Leadership in Sustainable Infrastructure

Leadership en Infrastructures Durables

Vancouver, Canada

May 31 – June 3, 2017/ *Mai 31 – Juin 3, 2017*

FRAMEWORK OF AUTOMATED AS-BUILT PHOTO CONTENTS AND CONTEXT RETRIEVAL USING 4D BIM

Park, Jaehyun[1] and Cai, Hubo[2,3]
[1] Lyles School of Civil Engineering, Purdue University, 550 Stadium Mall Drive, West Lafayette, IN 47907, United States
[2] Division of Construction Engineering & Management, Lyles School of Civil Engineering, Purdue University, 550 Stadium Mall Drive, West Lafayette, IN 47907, United States
[3] hubocai@purdue.edu

Abstract: Daily as-built photography has been considered an efficient means of capturing construction progress status in a timely manner. During the construction phase, many as-built photos are taken; however, the contents and context of each photo is manually grasped. Thus, this practice tends to be time-consuming and prone to error. A promising alternative to the current practice is four-dimensional (4D) building information modeling (BIM) (3D BIM plus time), a powerful tool used to visualize construction plans and milestones. The potential use of 4D (BIM) is suggested to support progress monitoring and tracking when daily construction progress is reflected in the 4D BIM model every day. Thus, the 4D BIM model is the best source for describing as-built photos from a particular day. This paper presents a framework of automated as-built photo contents and context retrieval using 4D BIM. Content-based image retrieval (CBIR) technology is introduced and tested for the framework.

1 INTRODUCTION

Project participants record as-built statuses frequently to capture construction progress during the construction phase (Rankohi and Waugh 2014). In particular, daily as-built photography has been considered an efficient means for capturing as-built statuses in a timely manner (Goeder and Meadati 2008; Rankohi and Waugh 2014). A large number of as-built photos is generated during the construction phase, and each photo includes its contents and context. However, in the current practice, the contents and context of the as-built photos must be manually grasped. This procedure has tended to be time-consuming and prone to error (Hegazy and Abdel-Monem 2012).

A promising alternative to this practice is four-dimensional (4D) building information modeling (BIM) (3D BIM plus time), a powerful tool for visualizing construction plans and milestones. A potential application of 4D BIM is to support construction progress monitoring and tracking, based on the premise that daily construction progress is incorporated into the 4D BIM model every day (Staub-French and Khanzode 2007). In this way, the 4D BIM model would contain various pieces of valuable information about the construction tasks (Pătrăucean et al. 2015) and would comprise a rich source of data for describing as-built photos from a particular day.

This paper presents a framework for automated as-built photo contents and context retrieval using 4D BIM that enables automatic detection of the photo-taken location and retrieval of the photo contents and context. For the proposed framework, content-based image retrieval (CBIR) technology is utilized to find the closest photo-taken location through comparisons between as-built photos and rendering images from the daily 4D BIM model. The remainder of this paper is organized as follows. First, the CBIR technology is reviewed.

Subsequently, fundamental principles and the workflow of the proposed framework are explained in detail. Then, a test based on CBIR technology is conducted to verify the feasibility of this framework. Finally, conclusions and suggestions for further research are provided.

2 LITERATURE REVIEW

This section summarizes the concepts of content-based image retrieval (CBIR) technology and vector space modeling to develop the framework of automated as-built photo documentation using daily 4D BIM.

2.1 Content-Based Image Retrieval (CBIR) Technology

CBIR is the methodology of searching digital images in a database based on the contents of the image rather than the metadata (e.g., keywords, tags, or descriptions associated with the image) (Lux and Marques 2013). The term “content” generally means colors, shapes, textures, or any other information that can be derived from the image itself (Lux and Chatzichristofis 2008). Based on this definition, various descriptors have been proposed for color, texture, and shape, as shown in Table 1 (Sidhu and Saxena 2015).

Table 1: Various CBIR descriptors

Content	Descriptors
Color	Color histograms, color correlogram, dominant color descriptor (DCD), color coherence vector (CCV), weighted average
Texture	Gray-level co-occurrence matrix (GLCM), Markov random field (MRF) model, simultaneous auto-regressive (SAR) model, local binary pattern (LBP), edge histogram descriptor (EHD), Gabor filter, Gabor wavelets, wavelet transform, Tamura features
Shape	Pseudo-Zernike moments (PZM), polar complex exponential transform (PCET), polar cosine transform (PCT), pyramid histogram of oriented gradients (PHOG)

In this research, as-built photos are compared to generated rendering images from the daily 4D BIM model. Hence, color descriptors do not consider the development of the proposed framework. Among CBIR descriptors, according to the previous research, EHD is suitable for texture and shape-based image retrieval and can be used for a type of coarse representation of object edges (Xu and Zhang 2006). PHOG outperforms other shape-based descriptors (Fierro-Radilla et al. 2015). Consequently, for the proposed framework, two descriptors (i.e., EHD for texture detection and PHOG descriptor for shape detection) are tested to verify the feasibility of the using them to search the most similar image (used for detecting photo-taken location based on CBIR technology) and to find the most appropriate descriptor for the proposed framework.

2.2 Vector Space Modeling

Vector space modeling is one of the most prominent concepts to evaluate similarities between images in the CBIR technology (Lux and Marques 2013). In the vector space model, the similarity between images I_i and I_j is determined by the cosine coefficient $s(I_i, I_j)$, as shown in Equation 1.

$$[1]\quad s(I_i, I_j) = \frac{I_i I_j}{|I_i||I_j|}$$

The target image, for example, has a term vector of I_t = (4, 3, 1, 1, 0, 0, 0, 2, 2, 4, 2, 1) through the retrieval process using a certain descriptor. Assuming that the database has three images with the same vector structures (i.e., I_1 = [9, 6, 3, 2, 2, 2, 0, 2, 0, 0, 0, 0], I_2 = [1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0], and I_3 = [1, 0, 0, 0, 2, 0, 0, 2, 1, 1, 1, 1]) through the retrieval process, the similarity between target image and the first, second, and third image can be determined as shown in Equations 2,3, and 4.

$$[2]\quad s(I_t, I_1) = \frac{I_t I_1}{|I_t||I_1|} = \frac{(4\times9+3\times6+1\times3+1\times2+2\times2)}{(7.5\times11.9)} = 0.71$$

$$[3]\quad s(I_t, I_2) = \frac{I_t I_2}{|I_t||I_2|} = \frac{(4\times1+3\times1+1\times1+1\times1+2\times1+2\times1)}{(7.5\times3.0)} = 0.58$$

$$[4]\quad s(I_t, I_3) = \frac{I_t I_3}{|I_t||I_3|} = \frac{(4\times1+2\times2+2\times1+4\times1+2\times1+1\times1)}{(7.5\times3.6)} = 0.63$$

Based on the cosine coefficient that was calculated through these equations, the first image (0.71) is the most similar image to the target image.

3 PROPOSED FRAMEWORK

A framework was created in this study to automatically retrieve contents and context of as-built photos. Figure 1 illustrates the framework and its workflow. The framework consists of four phases: image data set generation, similar image search, contents and context retrieval.

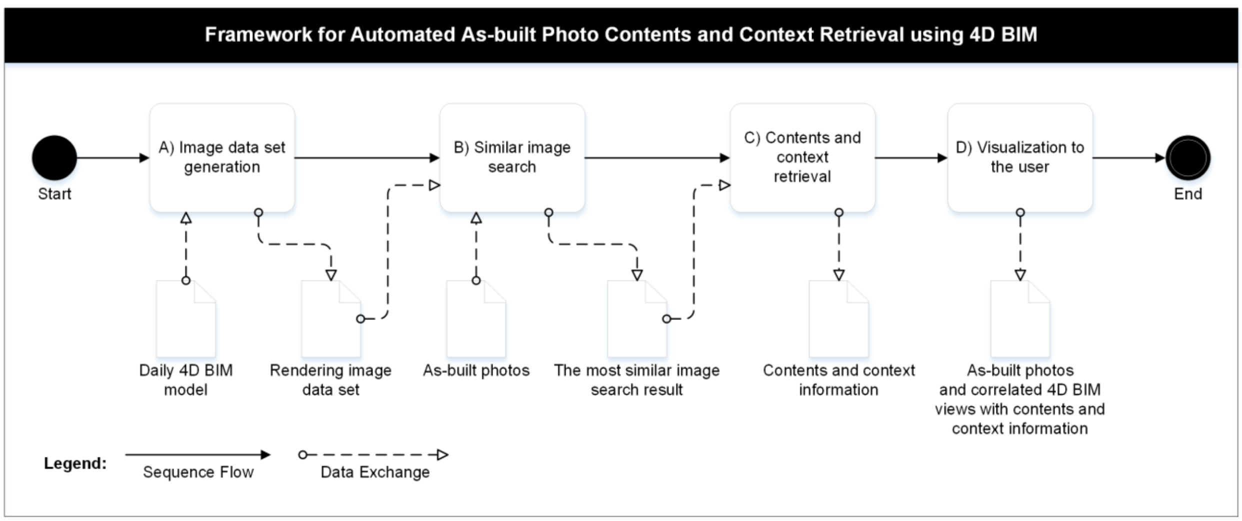


Figure 1: Methodology workflow

This framework has two assumptions: the 4D BIM model is developed and updated every day to reflect daily construction progresses, and daily as-built photos are documented and managed using a database management system (DBMS).

3.1 A) Image Data Set Generation

The closest photo-taken location is determined through searching the most similar 4D BIM rendering images, as to the target as-built photos, using CBIR technology in the proposed framework. Consequently, to utilize CBIR technology, meta-image data for searching is required. The image data set is generated from the 4D BIM model. When each rendering image is generated from the 4D BIM model, model location is recorded into each generated image for the further phase.

3.2 B) Similar Image Search

In the literature review chapter, two descriptors (EHD and PHOG) were selected for a CBIR-based, similar image search methodology. Using these descriptors, the cosine coefficient of each rendering image that is generated from the 4D BIM model as a meta-image data set is calculated to find the most similar image to the target as-built photo.

3.3 C) Contents and Context Retrieval

After the search process, the most similar 4D BIM image to the given as-built photo is detected. The last task before visualization to the user is identifying the BIM objects that appear in the searched 4D BIM image. In the 4D BIM model, each BIM object is linked with project tasks and has its own properties. Hence, photo contents and context can be retrieved when the objects that appeared in the searched image is identified through this phase.

3.4 D) Visualization to the User

As the last phase, as-built photos and retrieved photo contents and context (e.g., project tasks, construction progress information, BIM object lists, and photo-taken location in the daily 4D BIM model) are visualized to ther user.

4 VERIFICATION

The validation of the proposed framework for automated as-built photo contents and context retrieval using 4D BIM is described in this section. The verification tasks were conducted to ensure that the photo-taken location of the given as-built photo was automatically determined from the CBIR-based methodology. For these tasks, the 4D BIM model and two specific as-built photos (i.e., one for inside and one for outside) from August 9, 2016 were used, as shown in Figure 2.

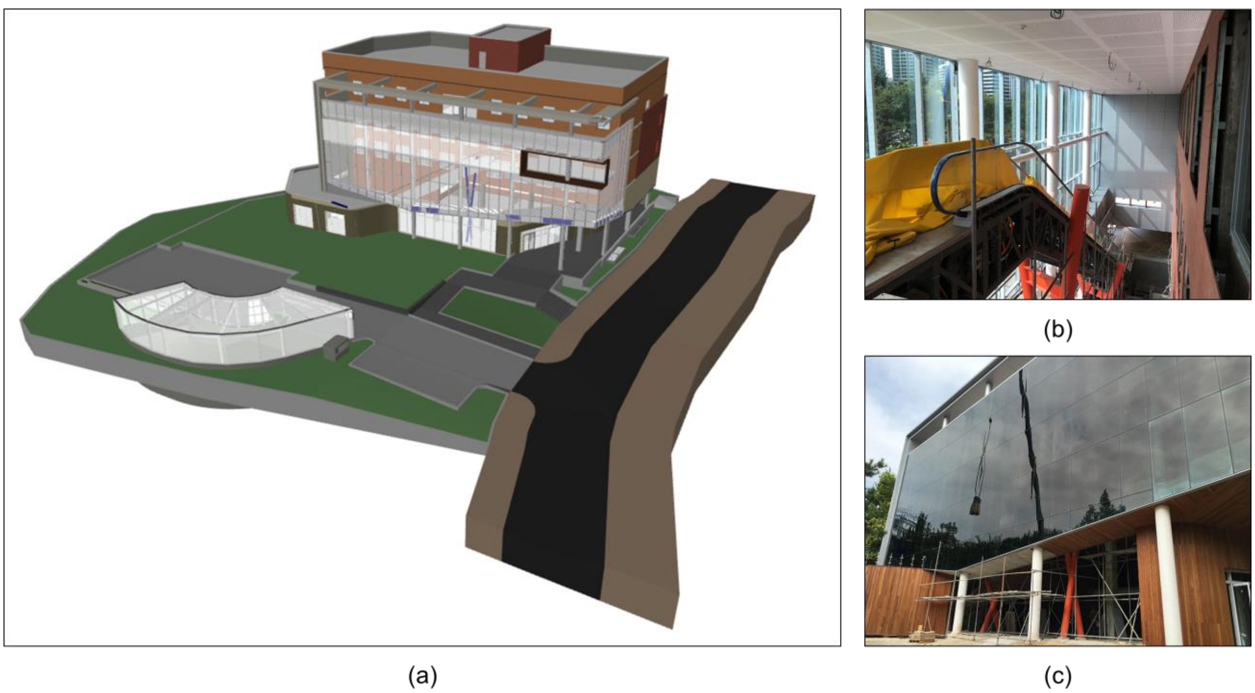


Figure 2: Daily 4D BIM model and as-built photos from August 9, 2016: (a) daily 4D BIM model; (b) inside as-built photo (escalator installation); (c) outside as-built photo (curtainwall installation)

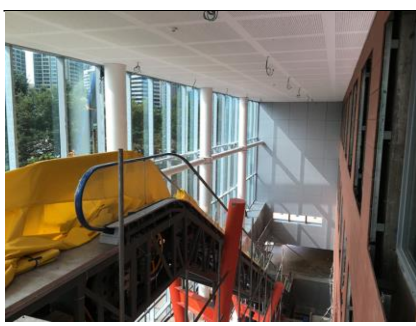
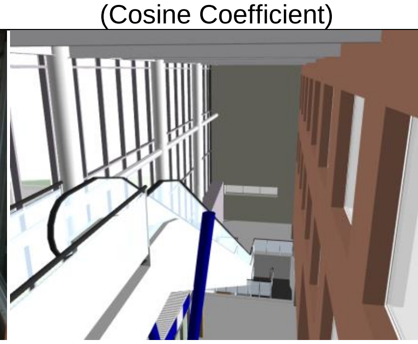
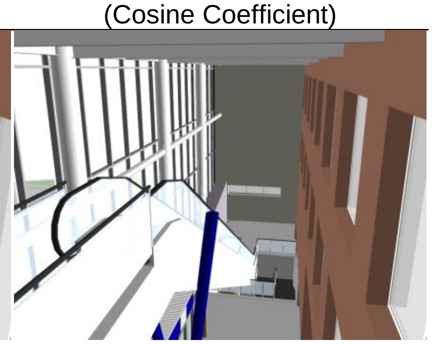
For these validation tasks, approximately 2,000 images were generated as an image data set from the daily 4D BIM model of August 9, 2016. Especially, for the image generation, 3D coordinates-based grid structure was used to the 4D BIM model. The Lucene Image Retrieval (LIRE) (LIRE 2017), an open-source CBIR project, was utilized to search the most similar image for detecting photo-taken locations using two descriptors (i.e., EHD and PHOG), as shown in Figure 3.

Figure 3: Example of the most similar image search process using LIRE (Case 1, EHD)

4.1 Case 1: Escalator Installation (inside)

This section presents the CBIR-based image search results of the as-built photo, shown in Figure 2 (b). A CBIR-based similar image search task using EHD and PHOG was executed to find the closest photo-taken location in the 4D BIM model. Table 2 presents CBIR-based similar image search results of each descriptor.

Table 2: CBIR-based most similar image search results (Case 1)

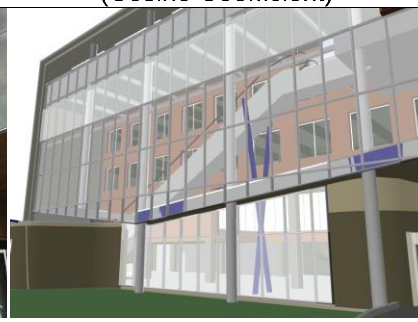
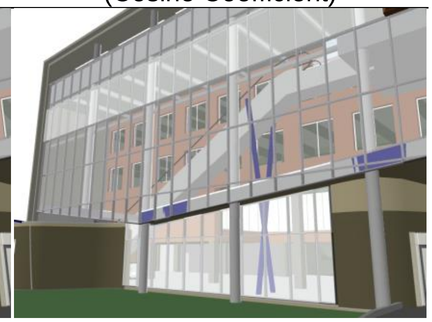
As-built photo	Most similar image search results	
	EHD (Cosine Coefficient)	PHOG (Cosine Coefficient)
		
	(0.399805)	(0.312988)

The CBIR-based similar image search results using both EHD and PHOG presented considerably accurate results. Two descriptors derived the same result. However, the calculated cosine coefficient values were different. The value from the EHD (0.399805) was bigger than that from the PHOG (0.312988).

4.2 Case 2: Curtainwall Installation (outside)

Similar to the previous section, the section presents the image search results using the as-built photo, shown in Figure 2 (c). Table 3 presents the results.

Table 3: CBIR- based most similar image search results (Case 2)

As-built photo	Most similar image search results	
	EHD (Cosine Coefficient)	PHOG (Cosine Coefficient)
		
	(0.410531)	(0.450000)

The CBIR-based similar image search results from the two descriptors presented considerably accurate results, just as in the previous case. However, the calculated cosine coefficient values showed a different aspect. The value from the PHOG (0.450000) was larger than that from the EHD (0.410531).

4.3 Discussion

The verification tasks show that the CBIR-based photo-taken location search is a feasible method for the proposed framework. However, cosine coefficient values showed different results although the two descriptors derived the same results for two cases. Thus, combining two descriptors for the most similar image search process can be considered for deriving more accurate results.

Meanwhile, the two cases illustrated considerably accurate results; these were not exactly the same as the given as-built photos. It can be concluded that a limited image data set generates errors and that this problem can be resolved by increasing the content of the 4D BIM image data set.

5 CONCLUSION

This paper presented a framework for automated as-built photo contents and context retrieval using 4D BIM. The framework consists of four phases: image data set generation, similar image search, contents and context retrieval, and visualization to the user. In the proposed framework, CBIR technology is used to obtain the photo-taken locations of the given as-built photos. The validation tasks were conducted to ensure the feasibility of the CBIR-based framework and the results showed that CBIR technology can identify the closest photo-taken locations. Ongoing research is being conducted to automatically generate image data sets from the 4D BIM model and retrieve photo contents and context that is described in the given as-built photos.

References

Fierro-Radilla, A.N., Nakano-Miyatake, M., Perez-Daniel, K., Perez-Meana, H., Garcia-Ugalde, F., Cedillo-Hernandez, M. 2015. Image Retrieval System Based on Combination of Color, Texture and Shape Features. The 7th International Conferences on Advances in Multimedia (MMEDIA 2015), Barcelona, Spain, 37-43.

Goedert, J.D., and Meadati, P. 2008. Integrating Construction Process Documentation into Building Information Modeling. *Journal of Construction Engineering and Management*, ASCE, **134**(7): 509-516.

Hegazy, T., and Abdel-Monem, M. 2012. Email-based system for documenting construction as-built details. *Automation in Construction*, Elsevier, **24**: 130-137.

LIRE. 2017. LIRE: Lucene Image Retrieval. <http://www.lire-project.net> (Feb. 8, 2017).

Lux, M., and Marques, O. 2013. Visual Information Retrieval using Java and LIRE. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, **5**(1): 1–112.

Lux, M., and Chatzichristofis, S.A. 2008. LIRe: Lucene Image Retrieval - An Extensible Java CBIR Library. *Proceedings of the 16th International Conference on Multimedia 2008*, Vancouver, British Columbia, Canada, 1085–1088.

Pătrăucean, V., Armeni, I., Nahangi, M., Yeung, J., Brilakis, I., and Hass, C. 2015. State of research in automatic as-built modelling. *Advanced Engineering Informatics*, Elsevier, **29**(2): 162-171.

Rankohi, S., and Waugh, L. 2014. Image-Based Modeling Approaches for Projects Status Comparison. *CSCE 2014 General Conference*, Halifax, Canada.

Staub-French, S. and Khanzode, A. 2007. 3D and 4D Modeling for Design and Construction Coordination: Issues and Lessons Learned. *Journal of Information Technology in Construction (ITcon)*, **12**: 381–407.

Sidhu, S., and Saxena, J. 2015. Content Based Image Retrieval A Review. *International Journal of Research in Computer Applications and Robotics*, **3**(5): 84-88.

Xu, F., and Zhang, Y. 2006. Evaluation and comparison of texture descriptors proposed in MPEG-7. *Journal of Visual Communication & Image Representation*, Elsevier, **17**: 701-716.