The published version is found in the ASCE Library here: https://doi.org/10.1061/9780784482421.034
Wu, J., and Zhang, J. (2019). "Introducing geometric signatures of architecture, engineering, and construction objects and a new BIM dataset." Proc., 2019 ASCE International Conference on Computing in Civil Engineering, ASCE, Reston, VA, 264-271.

Introducing Geometric Signatures of Architecture, Engineering, and Construction Objects and a New BIM Dataset

Jin Wu, S.M.ASCE,[1] and Jiansong Zhang, Ph.D., A.M.ASCE[2]

[1]Automation and Intelligent Construction Lab, Purdue University, West Lafayette, IN 47907; PH (301) 275-1272; email: wu1275@purdue.edu
[2]Automation and Intelligent Construction Lab, Purdue University, West Lafayette, IN 47907; PH (765) 494-1574; FAX (765) 496-2246; email: zhan3062@purdue.edu

ABSTRACT

Object signatures have been widely used in object detection and classification. Following a similar idea, the authors developed geometric signatures for architecture, engineering, and construction (AEC) objects such as footings, slabs, walls, beams, and columns. The signatures were developed both scientifically and empirically, by following a data-driven approach based on analysis of collected building information modeling (BIM) data using geometric theories. Rigorous geometric properties and statistical information were included in the developed geometric signatures. To enable an open access to BIM data using these signatures, the authors also initiated a BIM data repository with a preliminary collection of AEC objects and their geometric signatures. The developed geometric signatures were preliminarily tested by a small object classification experiment where 389 object instances from an architectural model were used. A rule-based algorithm developed using all parameter values of 14 features from the geometric signatures of the objects successfully classified 336 object instances into the correct categories of beams, columns, slabs, and walls. This higher than 85% accuracy showed the developed geometric signatures are promising. The collected and processed data were deposited into the Purdue University Research Repository (PURR) for sharing.

INTRODUCTION

Object signatures have been widely used in object detection and classification. For example, Wang and Zhang (2010) used decomposed spatial temporal-signatures to enable dynamic 3D objects perception from digital images by computers. Tu et al. (2017) proposed a fusion method to combine different object signatures such as appearance and motion cues for salient object detection from videos. In a geometric study, He and Peng (2011) proposed a shape retrieval system using geometric signatures that are invariant under translation. Hoff and Olver (2012) proposed further extensions of these invariant signatures with a compromise between local and global identifying properties to enable the recognition of non-congruent curves. Both versions of geometric signatures can be used to support the detection and recognition of real-world objects.

Following a similar idea with the above works, the authors developed geometric signatures for architecture, engineering, and construction (AEC) objects such as footings, slabs, walls, beams, and columns. The signatures were developed both scientifically and empirically, by following a data-driven approach based on analysis of building information modeling (BIM) data using geometric theories. These signatures include features of two types: (1) rigorous geometric properties such as cross-sectional profile, extrusion direction, dimensional ratios, boundary line connection angles, lengths, and turn directions, and (2) statistical information such as number of sub-components, number of faces, and the number of straight lines and curves. Because geometric information usually takes a large portion of BIM data (Zhang 2018), these signatures provide robust information of AEC objects that can be used to support engineering and management analysis such as quantity takeoff and structural analysis, in order to support seamless and universal interoperability of BIM. To enable open access to BIM data using these signatures, the authors also initiated a BIM data repository with a preliminary collection of AEC objects and their geometric signatures.

BACKGROUND

Neural signatures. The structure and functions of human brains are still underexplored. But the way that a human brain detects objects gradually got unraveled, i.e., patterns consisted of features help people recognize objects (Brandman and Peelen 2017). For example, Johnson and Olshausen (2003) observed this object detection process by a human brain through experiments. In their experiment, two ways to measure event-related potential (ERP), i.e., the electrophysiological response to a stimulus, were used to correlate brain activities with object recognition. Two types of components in an ERP of natural images were discovered: early presentation-locked signal and later recognition-related component, respectively. An early presentation-locked signal indicates low-level feature differences between images. A later recognition-related component covariates with the subsequent reaction time. Their experiment inferred that the second type of neural signatures for image recognition have a substantially later and variable time of onset comparing to the first type, which provides insights into object detection by human brains using neural signatures.

Mathematical signatures. Compared to neural signatures described above, mathematical signatures follow a more concise and rigid formulation. The creation of mathematical signatures followed a rigorous procedure of definition and proof, based on element axioms and complicated deductions. For example, the mathematical signature of a circle includes the following two rules depicting the patterns of features: (1) the set of points that forms the circle must be in the same plane; (2) the boundary (circumference) must be equidistant to a center point (Coolidge 1902). Furthermore, Daniyarova et al. (2012) showed that the entire properties of algebraic universal geometry can be carried over to the case of an arbitrary geometric signature without essential changes.

Other signatures. The idea of signatures is widely used and has varieties of different nature. For example, Stow et al. (2012) proposed frequency distribution signatures for use in the classification of within-object pixels. Nelson and Sokkappa (2008) proposed radiation signatures generated using a statistical model to detect nuclear threats. Marat and Ltti (1996) established object signatures for object classification and showed that the amount of context learned had an important effect in object recognition results.

The published version is found in the ASCE Library here: https://doi.org/10.1061/9780784482421.034
Wu, J., and Zhang, J. (2019). "Introducing geometric signatures of architecture, engineering, and construction objects and a new BIM dataset." Proc., 2019 ASCE International Conference on Computing in Civil Engineering, ASCE, Reston, VA, 264-271.

Shared data for research and development. A Shared dataset (e.g., ImageNet, Flickr) not only provides people with resources for research, but also enhances the synergistic effect of research efforts from different teams by providing a common ground for comparison and discussion. It has been a common practice in computer science domains and helped advance research discoveries and technology development. For example, in the computer vision domain, Guillaumin et al. (2014) developed automated annotation methods for images using ImageNet. Yin et al. (2009) explored social tagging graph-based web object classification using Flickr. In the natural language processing domain, Reid et al. (2018) developed social science interpretation methods based on decompounded lexicon induction technics, through the use of a Consumer Complaint Database in their development.

Open BIM repository. A few data repositories exist in the architecture, engineering, and construction (AEC) domain. For example, Dimitrov and Golparvar-Fard (2014) established the Construction Materials Library that contains 3,000 material samples that were collected from 5 construction sites and 2 existing buildings. Varying degrees of viewpoint, scale, and illumination were recorded during the collection period spanning seven months. For IFC data, the "Open IFC Model Repository" (Dimyadi and Henderson 2012) provides 105 models of building elements. The NBS National BIM Library (2018) provides 6,660 IFC data instances. While useful for research and development, these data repositories mostly provide data models without detailed analysis. In comparison, the authors created a new IFC data repository which provides IFC data at the object level that were processed with their geometric signatures. These data can be used for various analyses and developments such as AEC object classification to detect misuse of IFC entities in BIM.

PROPOSED GEOMETRIC SIGNATURES OF AEC OBJECTS

IFC-based BIM data extraction. IFC data follows the EXPRESS specification (BuildingSmart 2018). Information in IFC is stored in entities and attributes, including numeric values and relationships that refer entities to one another, forming a hierarchical structure of interrelated objects. To extract information from IFC data using the cross-references between entities, Won et al. (2013) proposed an IFC data extraction algorithm that can extract targeted parts of an IFC model, e.g., extracting a slab from a building or a pier from a bridge. The authors used a similar algorithm to extract objects from IFC models (referred to as AEC objects hereafter) and store each extracted object into one stand-alone IFC file. The authors collected data from two different sources: (1) the "Common Building Information Model Files" published by buildingSMARTalliance of the National Institute of Building Sciences (East 2013), and (2) Revit models exported as IFC data files. The data includes models of a duplex apartment, three architecture projects, and a technical school building.

Geometric signatures of AEC objects. The authors analyzed the contents and structures of the extracted AEC objects by tracing their 3D geometric representations in IFC such as "swept solid" and "boundary representation (Brep) bodies" and found two sets of properties to describe the objects' geometries. The first set includes geometric properties that come from mathematical definitions and geometric theorems. The second set includes statistical measures of components in an object's geometry. Combining both sets, the authors proposed a systematic set of features as

The published version is found in the ASCE Library here: https://doi.org/10.1061/9780784482421.034
Wu, J., and Zhang, J. (2019). "Introducing geometric signatures of architecture, engineering, and construction objects and a new BIM dataset." Proc., 2019 ASCE International Conference on Computing in Civil Engineering, ASCE, Reston, VA, 264-271.
geometric signatures of AEC objects, including number of sub-components, number of faces, cross-sectional profile, extrusion direction, dimensional ratios, number of straight lines and curves, boundary line connection angles, lengths, and turn directions.

The most frequently used shape in our collected objects is rectangular parallelepiped. Table 1 summaries the features of walls, slabs, footings, columns, and beams in this shape. The values of number of sub-components (NoSC), number of faces (NoF), cross-sectional profile (CSP), number of straight lines and curves (NoSLC), boundary line connection angles (BLCA), lengths (BLCL), and turn directions (BLCTD) are the same across these five types of objects. For extrusion direction (ED), beams have the value of "horizontal" where all other four types have the value of "vertical." Dimensional ratios (DR) are mainly distinguishing different types of objects from each other in this case.

Table 1. Geometric signatures of AEC objects in a rectangular parallelepiped shape.

Feature		Value range for each object type				
		Wall	**Slab**	**Footing**	**Column**	**Beam**
Number of sub-components (NoSC)		1	1	1	1	1
Number of faces (NoF)		6	6	6	6	6
Cross-sectional profile (CSP)		Rectangle				
Extrusion direction (ED)		Vertical	Vertical	Vertical	Vertical	Horizontal
Dimensional ratios (DR) H: height W: width L: length	L:H	[0.1228, 99.3807]	[11.6545, 3000.1260]	[1.5000, 3.8000]	[0.1500, 0.4444]	[0.0091, 0.1655[1]]
	W:H	[0.0175, 0.6847]	[5.0000, 171.5385]	[0.6667, 3.8000]	[0.1500, 0.4444]	[0.0046, 0.0587]
	L:W	[3.2143, 294.6825]	[1.0683, 40.1069]	[1.0000, 2.2500]	[1.0000, 1.0000]	[1.0000, 4.000]
Number of straight lines and curves (NoSLC)		12 straight lines				
Boundary line connection angle (BLCA)		90 degrees				
Boundary line length type (BLCL)	Mathematical (theory)	Three lengths. Each length has four lines of that length.				
	IFC (swept solid)	For swept solid, the 2D shape has two lengths: width and length.				
Boundary line connection turn direction (BLCTD)		90 degrees with the same direction (all right turns or all left turns)				

[1] Height is the extrusion depth, which is horizontally aligned.

The second most frequently used shape in the collected data is the cylinder. Table 2 shows feature values in the geometric signatures of cylinder-shaped footings, columns, and beams. Similar to objects in a rectangular parallelepiped shape, objects in a cylinder shape have the same values for all features except for extrusion direction and dimensional ratios. In addition, ring-shaped and I-shaped objects were also processed, the feature values in the geometric signatures of which are summarized in Table 3 and Table 4, respectively. A typical I-shape cross-sectional

The published version is found in the ASCE Library here: https://doi.org/10.1061/9780784482421.034
Wu, J., and Zhang, J. (2019). "Introducing geometric signatures of architecture, engineering, and construction objects and a new BIM dataset." Proc., 2019 ASCE International Conference on Computing in Civil Engineering, ASCE, Reston, VA, 264-271.
profile and a typical ring shape cross-sectional profile are illustrated in Figures 1 and 2, respectively.

Table 2. Geometric signatures of AEC objects in a cylinder shape.

Feature	Value range for each object type		
	Footing	**Column**	**Beam**
NoSC	1	1	1
NoF	3	3	3
Face type (FT)	2 circles, one curved rectangle	2 circles, one curved rectangle	2 circles, one curved rectangle
CSP	Circle	Circle	Circle
ED	Vertical	Vertical	Horizontal
DR (R:H) R: radius H: height	[0.0250, 0.0417]	[0.0395, 0.1230]	[0.0043, 0.0047]
NoSLC	2 curves: 2 circles	2 curves: 2 circles	2 curves: 2 circles
BLCA	90 degrees	90 degrees	90 degrees
BLCTD	90 degrees with the same direction (all right turns or all left turns)		

Table 3. Geometric signature of AEC objects in a ring shape.

Feature	NoSC	NoF	FT	CSP	ED	DR[1]		NoSLC	BLCA	BLCTD
Value range	1	4	2 rings, two curved rectangles	Ring	Horizontal	R:H	[0.0124, 0.0157]	2 curves and 2 circles	90 degrees	90 degrees with the same direction (all right turns or all left turns)
						T:H	[0.0017, 0.0018]			

[1] H: height, R: radius, T: thickness.

Table 4. Geometric signature of AEC objects in an I-Shape.

Feature	NoSC	NoF	FT	CSP	ED	DR[1]		NoLC	Con	Turn
Value range	1	14	12 rectangles, 2 I-shaped faces	I-shape	Horizontal	OD:ED	[0.0842, 0.1023]	2 curves and 2 circles	90 degrees	90 degrees with the same direction (all right turns or
						OW:ED	[0.0432, 0.0988]			
						R:ED	[0.0031, 0.0042]			

The published version is found in the ASCE Library here: https://doi.org/10.1061/9780784482421.034
Wu, J., and Zhang, J. (2019). "Introducing geometric signatures of architecture, engineering, and construction objects and a new BIM dataset." Proc., 2019 ASCE International Conference on Computing in Civil Engineering, ASCE, Reston, VA, 264-271.

	FT:ED	[0.0031, 0.0029]	all left turns)
	WT:ED	[0.0049, 0.0165]	

[1] R: radius, FT: flange thickness, WT: web thickness, OD: overall depth, OW: overall width, ED: extrusion depth.

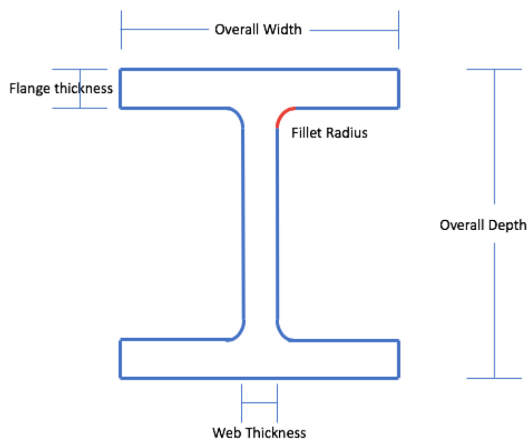

Figure 1. Cross-sectional profile of an I-shape.

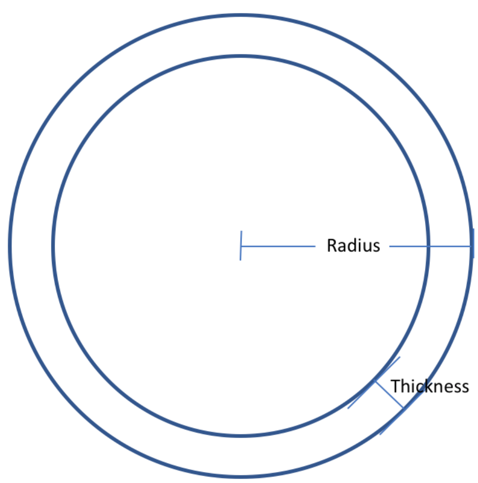

Figure 2. Cross-sectional profile of a ring shape.

DEVELOPED BIM DATA REPOSITORY

After the geometric signatures were developed, the authors extracted features of 1,252 object instances from the collected data. Figure 3 shows a part of the processed data, including 14 features of 14 object instances among 1,252 object instances of regular shapes.

The published version is found in the ASCE Library here: https://doi.org/10.1061/9780784482421.034
Wu, J., and Zhang, J. (2019). "Introducing geometric signatures of architecture, engineering, and construction objects and a new BIM dataset." Proc., 2019 ASCE International Conference on Computing in Civil Engineering, ASCE, Reston, VA, 264-271.

In addition, to show an example use of the data set, the authors developed a rule-based algorithm for object classification, which was tested on 389 object instances from an architectural model. The algorithm used all parameter values from the 14 features and successfully classified 336 object instances into the correct categories of beams, columns, footings, slabs, and walls, resulting in an 88.67% accuracy (Table 5). The collected and processed data were deposited into the Purdue University Research Repository (PURR) (Wu and Zhang 2018).

Table 5. Classification results of object instances in an architectural model

Object type	Correctly classified	Ground truth	Accuracy
Beam	12	37	32.43%
Column	170	176	95.59%
Footing	0	0	Na
Slab	14	30	46.67%
Wall	140	146	95.89%
Overall	336	389	86.38%

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Stepline	Name	Type	Rec_L	Rec_W	Rec_H	Cir_R	Cir_H	R_R	R_T	I_W	I_H	I_D	I_R	I_WT	I_FT	Done
2	23286	#23286 IfcFooting	Footing	18.283	0.9	0.3											Done
3	23369	#23369 IfcFooting	Footing	8.383	0.9	0.3											Done
4	23408	#23408 IfcFooting	Footing	17.383	0.9	0.3											Done
5	23446	#23446 IfcFooting	Footing	7.483	0.9	0.3											Done
6	23485	#23485 IfcFooting	Footing	4.1915	0.9	0.3											Done
7	23524	#23524 IfcFooting	Footing	2.2	0.9	0.3											Done
8	36892	#36892 IfcBeam	Beam			6.18189					0.178	0.347	6.18189	0.0128	0.0128	0.06795	Done
9	37456	#37456 IfcBeam	Beam			6.18189					0.178	0.347	6.18189	0.0128	0.0128	0.06795	Done
10	96748	#96748 IfcSlab	Slab				150	6002									Done
11	96819	#96819 IfcSlab	Slab				150	6002									Done
12	96879	#96879 IfcSlab	Slab				150	6002									Done
13	96939	#96939 IfcSlab	Slab				150	6002									Done
14	96999	#96999 IfcSlab	Slab				150	6002									Done

Figure 3. A snapshot of the developed dataset (partial).

CONCLUSION

The authors developed AEC object geometric signatures for regular shapes of footings, slabs, walls, beams, and columns, including rectangular parallelepiped, cylinder, ring, and I-shape. These geometric signatures were developed in a scientific and empirical manner following a data-driven approach. Mathematical definitions, geometric theorems, and statistical counts of components were used in the signatures. The developed signatures provide useful information for describing the characteristics of AEC objects' geometries. The authors collected a set of 1,252 AEC object instances and processed their geometric signatures. The dataset was shared through the Purdue University Research Repository (PURR) to provide analyzed AEC objects data as a benchmark and common ground for establishing shared tasks in future BIM research.

ACKNOWLEDGEMENTS

The published version is found in the ASCE Library here: https://doi.org/10.1061/9780784482421.034
Wu, J., and Zhang, J. (2019). "Introducing geometric signatures of architecture, engineering, and construction objects and a new BIM dataset." Proc., 2019 ASCE International Conference on Computing in Civil Engineering, ASCE, Reston, VA, 264-271.
The author would like to thank the National Science Foundation (NSF). This material is based on work supported by the NSF under Grant No. 1745374. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the NSF.

REFERENCES

Brandman, T. and Peelen, M. (2017). "Interaction between Scene and Object Processing Revealed by Human fMRI and MEG Decoding." *J. Neuroscience*., 37(32), 7700-7710.

BuildingSmart. (2018) "International Home of OpenBIM. " <http://www.buildingsmart-tech.org/> (Nov. 18, 2018)

Coolidge, J.L. (1902) "Euclid's elements." *Bulletin of the American Mathematical Society*, 8(5), 216-220.

Daniyarova, E., Myasnikov, A., and Remeslennikov, V. (2012). "Algebraic geometry over algebraic structures. V. The case of arbitrary signature." *Algebra and Logic*, 51(1), 28-40.

Dimitrov, A. and Golparvar-Fard, M. (2014). "Vision-based material recognition for automated monitoring of construction progress and generating building information modeling from unordered site image collections." *Advanced Engineering Informatics*, 28(1), 37-49.

Dimyadi, J., and Henderson, S. (2012). Open IFC Model Repository. < http://openifcmodel.cs.auckland.ac.nz/> (Nov 18, 2018).

East, E.W. (2013). "Common building information model files and tools." < https://www.nibs.org/page/bsa_commonbimfiles?> (Nov. 18, 2018).

He, J., and Peng, Y. (2011). "A Shape Retrieval Study Based on Geometric Signature." *Advanced Materials Research*, 411, 597-601.

Hoff, D. and Olver, P. (2013). "Extensions of Invariant Signatures for Object Recognition." *J. Mathematical Imaging and Vision*, 45(2), 176-185.

Guillaumin, M., Küttel, D., Ferrari, V. (2014). "ImageNet Auto-Annotation with Segmentation Propagation." *International Journal of Computer Vision*, 110(3), pp.328-348.

Johnson, J. and Olshausen, B. (2003). "Timecourse of neural signatures of object recognition." J. *Vision*, 3(7), 499-512.

Marat and Ltti (1996). "Influence of the amount of context learned for improving object classification when simultaneously learning object and contextual cues." *Visual Cognition*, 20(4-5), 580-602.

NBS of UK. (2014). "NBS National BIM Library." <http://www.nationalbimlibrary.com/> (Nov. 18, 2018).

Nelson, K. and Sokkappa, P. (2008). "A Statistical Model for Generating a Population of Unclassified Objects and Radiation Signatures Spanning Nuclear Threats." *International Nuclear Information System*, 40(17).

Stow, D.A., Toure, S.I., Lippitt, C.D., Lippitt, C. L., and Lee, C. (2012). "Frequency distribution signatures and classification of within-object pixels." *International J. Applied Earth Observations and Geoinformation*, 15, 49-56.

Tu, Z., Guo, Z., Xie, W., Yan, M., Veltkamp, R.C., Li, B., and Yuan, J. (2017) "Fusing disparate object signatures for salient object detection in video." *Pattern Recognition*, 72, 285-299.

Wang, Y., and Zhang, K. (2010) "Decomposing the spatiotemporal signature in dynamic 3D object recognition*." J. Vision*, 10(10), 2.

The published version is found in the ASCE Library here: https://doi.org/10.1061/9780784482421.034
Wu, J., and Zhang, J. (2019). “Introducing geometric signatures of architecture, engineering, and construction objects and a new BIM dataset.” Proc., 2019 ASCE International Conference on Computing in Civil Engineering, ASCE, Reston, VA, 264-271.

Won, J., Lee, G., and Cho, C. (2013). “No-schema algorithm for extracting a partial model from an IFC instance model.” *J. Comput. Civ. Eng.*, 27(6), 585-592.

Wu, J. and Zhang, J. (2018), “BIM database with invariant features.” (DOI: 10.4231/R7MG7MSG). <https://purr.purdue.edu/projects/bimdatareposit/databases/> (Dec. 1, 2018).

Yin, Z., Li, R., Mei, Q., and Han, J. (2009). “Exploring social tagging graph for web object classification.” *Proc., 15th ACM SIGKDD inter. conf. on knowledge discovery and data mining,* ACM, New York, 957-966.

Zhang, J. (2018). “Towards systematic understanding of geometric representations in BIM standard: an empirical data-driven approach.” *Proc., ASCE Construction Research Congress*, ASCE, Reston, VA, 96-105.