 with different number of sampling points. While the execution time increases as the number of sampling points increased (Figure 2(b)), the extracted minimum under-clearance does not decrease with the increasing of the number of sampling points. It reflects the fact that randomly sampling points either in the superstructure of the bridge or the road under the bridge cannot ensure the calculated under-clearance will reach a global minimum. This observation shows the necessity of better understanding how larger number of measurements helps the decision making of bridge management agencies: more time does not mean better information.



(a) (b)

Figure 2. Extracted minimum under-clearances and execute time: (a) distribution of extracted minimum under-clearances; (b) execution time

Case 2: Extracting Cross-Section Areas of Bridge Piers. This workflow extracts the cross-sections of bridge piers for identifying weak parts of bridge substructures in safety assessment. Figure 3 shows this workflow.

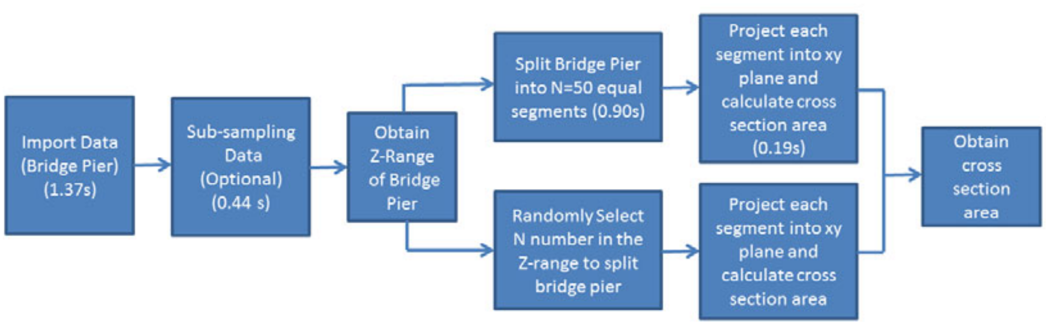
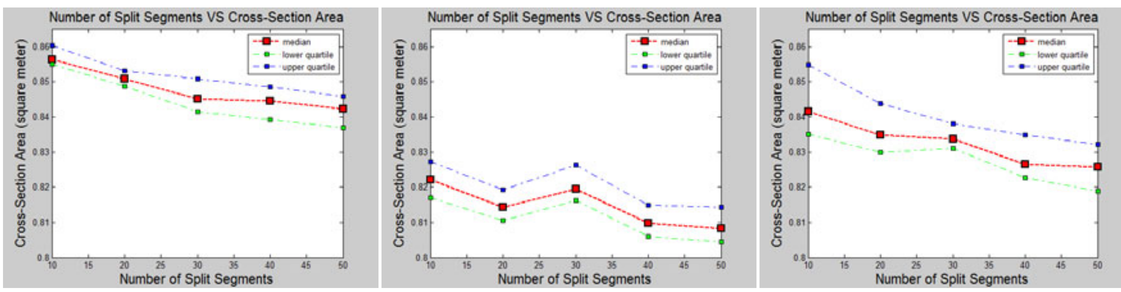


Figure 3. Workflow for extracting cross-section areas of bridge piers

Given data points of bridge piers, this workflow generates slices of data points at points sampled on the axis of bridge piers, and then estimates the cross section areas based on those slices of data points. The sampled points are either uniformly spacing on the axis of a pier, or randomly sampled. Figure 4 shows the distribution of extracted cross-section area according to the number of sampled slices. It shows how the number of slices influences the lower quartile, median and upper quartile of the extracted cross-section areas of a column. The authors studies four columns (a,b,c in Figure 4) and found that different data qualities on different columns can significantly influence the results of cross-section analysis. Overall, the cross-section area decreases slightly with the increases of the number of slices. This phenomenon indicates that dense cross-section measurements along the axis of a column can capture some parts that having less 3D data points due to occlusions or sharp incidence angles between the laser and pier surfaces. Adding those cross-sections generated from sparse and partially occluded cross-sections will decrease the statistical values about cross-section areas shown in figure 4. The extracted cross-section areas of Pier 1 and 3 are quite close to each other. For Pier 2, the extracted cross-section areas are much smaller. This phenomenon also caused by the density of collected point clouds. As we found, this also has relation with scan planning.



(a) (b) (c)

Figure 4. Extracted cross-section area for bridge piers: (a) pier one; (b) pier two; and (c) pier three

CONCLUSIONS

Applying state-of-art technology, e.g., spatial information collected by laser scanner, for bridge management has the potential to reduce the time for inspection results distribution. In this paper, we studied the relationship between time and quality of spatial data processing using the computing framework we proposed. We reached two conclusions. First, the proposed methodology has the potential to formalize and regulate ad-hoc workflows written or executed by both researchers and engineers. In the future, this methodology will support the development of spatial data processing workflow mining and learning systems for automatically discover time-quality relationships of spatial data processing workflows. Second, the case studies presented show that time-quality relationship is complicated, and is influence by both data quality and workflows themselves. Therefore, there is still a great research gap between using spatial information (e.g., scanning point clouds) for bridge management. In the future, the authors will use the theory of computational complexity (Big-O notation) for examining the time-quality trade-offs of workflows on network platforms.

REFERENCES

ASCE. (2013). ASCE's Failure to Act economic report series. Retrieved from http://www.asce.org/economicstudy/

Dai, K., Boyajian, D., Liu, W., Chen, S.-E., Scott, J., and Schmieder, M. (2013). Laser-Based Field Measurement for a Bridge Finite Element Model Validation. *Journal of Performance of Constructed Facilities*, 130605214305002. doi:10.1061/(ASCE)CF.1943-5509.0000484

Freire, J., Silva, C. T., Callahan, S. P., Santos, E., Scheidegger, C. E., and Vo, H. T. (2006). Managing rapidly-evolving scientific workflows. In L. Moreau and I. Foster (Eds.), *IPAW'06 Proceedings of the 2006 international conference on Provenance and Annotation of Data* (Vol. 4145, pp. 10–18). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/11890850

Golparvar-Fard, M., Bohn, J., and Teizer, J. (2011). Evaluation of image-based modeling and laser scanning accuracy for emerging automated performance monitoring techniques. *Automation in Construction*, *20*(8), 1143–1155. Retrieved from http://www.sciencedirect.com/science/article/pii/S0926580511000707

Iowa Department of Transportation. (2012). A draft policy for Accelerated Bridge Construction (ABC) project developmen. Ames, IA: Iowa DOT. Retrieved from http://www.iowadot.gov/bridge/DraftABCPolicyGuidelines.pdf

Jaselskis, E., Gao, Z., and Walters, R. (2005). Improving transportation projects using laser scanning. *Journal of Construction Engineering and Management*, *131*(3), 377–384. Retrieved from http://link.aip.org/link/?JCEMD4/131/377/1

Kindler, E., Rubin, V., and Schäfer, W. (2006). Incremental Workflow Mining Based on Document Versioning Information. In M. Li, B. Boehm, and L. J. Osterweil (Eds.), *Unifying the Software Process Spectrum* (Vol. 3840, pp. 287–301). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/11608035

Medeiros, A. K. A. de, Aalst, W. M. P. van der, and Weijters, A. J. M. M. (2003). Workflow mining: Current status and future directions. In R. Meersman, Z. Tari, and D. C. Schmidt (Eds.), *On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE* (Vol. 2888, pp. 389–406). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/b94348

Tang, P., and Akinci, B. (2012a). Automatic execution of workflows on laser-scanned data for extracting bridge surveying goals. *Advanced Engineering Informatics*, *26*(4), 889–903. doi:10.1016/j.aei.2012.07.004

Tang, P., and Akinci, B. (2012b). Formalization of workflows for extracting bridge surveying goals from laser-scanned data. *Automation in Construction*, *22*(3), 306–319.

Tang, P., and Alaswad, F. S. (2012). Sensor Modeling of Laser Scanners for Automated Scan Planning on Construction Jobsites. In Amr Kandil and H. Cai (Eds.), *Construction Research Congress 2012*. West Lafayette, IN: ASCE.

Tang, P., Shen, Z., Kannan, O., and Cho, Y. K. (2013). As-Built Error Modeling for Effective 3D Laser Scanning on Construction Sites. In I. Brilakis (Ed.), *2013*

ASCE Workshop of Computing in Civil Engineering. Los Angeles, CA, USA: American Society of Civil Engineers.

Yen, K. S., Akin, K., Ravani, B., and Lasky, T. A. (2011). *Accelerated Project Delivery: Case Studies and Field Use of 3D Terrestrial Laser Scanning in Caltrans Projects: Phase I – Training and Materials*. *Aerospace Engineering* (p. 77). Davis, CA. Retrieved from http://ahmct.ucdavis.edu/pdf/UCD-ARR-08-06-30-06.pdf