

CLOI: A Shape Classification Benchmark Dataset for Industrial Facilities

Eva Agapaki, S.M. ASCE,¹ Alex Glyn-Davies,^{2*} Sara Mandoki,^{3*} and
Ioannis Brilakis, Ph.D., M.ASCE⁴

¹Ph.D. Candidate, Department of Engineering, University of Cambridge, CB2 1PZ, U.K.; e-mail: ea437@cam.ac.uk

²Undergraduate student, Department of Engineering, University of Cambridge, CB2 1PZ, U.K.; e-mail: ag933@cam.ac.uk

³ Undergraduate student, Department of Engineering, University of Cambridge, CB2 1PZ, U.K.; e-mail: sm2287@cam.ac.uk

⁴ Laing O'Rourke Reader, Department of Engineering, University of Cambridge, CB2 1PZ, U.K.; e-mail: ib340@cam.ac.uk

*equal contribution

ABSTRACT

Generation of digital models of existing industrial facilities is labor intensive and expensive. The use of state-of-the-art deep learning algorithms can assist to reduce the modelling time and cost. However large databases of labelled, laser-scanned industrial facilities do not exist to date, henceforth training of deep learning models is not possible. Our paper solves this problem by proposing a new benchmark dataset, which consists of five labelled industrial plants. The labelling schema that we followed for the generation of this dataset is based on the frequency of appearance of industrial object types. We labelled the ten most frequent industrial object shapes as identified in previous work. We present CLOI (Channels, L-shapes, circular sections, I-shapes): a richly annotated large-scale repository of shapes represented by labelled point clusters. CLOI has more than 140 million hand labelled points and serves as the foundation for researchers who are interested in automated modelling of industrial assets using deep learning algorithms.

INTRODUCTION

Operation phase Digital Twins (DTs) are digital replicas of real-world assets (Boschert et al., 2018; Parrott and Lane, 2017). These assets can be tunnels, buildings, bridges or any other man-made assets of the built environment. A DT contains all life-cycle information of the physical asset. Geometric modelling is the “bottleneck” during the generation of DTs for any asset given how costly and time consuming it is. Geometric modelling involves the following steps: (a) primitive shape detection, (b) shape classification and (c) fitting. The modelling time for (b) is significant especially for cluttered environments with many object categories like industrial plants (Agapaki et al., 2018).

Improving the effectiveness of shape classification algorithms that take as input 3D Lidar data remains a challenge towards high level scene understanding solutions for built environments.

Given the large amounts of data that assets of the built environment constitute, there is a strong need for data-hungry (deep) learning methods that are tailored for 3D data. However, there are significant challenges to be solved to efficiently feed the deep learning algorithms with reliable data for shape classification. One of those is the generation of large benchmark datasets for training classification algorithms.

Manual extraction of thousands of point clusters from point clouds for a multi-classification algorithm is a tedious process that prohibits training for the application of deep learning algorithms. The collected point clouds need to be annotated for supervised learning multi-classifiers. A technique of hand labelling through crowdsourcing has emerged for images (Silberman et al., 2012; Song et al., 2015). For this purpose, crowdsourcing platforms like Amazon Mechanical Turk (Amazon Mechanical Turk, 2018) or LabelMe (Russell et al., 2008) have been developed. However, it is more difficult to accomplish this task for point clouds due to the varying density and difficulty to interpret cluttered 3D scenes for untrained users.

Industrial scenes are a significant example of complex scenes with thousands of object categories that make hand-labelling even more time-consuming. We focused on determining the most labor intensive industrial object shapes to model in our previous work (Agapaki et al., 2018). These are in descending order of labor intensiveness: electrical conduit, straight pipes, circular hollow sections (CHSs), elbows, channels, solid bars, I-beams, angles, flanges and valves. We introduce a new point cloud database called “CLOI”. We provide the (to the best of our knowledge) hitherto largest collection of terrestrial laser scans of industrial plants with point-level semantic ground truth annotation. In total, it consists of 140 million points and class labels for each point.

Table 1 shows the total number of labelling hours needed for the generation of CLOI and total number of shapes annotated by each researcher. This means that around 4 min/shape were needed for hand-labelling.

Table 1 Total number of labor hours, total number of shapes and average labelling time per shape

	Researcher 1	Researcher 2	Researcher 3
Total number of hours	267.5	162.5	233
Total number of shapes	3746	1780	4869
Average minutes/shape		4.2	

It is easily distinguishable that the time needed for manual classification of each point cluster in large point cloud datasets is a time-consuming procedure. This necessitates the need for automated shape classification. This paper is the first to generate a benchmark labelled dataset for industrial facilities that will enable the use of supervised deep learning algorithms. We first describe our labelling interface and generation of CLOI. We then show that CLOI can serve as a useful resource for industrial shape classification and detection applications with much higher overall number of labelled points compared to those already available to the research community for other applications.

BACKGROUND

State of research. Shape classification has been solved for the case of images (Krizhevsky et al., 2012; Russakovsky et al., 2015). However, different from images that have representations as 2D

pixel arrays, point clouds have various 3D representations and varying point densities. Neural network approaches can learn in arbitrarily complex domains but typically rely on simulated datasets that reach the order of a few thousands so far (C. R. Qi et al., 2017(a); R. Qi et al., 2017(b); Shilane et al., 2004) or datasets collected from the web (Song et al., 2015; Xu et al., 2015). A recent example, SynthCam3D, gathers a library of synthetic indoor scenes collected from various online 3D repositories of offices and indoor building scenes (Handa et al., 2015). Often, this is unrealistic in real-world data as, in general, data collection is time-consuming, inaccessible especially due to hazardous or radioactive chemicals in industrial environments that even drones may not be able to capture. Benchmark datasets of large-scale indoor spaces are limited to buildings and offices (Armeni et al., 2016; Dai et al., 2017). The former has around 10,000 labelled point clusters of indoor objects whereas the latter consists of more than 1,500 scans of indoor rooms.

Benchmark datasets of outdoor scenes are widely needed in the geospatial data community. Segmented and classified point clouds are used for vehicle localization (Maddern et al., 2015) and navigation (Chu et al., 2017). Two densely annotated point clouds of urban outdoor scenes are semantic3D.net (Hackel et al., 2017) and Paris-Lille-3D (Roynard et al., 2018). The former consists of dense point clouds of a wide range of outdoor scenes (churches, streets, railroad tracks, squares, villages, soccer fields and castles) whereas the latter is composed of point clouds of outdoor scenes in Paris and Lille with more than 50 classes.

Gaps in knowledge and research questions. Considering the state of practice and body of research reviewed above, existing benchmark datasets are used for classification of indoor and outdoor scenes; however no benchmark dataset exists for industrial plants. It is therefore still unclear how to use deep learning algorithms without the existence of an annotated dataset for these environments. The aim of this work is to solve this gap in knowledge by answering the following research questions: (a) how can we close the data gap to help unleash the full potential of deep learning methods for 3D labelling tasks on point clouds of industrial plants and (b) how can this annotated dataset be useful for other applications?

RESEARCH METHODOLOGY

The research described in this paper focuses on the generation of the first publicly available dataset of labelled point clusters of industrial plants. We describe the generation of CLOI in three basic steps: (a) data preparation for registered point cloud generation, (b) labelled point cluster extraction and (c) refinement of labelled clusters.

Data collection. CLOI consists of 10 classes that cover a wide range of industrial scenes (both indoor and outdoor). Five laser scanned industrial facilities were used for the generation of CLOI as shown in Figure 1. Three facilities were warehouses, one was a petrochemical plant and the fifth an oil processing refinery. These facilities are anonymized since rights are reserved by AVEVA Group Plc. and British Petroleum. All datasets were obtained using static terrestrial laser scanners. The scanner setup and scan frequency of these facilities is not available, since data was collected by industrial partners.

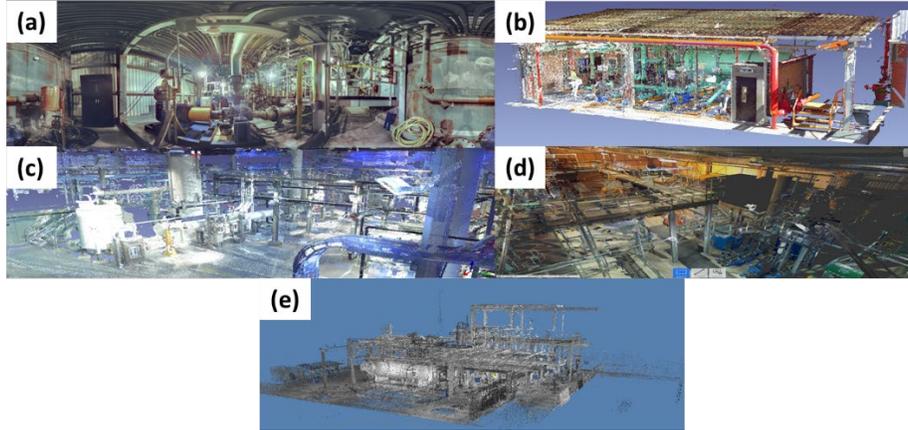


Figure 1 (a) Oil refinery, (b) warehouse 1, (b) warehouse 2, (d) warehouse 3 and (e) processing unit.

Assumptions. CLOI consists of the most important industrial shapes as identified in Agapaki et al., 2018. We provide the definitions of these shapes below. We define *angles* as roll-formed steel angles that have “L” shape. The legs of the “L” shape have equal or unequal length. *Channels* refer mostly to steel beams with C-shape. *CHSs* refer to cylinders that support pipes, cylindrical structural columns and handrails. *Conduit* refers to the tubes that protect electric wiring. *Elbows* are tubes that have 90° or 45° angle. *Flanges* refer to plates or rings at the end of pipes. We define *I-beams* as the structural steel beams that have I-shape. *Pipes* are tubes that carry liquids and gases. *Valves* refer to all the devices that control the flow of liquids through the pipelines. We cover all types of valves across our datasets (globe, ball, gate, butterfly, diaphragm, plug, check, needle, pinch valve). *Other* refers to any other point clusters that do not belong in the above-mentioned classes.

Data preparation. The first step in our pipeline is to prepare and register the laser scanned point clouds, so that they can be used in the commercial, manual labelling platform for industrial plants, LFM (“LFM Software,” 2016). The organogram of the dataset is presented in Figure 2. The source data of our laser scanned facilities are scans, panoramic images and Bubble Views (photorealistic views) generated in LFM. We first manually removed outliers (vegetation, reflectance outside windows) and registered the laser scans using Cloud to Cloud registration in LFM Gateway. LFM Gateway is then used to generate a dataset from all the source registered scans. The generation process involves to spatially index all the returns in the registered scans into a single dataset [*lfd and points*points] along with the scan positions. We then used LFM Server to access the single dataset to access the scanner returns to define the point clusters, removing the need to define clusters for each scan.

Labelled point cluster extraction. The researchers manually defined bounding volumes of the 10 most important industrial shapes and extracted the points inside these volumes in ASCII format. An example of a user selected bounding volume and the extracted shape is shown in Figure 3. The labelled clusters are organized in separate folders based on the shape that they belong to and can be directly used for training supervised learning algorithms.

Refinement of labelled clusters. The researchers inspected the labelled clusters, removed outliers manually and refined the clusters by using the Statistical Outlier Removal (SOR) filter (Rusu and

Cousins, 2011). There were cases where the rectangular bounding volumes could not capture irregular or adjacent shapes. For these cases, the researchers used CloudCompare (“Cloudcompare,” 2016) to refine the selection of points by defining polygons around the object of interest. The extracted point clusters contain XYZ and intensity values. The labelled point clusters are then demolished, the laser returns are marked as demolished [*points] and the geometric definition of the demolition is stored in the [demolition.sqlite] files (Figure 2).

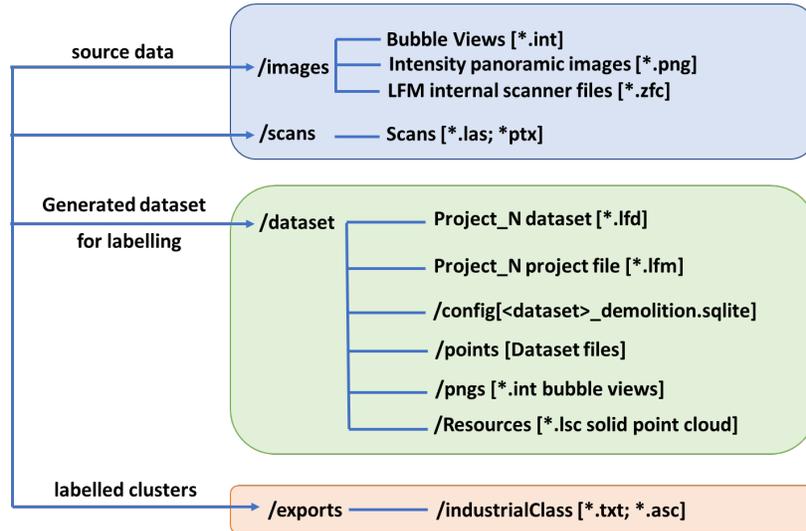


Figure 2 Organogram of CLOI

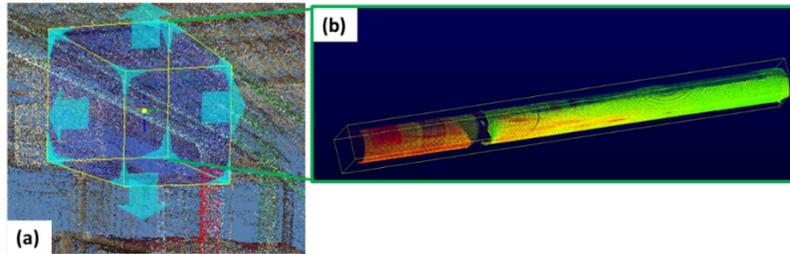


Figure 3 (a) Example of user selected bounding volume in LFM Server and (b) corresponding labelled point cluster of a pipe

RESULTS

CLOI consists of 12,125 shapes distributed in 10 classes. The frequency of appearance of each class across the five industrial facilities is shown in Figure 4(a) and the distribution of shapes in CLOI is shown in Figure 4(b). We observe that there is variation in the frequency of appearance of channels, conduit and CHSs (~10-25%) across different datasets. This is attributed to the different use of each industrial plant. We also compare the distribution of shapes in CLOI with the average distribution of the same shapes as identified in our previous work (Agapaki et al., 2018). We observe that the percentages of all shapes in CLOI align with those in our previous work except those of elbows (15% in CLOI and 5% in Agapaki et al., 2018) and other (12.5% in CLOI and 5% in Agapaki et al., 2018). These disparities can be attributed to the diversity of industrial scenes.

Comparison with existing benchmarks. We compare the volume of data of the largest existing benchmark datasets to CLOI. We observe that the total number of labelled clusters in CLOI is an order of magnitude larger compared to the S3DS (indoor scenes), Paris-Lille-3D and Semantic3D.net (outdoor scenes) as shown in Table 2.

Table 2 Comparison of CLOI with other benchmark datasets

Dataset	Total number of labelled clusters
S3DS	5,862
Paris-Lille-3D (Roynard et al., 2018)	2,479
Semantic3D.NET	~1,000
CLOI	12,125

The quality of manual classification and quantity of data in the CLOI dataset will enable deep learning application on the labelled clusters. The choice of many different data sources was based not only on detecting numerous shape classes present in parts of an industrial plant but also in different industrial environments. Each industrial facility serves as a training base for tests on the others. The classification benchmark dataset is released with this publication and can be found in this link: <https://drive.google.com/drive/folders/1Nale4OK6BCdAZ94w1BVdBokW2wfgwS1>¹.

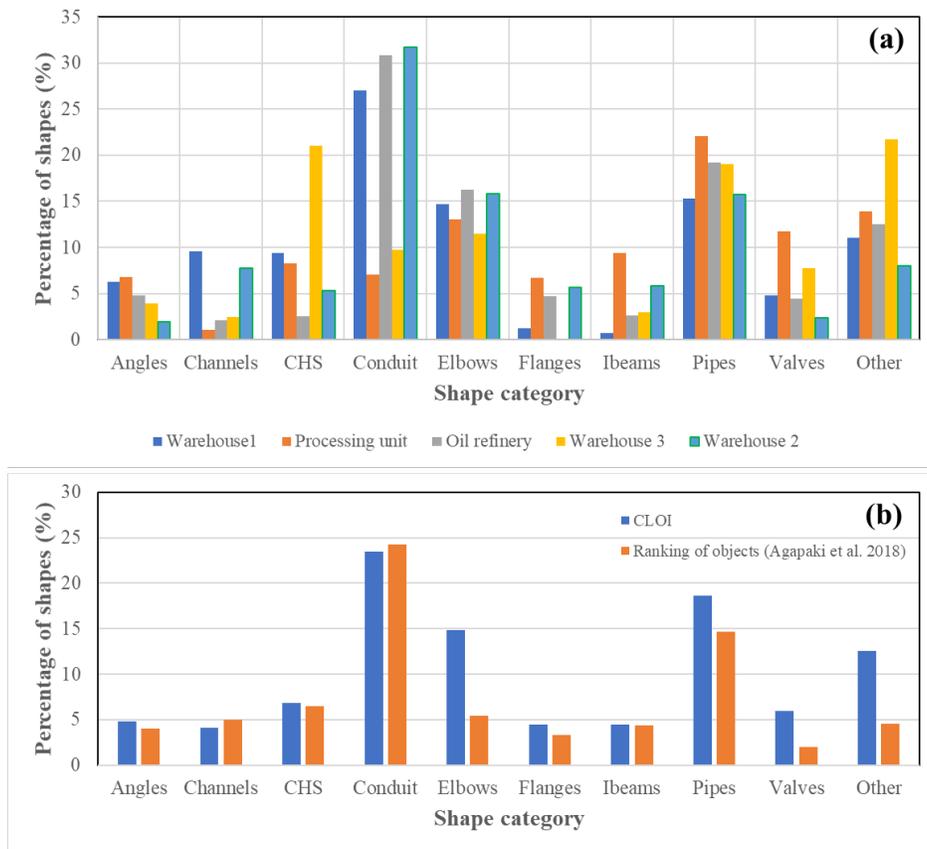


Figure 4 Frequency of appearance of the 10 industrial shapes of CLOI

¹ DOI released after publication: 10.6084/m9.figshare.7355141

CONCLUSION

In this paper we presented the CLOI dataset and the interest of having this benchmark dataset for the scientific community to improve the techniques of automated classification of industrial scenes. Important characteristics of our benchmark are high quality of segmentation and classification (the annotation of each point was done manually), the wide variety of most important industrial object classes and the large volumes of labelled clusters that makes it useful for deep learning applications. It is important to note that CLOI is the largest annotated dataset compared to its predecessors in other applications and the first of its kind on point clouds of industrial plants. The use of this dataset is not only limited on industrial shape classification, but also on definitions of contextual relationships between the classified objects and object detection. Some other applications of CLOI include generation of the connectivity relationships of point clusters to create digital models of these facilities and image classification of the panoramic images taken from the laser scanners based on labels per point. Researchers can also use the “*Other*” object class to annotate other industrial shapes that they are interested in. We hope that CLOI will pave the way for deep learning methods in these applications.

ACKNOWLEDGEMENTS

We thank our colleague Graham Miatt, Senior Software Engineer in LFM Software from AVEVA Group Plc., who provided insight and expertise that greatly assisted this research. The research leading to these results has received funding from the Engineering and Physical Sciences Research Council (EPSRC). AVEVA Group Plc. and BP International Centre for Business and Technology (ICBT) partially sponsor this research under grant agreements RG83104 and RG90532 respectively. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the institutes above.

REFERENCES

- Agapaki, E., Miatt, G., Brilakis, I., 2018. Prioritizing object types for modelling existing industrial facilities. *Autom. Constr.* <https://doi.org/10.1016/j.autcon.2018.09.011>
- Amazon Mechanical Turk, 2018. Amazon Mechanical Turk [WWW Document]. URL <https://www.mturk.com/>
- Armeni, I., Sener, O., Jiang, H., Fischer, M., Savarese, S., 2016. 3D Semantic Parsing of Large-Scale Indoor Spaces, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 1534–1543.
- Boschert, S., Heinrich, C., Rosen, R., 2018. Next Generation Digital Twin, in: *Proceedings of TMCE 2018*.
- Chu, K.F., Magsino, E.R., Ho, I.W.H., Chau, C.K., 2017. Index Coding of Point Cloud-Based Road Map Data for Autonomous Driving, in: *IEEE Vehicular Technology Conference*. <https://doi.org/10.1109/VTCSpring.2017.8108280>
- Cloudcompare, 2016. Cloudcompare (version 2.9.alpha) [WWW Document]. URL <http://www.cloudcompare.org/> (accessed 1.31.19).
- Dai, A., Chang, A.X., Savva, M., Halber, M., Funkhouser, T., Nießner, M., 2017. ScanNet: Richly-annotated 3D reconstructions of indoor scenes, in: *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*. <https://doi.org/10.1109/CVPR.2017.261>

- Hackel, T., Savinov, N., Ladicky, L., Wegner, J.D., Schindler, K., Pollefeys, M., 2017. SEMANTIC3D.NET: A new large-scale point cloud classification benchmark, in: ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences. <https://doi.org/10.5194/isprs-annals-IV-1-W1-91-2017>
- Handa, A., Patraucean, V., Badrinarayanan, V., Stent, S., Cipolla, R., 2015. SynthCam3D: Semantic Understanding With Synthetic Indoor Scenes. arXiv Prepr. arXiv1505.00171 1–5. <https://doi.org/10.1109/CVPR.2016.442>
- Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet Classification with Deep Convolutional Neural Networks. Adv. Neural Inf. Process. Syst. <https://doi.org/http://dx.doi.org/10.1016/j.protcy.2014.09.007>
- LFM Software [WWW Document], 2016. URL <https://www.lfmsoftware.com/> (accessed 1.31.19).
- Maddern, W., Pascoe, G., Newman, P., 2015. Leveraging experience for large-scale LIDAR localisation in changing cities, in: Proceedings - IEEE International Conference on Robotics and Automation. <https://doi.org/10.1109/ICRA.2015.7139414>
- Parrott, A., Lane, W., 2017. Industry 4.0 and the digital twin: Manufacturing meets its match. Deloitte Univ. Press.
- Qi, C.R., Yi, L., Su, H., Guibas, L.J., 2017. PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space, in: Computer Vision and Pattern Recognition (CVPR).
- Qi, R., Su, H., K., M., J., G.L., 2017. PointNET: Deep Learning on Point Sets for 3D Classification and Segmentation, in: Computer Vision and Pattern Recognition (CVPR).
- Roynard, X., Deschaud, J.-E., Goulette, F., 2018. Paris-Lille-3D: A Point Cloud Dataset for Urban Scene Segmentation and classification, in: CVPR 2018.
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., Fei-Fei, L., 2015. ImageNet Large Scale Visual Recognition Challenge. Int. J. Comput. Vis. 115, 211–252. <https://doi.org/10.1007/s11263-015-0816-y>
- Russell, B.C., Torralba, A., Murphy, K.P., Freeman, W.T., 2008. LabelMe: A database and web-based tool for image annotation. Int. J. Comput. Vis. <https://doi.org/10.1007/s11263-007-0090-8>
- Rusu, R.B., Cousins, S., 2011. 3D is here: Point Cloud Library (PCL), in: Proceedings - IEEE International Conference on Robotics and Automation. <https://doi.org/10.1109/ICRA.2011.5980567>
- Shilane, P., Min, P., Kazhdan, M., Funkhouser, T., 2004. The Princeton Shape Benchmark, in: Proceedings - Shape Modeling International SMI 2004. pp. 167–178. <https://doi.org/10.1109/SMI.2004.1314504>
- Silberman, N., Hoiem, D., Kohli, P., Fergus, R., 2012. Indoor segmentation and support inference from RGBD images, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). https://doi.org/10.1007/978-3-642-33715-4_54
- Song, S., Lichtenberg, S.P., Xiao, J., 2015. SUN RGB-D: A RGB-D scene understanding benchmark suite, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. <https://doi.org/10.1109/CVPR.2015.7298655>
- Xu, K., Kim, V.G., Huang, Q., Kalogerakis, E., 2015. Data-Driven Shape Analysis and Processing. Comput. Graph. Forum xx, 1–27. <https://doi.org/10.1111/cgf.12790>