BUILDING MODEL OBJECT CLASSIFICATION FOR SEMANTIC ENRICHMENT USING GEOMETRIC FEATURES AND PAIRWISE SPATIAL RELATIONS

Ling Ma[1], Rafael Sacks[2], and Uri Kattel[3]

Abstract: Semantic enrichment is a process of supplementing/correcting information in a poorly prepared BIM model. Object classifications are essential information, but are commonly missing or incorrectly represented when transferring a BIM model or creating a model using tools customized for other domains in design. Automated compilation of 'as-is' BIM models from point cloud data also requires object classification, as well as 3D reconstruction. We present a systematic approach to classifying objects in a BIM model, for use in future semantic enrichment systems. Previous work on object classification in BIM model enrichment was restricted by its limited ability to accurately interpret geometric and spatial features and by the constraints of Boolean logic rules and the rule compilation process. To address these issues, we propose a procedure for establishing a knowledge base that associates objects with their features and relationships, and a matching algorithm based on a similarity measurement between the knowledge base and facts. An implementation on a synthetic bridge model shows that whereas some objects can be classified by shape features alone, most objects require the use of spatial relations for unique classification. Spatial context is more likely uniquely identify an object than shape features are.

Keywords: Building information modelling, semantic enrichment, geometric feature, spatial relation, object classification.

1 INTRODUCTION

Semantic enrichment for building models addresses the information interoperability problem (Eastman et al., 2008), which continues to hamper the use of Building Information Modelling (BIM) despite the success in implementing a standard data exchange schema - the Industry Foundation Classes (IFC) (BuildingSmart, 2013). Professions in different domains use different BIM software applications, and models built in one software can be exchanged using the IFC file intermediary, but in almost all cases, some of the information is lost or incorrectly presented.

Information typically missing in an exchanged BIM model include the objects' classification, relationships and other alphanumerical data. Objects may be misclassified due to BIM software limitations. For example, precast joint fillers do not exist in many BIM tools' library, so the modeller is forced to use an object that has similar shape (e.g., a

---

[1] Lecturer, School of Art, Design and Architecture, University of Huddersfield, Huddersfield, United Kingdom, l.ma@hud.ac.uk

[2] Associate Professor, Faculty of Civil and Environmental Engineering, Technion - Israel Institute of Technology, Haifa, Israel, cvsacks@techunix.technion.ac.il

[3] Graduate student, Faculty of Civil and Environmental Engineering, Technion –Israel Institute of Technology, Haifa, Israel, uri.kattel@gmail.com

column) to represent it (Belsky et al. 2015). In this case, the model conveys wrong information.

Semantic enrichment is also useful in the area of compilation of 'as-is' or 'as-built' BIM models from spatial data collected on site through state-of-the-art surveying technologies, such as laser scanning and photo/videogrammetry (Zeibak-Shini et al., 2016, Brilakis et al., 2010). These large data sets must be converted into 3D primitives and then identified as context-specific objects. Current practice requires intensive operations by experienced BIM modellers, and the problem has attracted many research efforts to automate the procedure (Bosche and Haas, 2008, Zhang et al., 2015). However, beyond creating the 3D geometry, classification of the 3D objects remains a challenge (Brilakis et al., 2010, Tang et al., 2010).

Development of a semantic enrichment system has been inspired and informed by the research of semantic query languages for BIM (Mazairac and Beetz, 2013), semantic rule-checking systems for BIM (Pauwels et al., 2011, Eastman et al., 2009), and BIM model query & comparison using spatial and topological relationships (Daum and Borrmann, 2014, Borrmann and Rank, 2009, Tang et al., 2016). SeeBIM, namely Semantic Enrichment Engine for BIM, (Belsky et al., 2015) is an early software prototype whose primary aim was to establish the feasibility of the approach. The tool parses an IFC file to extract objects' shapes, relationships and other attributes. It then applies forward chaining rules to infer additional facts about the model. The rules are compiled in advance by experts in the domain of interest. It records the results in an enriched IFC file. Experiments conducted using SeeBIM for two domains – precast concrete modelling (Belsky et al., 2015) and automated detailed design (Aram, 2015) – showed how the approach could be used to add information to a BIM model.

However, SeeBIM has an important limitation. The rule compilation process in SeeBIM is essentially a social exercise that entails interviewing domain experts to elicit their knowledge and compiling it in the form of if-then rules, in which the truth values of variables may only be yes or no. However, the complexity of object classification makes such Boolean logic very limiting, as some rules may not have unequivocal true/false values for all types of objects without exception. Hence, a fuzzy similarity measurement should be devised for the object inference, and a formalized process to compile such logic is needed.

We address this issue in semantic enrichment for BIM in the area of object classification. The procedure that was devised is described in the next two sections. The following section illustrates the procedure by application to classify the objects of a synthetic bridge model. Finally, the conclusions and future work are discussed.

2   COMPILATION OF OBJECT CLASSIFICATION KNOWLEDGE MATRICES

Domain experts usually recognize an object based on its physical features and its spatial context. For example, a beam is usually a horizontal object with a longitudinally extruded cross-section (shape features), perpendicular to its supporting columns (pairwise spatial relations). Examples of shape features include centre of mass, volume, number of faces, and relative size (e.g., A is bigger than B); examples of spatial relations include contact, intersection, and orthogonal. The features and relations can be further categorized as singular features, pairwise features (e.g. relative size) and pairwise relations (e.g. spatial and topological relations).

Such knowledge can be encapsulated in a computer readable matrix structure. For example, consider a case where experts express their knowledge for identifying four types

of objects by judging whether they possess five kinds of singular features (Table 1) and four kinds of pairwise relations (Table 2). '1' signifies that this particular type of object always has this feature; '-1' signifies that it never has this feature; '0' signifies that it may or may not have the feature. Evaluation of relations between separate objects of the same type is included (e.g., a slab plate is usually in contact with another). Some pairwise relations are not commutative, such as the comparisons of objects' properties such as length and volume, so the same object pairs must appear twice in Table 2, but in different order.

Table 1: Synthetic singular feature matrix.

Row	Type	Feature 1	Feature 2	Feature 3	Feature 4	Feature 5
1	Type 1	0	0	-1	1	1
2	Type 2	-1	1	1	-1	0
3	Type 3	0	-1	-1	1	-1
4	Type 4	1	1	1	1	-1

Table 2: Synthetic pairwise relation matrix.

Row	Object Pair	Relation 1	Relation 2	Relation 3	Relation 4
1	Type 1, Type 1	1	0	0	-1
			......		
4	Type 1, Type 4	1	0	1	1
			......		
13	Type 4, Type 1	1	0	0	1
			......		
16	Type 4, Type 4	-1	0	0	0

Each row of these tables is regarded as a vector in the corresponding space; an object is identified by one vector in the singular feature space, while it is identified by $2m - 1$ vectors in the pairwise relation space if there are m different object types.

3   FEATURE MATCHING

The same singular features and pairwise relations can be evaluated for all the objects in a 3D model, which results in two fact matrices (see synthetic data in Table 3 and Table 4). Unlike the knowledge matrices (i.e. Table 1 and Table 2), the fact matrices do not have cells with '0' value or cells of self-pointed pairs, because the feature computation result is unambiguous and an object's self-relations are redundant.

A vector $\mathbf{V}$ (a row) in all these matrices (tables) can be normalized as $\hat{\mathbf{V}} = \mathbf{V}/|\mathbf{V}|$ so that it can be easily compared to another in a unit length. For example, object type 1's unit singular feature vector is $(0.00, 0.00, -0.58, 0.58, 0.58)$ and the object 1's unit singular feature vector is $(-0.45, 0.45, -0.45, 0.45, 0.45)$. The angle between them is computed by

their inner product. The greater the inner product is (i.e. smaller angle), the greater the matching degree between the object and the object type (as shown in Table 5);

Table 3: Synthetic singular feature fact matrix.

Object	Feature 1	Feature 2	Feature 3	Feature 4	Feature 5
Object 1	-1	1	-1	1	1
Object 2	-1	1	1	-1	-1
			......		
Object 11	1	1	1	1	-1

Table 4: Synthetic pairwise relation fact matrix.

Object Pair	Relation 1	Relation 2	Relation 3	Relation 4
Object1,Object2	-1	-1	1	1
		......		
Object1,Object11	1	1	1	1
		......		
Object11,Object1	-1	-1	-1	1
		......		
Object11,Object10	1	1	1	-1

Table 5: Degree of matching between objects and object types.

Object Pair	Type 1	Type 2	Type 3	Type 4
Object1	0.77	0.00	0.00	-0.20
		......		
Object11	-0.77	0.45	0.00	-0.20

The inner product of an object pair's relation vector and that of an object type pair signifies the degree of matching between this object pair and that pair of object types (results shown in Table 6). One object appears $2(n-1)$ times in this matrix if the model has n objects; one object type appears $2m-1$ times, if the model has m types of objects; the mapping between one object and one type appear in a group of $2m(n-1)$ cells. Therefore, the pair matching result can be further consolidated into object matching result by taking the average value of each group of such cells (result was shown in Table 7).

Table 6: Degree of matching between object pairs and object type pairs.

Object Pair	Type 1, 1	Type 1, 2	......	Type 4, 3	Type 4, 4
Object1,Object2	0.00	0.00		0.00	0.50
Object1,Object3	0.00	0.71		0.00	-0.50
		......			
Object11,Object9	0.00	0.00		-0.71	-0.50
Object11,Object10	0.71	-0.71		0.00	-0.50

Table 7: Degree of object matching based on pair matching result

Object Pair	Type 1	Type 2	Type 3	Type 4
Object1	0.07	0.06	0.00	0.03
		......		
Object11	0.08	-0.03	0.13	-0.05

Comparison of the results in Table 5 and Table 7 show that the results of singular feature based matching and that of pairwise based matching may be consistent (e.g. Object 1 is classified as Type 1 in both cases) or inconsistent (Object 11 is divergently classified as Type 2 and as Type 3 respectively). A method for fusing the two results is needed, but is beyond the scope of this paper.

4   CASE STUDY

The proposed approach was developed under the EU Infravation funded SeeBridge (Semantic Enrichment Engine for Bridges) project[4]. A synthetic bridge model was manually compiled for testing. It consisted of 59 bridge objects of eight different types: 1) Primary Girder; 2) Transverse Beam; 3) Deck Slab; 4) Shear Key; 5) Abutment; 6) Column; 7) Bearing; and 8) Safety Barrier, as shown in Figure 1. For compiling the knowledge matrices of such concrete girder bridges, bridge engineers suggested the use of singular features as follows:

1. Is the object's extrusion axis along the road axis?
2. Is the object's extrusion axis along the transverse axis?
3. Is the object's extrusion axis vertical?
4. Is the object convex?

and pairwise relations as follows:

1. Are the two objects in contact?
2. Are the two objects parallel?
3. Is object 1's centroid higher than that of object 2?

---

[4] SeeBridge website: http://seebridge.net.technion.ac.il/

4. Is object 1's bottom face lower than that of object 2?
5. Is object 1's length greater than that of object 2?
6. Is object 1's volume greater than that of object 2?
7. Is object 1 above object 2?
8. Is object 1 closer to the centre road axis than object 2?
9. Is object 1 closer to the centre transverse axis than object 2?
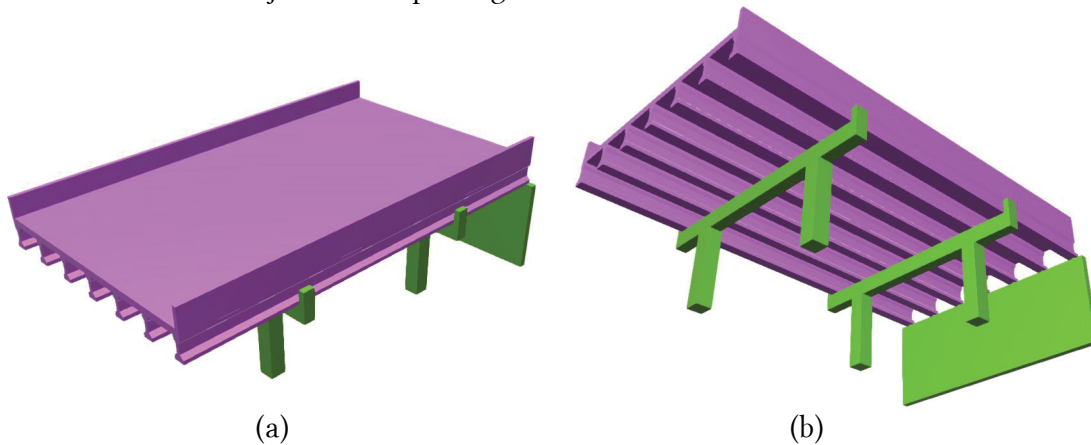10. Do the two objects overlap along the vertical axis?

(a)   (b)

Figure 1 (a) top view and (b) bottom view of the synthetic bridge model

In SeeBIM (Belsky et al., 2015), an object's axis-aligned bounding box was used as an approximate shape of it, which enlarged the object's boundary and dimensions when it is non-axis aligned, with a consequent result that many spatial topology relations would be incorrectly estimated. In this work, the above listed singular shape features were derived based on the minimal volume bounding box (Jylanki, 2015). Object volumes were computed using voxelization (Wang and Kaufman, 1993) and a tetrahedron based algorithm (Zhang and Chen, 2001). Object centroids were computed by a weighted average of its face centroids with the weights being the face areas. In addition, Query Language for 4D BIM models (QL4BIM) (Daum and Borrmann, 2013) was adopted for checking the contact and directional relations based on a ray casting and projection algorithm, which works on both convex and non-convex shapes.

Only the Girder has a unique feature vector in the singular feature space, because it is the only object type that is not convex. Other objects are indistinct, for example, the Transverse Beam and the Abutment have the same vector, and the Deck Slab and Safety Barrier have the same vector. As a result, the singular feature based matching only guarantees the correct classification of the 21 girders in the model (59 objects). In contrast, based on the pairwise relation matching, Transverse Beams, Shear Keys, Abutments, Columns and Safety Barriers are all identified correctly (100%), and 71.43% of the Bearings are correctly identified. The pairwise relation based matching serves as an important supplement to the singular feature based matching.

5   CONCLUSION

3D object classification is a fundamental aspect of semantic enrichment for BIM. Previous work using rule-based object classification depended on the geometric accuracy of the model, and was restricted in finding unique classification rules. This paper reports an approach to developing knowledge matrices that encapsulates the singular features and

pairwise relations of different object types, and a matching-based algorithm for object classification. A synthetic bridge model was used for explaining the implementation.

In the example, the results of singular shape based matching and that of pairwise relation based matching were given. However, the problem of inconsistency between the two results remains unsolved, and should be addressed in future work. In addition, the selection of distinct features and pairwise relations were crucial in the classification approach, and currently depended on experienced domain experts. Future work will attempt to use machine learning technology to establish these features and relations for object classification.

6   ACKNOWLEDGMENTS

7   REFERENCES

Aram, S. 2015. A Knowledge-based system framework for semantic enrichment and automated detailed design in the AEC projects. PhD Dissertation, Georgia Tech.

Belsky, M., Sacks, R. & Brilakis, I. 2015. Semantic Enrichment for Building Information Modeling. *Computer-Aided Civil and Infrastructure Engineering,* 31**,** 261-274.

Borrmann, A. & Rank, E. 2009. Topological analysis of 3D building models using a spatial query language. *Advanced Engineering Informatics,* 23**,** 370-385.

Bosche, F. & Haas, C. T. 2008. Automated retrieval of 3D CAD model objects in construction range images. *Automation in Construction,* 17**,** 499-512.

Brilakis, I., Lourakis, M., Sacks, R., Savarese, S., Christodoulou, S., Teizer, J. & Makhmalbaf, A. 2010. Toward automated generation of parametric BIMs based on hybrid video and laser scanning data. *Advanced Engineering Informatics,* 24**,** 456-465.

buildingSMART. 2013. *Industry Foundation Classes Release 4 (IFC4)* [Online]. Available: http://www.buildingsmart-tech.org/ifc/IFC4/final/html/ [Accessed].

Daum, S. & Borrmann, A. 2013. Definition and Implementation of Temporal Operators for a 4D Query Language. *Proc. of the ASCE International Workshop on Computing in Civil Engineering.*

Daum, S. & Borrmann, A. 2014. Processing of Topological BIM Queries using Boundary Representation Based Methods. *Advanced Engineering Informatics*, 28(4), 272-286.

Eastman, C., Lee, J. M., Jeong, Y. S. & Lee, J. K. 2009. Automatic rule-based checking of building designs. *Automation in Construction,* 18**,** 1011-1033.

Eastman, C., Teicholz, P., Sacks, R. & Liston, K. 2008. *BIM Handbook*, John Wiley & Sons, Inc.

Jylanki, J. 2015. *An Exact Algorithm for Finding Minimum Oriented Bounding Boxes* [Online]. Available: http://clb.demon.fi/projects/an-exact-algorithm-for-finding-minimum-oriented-bounding-boxes [Accessed 2015].

Marairac, W. & Beetz, J. 2013. BIMQL: An open query language for building information models. *Advanced Engineering Informatics,* 27**,** 444-456.

Pauwels, P., van Deursen, D., Verstraeten, R., de Roo, J., de Meyer, R., van de Walle, R. & van Campenhout, J. 2011. A semantic rule checking environment for building performance checking. *Automation in Construction,* 20**,** 506-518.

Tang, P., Chen, G. & Shen, Z. 2016. A Spatial-Context-Based Approach for Automated Spatial Change Analysis of Piece-Wise Linear Building Elements. *Computer-Aided Civil and Infrastructure Engineering,* 31**,** 65-80.

Tang, P., Huber, D., Akinci, B., Lipman, R. & Lytle, A. 2010. Automatic reconstruction of as-built building information models from laser-scanned point clouds: A review of related techniques. *Automation in Construction,* 19**,** 829-843.

Wang, S. W. & Kaufman, A. E. Volume sampled voxelization of geometric primitives. Visualization, 1993. Visualization'93, Proceedings., IEEE Conference on, 1993. IEEE, 78-84.

Zeibak-Shini, R., Sacks, R., Ma, L. & Filin, S. 2016. Towards generation of as-damaged BIM models using laser-scanning and as-built BIM: First estimate of as-damaged locations of reinforced concrete frame members in masonry infill structures. *Advanced Engineering Informatics,* 30**,** 312-326.

Zhang, C. & Chen, T. Efficient feature extraction for 2D/3D objects in mesh representation. International Conference on Image Processing, 2001. IEEE, 935-938.

Zhang, G., Vela, P. A., Karasev, P. & Brilakis, I. 2015. A Sparsity-Inducing Optimization-Based Algorithm for Planar Patches Extraction from Noisy Point-Cloud Data. *Computer-Aided Civil and Infrastructure Engineering,* 30**,** 85-102.