OpenSketch: A Richly-Annotated Dataset of Product Design Sketches Supplemental Materials

YULIA GRYADITSKAYA, Université Côte d'Azur, Inria
MARK SYPESTEYN, Delft University of Technology, Faculty of Industrial Design Engineering
JAN WILLEM HOFTIJZER, Delft University of Technology, Faculty of Industrial Design Engineering
SYLVIA PONT, Delft University of Technology, Perceptual Intelligence lab
FRÉDO DURAND, MIT CSAIL, Inria
ADRIEN BOUSSEAU, Université Côte d'Azur, Inria

ACM Reference Format:
Yulia Gryaditskaya, Mark Sypesteyn, Jan Willem Hoftijzer, Sylvia Pont, Frédo Durand, and Adrien Bousseau. 2019. OpenSketch: A Richly-Annotated Dataset of Product Design Sketches Supplemental Materials. *ACM Trans. Graph.* 38, 6, Article 232 (November 2019), 6 pages. https://doi.org/10.1145/3355089.3356533

In this document we include additional details about the design of our dataset and its analysis. We also provide additional proof-of-concept applications in Section 7.

1 SELECTED SHAPES

We designed 9 shapes of varying complexity that we thought would require the techniques described in the paper in Section 3 to be drawn accurately (Figure 6, a-i). In particular:

Wobble surface contains a convex arch and a concave hole, which need descriptive cross-section lines to be well explained (OpenSketch paper, Section 3.1), even though those cross-sections are not necessary to construct the surface. It is also a symmetric shape that can benefit from rectangle duplication (OpenSketch paper, Section 3.2.2) and curve mirroring techniques (OpenSketch paper, Section 3.2.3).

House is composed of two levels of equal height, which can be created by either drawing a cuboid scaffold for one of the levels and duplicating it or by drawing a cuboid scaffold for both of the levels and dividing it (OpenSketch paper, Section 3.2.1 and 3.2.2). Locating the tip of the roof also requires dividing the façade into two equal parts (OpenSketch paper, Section 3.2.2).

Waffle iron is composed of two parts forming a hinge, whose rotational trajectory forms an ellipse when drawn in perspective (OpenSketch paper, Section 3.2.2). The two parts have rounded corners, which can be constructed using local cross-sections (OpenSketch paper, Section 3.2.3).

Mouse is a symmetric object with non-coplanar sides, which can be achieved by drawing one side and mirroring it with respect to a central cross-section plane (OpenSketch paper, Section 3.2.3).

Potato chip is a doubly-curved, symmetric smooth surface, which can be efficiently obtained by defining one or several projection planes and mirroring 3D curves (OpenSketch paper, Section 3.2.3).

Shampoo bottle has multiple curved sides forming non-planar creases at their intersections. Multiple cross-sections and projection lines can be used to construct this shape (OpenSketch paper, Section 3.2.3).

Tubes is an assemblage of three cylinders, each requiring drawing ellipses in perspective (OpenSketch paper, Section 3.2.1). Drawing the intersections between cylinders also requires projecting ellipses onto curved surfaces (OpenSketch paper, Section. 3.2.3).

Hairdryer is a beveled cylinder, which requires mirroring and projection lines for accurate construction (OpenSketch paper, Section 3.2.3). The handle can be positioned by drawing tangents to an elliptical cross-section of the body (OpenSketch paper, Section 3.2.1).

Vacuum cleaner is a compound object that can be decomposed into simple geometric scaffolds (OpenSketch paper, Section 3.2.1) before drawing detailed roundings using local cross-sections (OpenSketch paper, Section 3.2.3). Placing the elliptic hole on the sloped surface requires determining its minor and major axis.

We complemented these 9 shapes with a more complex ***kitchen mixer*** (OpenSketch paper, Figure 6, j), which appears in several sketch-based modeling and tutoring systems [Hennessey et al. 2017; Xu et al. 2014], and two shapes from the study by Cole et al. [2008] – ***bumps*** and ***flange*** (OpenSketch paper, Figure 6, k and l) – to allow comparison with this prior work. We selected these two shapes among the 12 used by Cole et al. because they resemble the man-made shapes that product designers frequently draw, and because they complement our shapes without adding unnecessary complexity. In particular, drawing ***flange*** requires dividing an ellipse into 6 equal arcs (OpenSketch paper, Section 3.2.2), which is not covered by our other shapes.

While all of the shapes present some form of symmetry, the two design sketching teachers commented that *"symmetric objects are very representative of industrial design sketching; indeed, many many products are symmetric"*.

Authors' addresses: Yulia Gryaditskaya, Université Côte d'Azur, Inria, yulia.gryaditskaya@gmail.com; Mark Sypesteyn, Delft University of Technology, Faculty of Industrial Design Engineering, M.Sypesteyn@tudelft.nl; Jan Willem Hoftijzer, Delft University of Technology, Faculty of Industrial Design Engineering, J.W.Hoftijzer@tudelft.nl; Sylvia Pont, Delft University of Technology, Perceptual Intelligence lab, S.C.Pont@tudelft.nl; Frédo Durand, MIT CSAIL, Inria, fredo@mit.edu; Adrien Bousseau, Université Côte d'Azur, Inria, adrien.bousseau@inria.fr.

0730-0301/2019/11-ART232 $15.00
https://doi.org/10.1145/3355089.3356533

2 DRAWING INTERFACE

We implemented a custom drawing interface compatible with pen tablets. The interface displays the three orthographic views on the side of the drawing canvas, either on the left or right depending on the dominant hand of the participant. The pen tablet records the trajectory and pressure of the pen along each pen stroke. We render each stroke as a semi-transparent polyline, with opacity and width linearly proportional to pressure. We set the maximum width to 1.5 pixels. Our interface also supports undo, but no eraser. Finally, we allow participants to sketch over the reference orthographic views to take measurements. The size of the canvas is selected automatically to maximize the drawing area with respect to the available screen resolution.

For the secondary presentation drawing task, we augment our interface with a layering system that displays the original sketch as an underlay over which participants trace their drawing. We display the underlay sketch in blue with user-adjustable opacity.

3 PARTICIPANTS

Table 1 details the level of experience of each participant. Table 2 lists the pair of objects assigned to one or two students.

Table 1. Experience of participants: number of years of studies for students and number of years of professional practice for professionals. Professionals 2,3 and 5 studied in the same design school as student participants.

Participant	Years	Participant	Years	Participant	Years
Student 1	< 1	*Student 6*	> 3	*Professional 2*	$2-5$
Student 2	< 1	*Student 7*	> 3	*Professional 3*	> 5
Student 3	$2-3$	*Student 8*	> 3	*Professional 4*	> 5
Student 4	> 3	*Student 9*	> 3	*Professional 5*	> 5
Student 5	> 3	*Professional 1*	< 1	*Professional 6*	> 10

Table 2. Each pair of object were drawn by one or two students, except ***House*** and ***Wobble surface*** which were drawn by all.

Pair of objects		Students
Mixer	***Waffle iron***	3
Hairdryer	***Mouse***	6 and 8
Tubes	***Bumps***	2 and 9
Shampoo bottle	***Vacuum cleaner***	1 and 7
Potato chip	***Flange***	4 and 5

All the participants completed the task remotely in multiple sessions using their own devices; 14 used different versions of Wacom tablets and one of the designers worked on a Genius tablet. One participant, *Professional 2*, experienced a compatibility issue with his tablet, which resulted in strokes that looked like polylines rather than smooth curves. This issue did not prevent the participant to produce compelling drawings, as shown in several figures of this paper. In total, we collected 107 concept and 107 presentation sketches from the first viewpoint and 105 concept and 104 presentation sketches from the second viewpoint. *Student 7* did not complete the drawings of the house model in both views and omitted the second viewpoint for the ***Shampoo bottle***. *Student 6* did not sketch the ***House*** from the second viewpoint. Finally, *Professional 4* did not complete the presentation drawing for the ***Waffle iron***.

Table 3. List of all labels in our dataset.

Lines types			
silhouette smooth	0	surfacing: mirroring	13
ridges visible	1	surfacing: temporary plane	14
vallleys visible	2	surfacing: projection lines	15
ridges occluded	3	proportions div. rectangle2	16
valleys occluded	4	proportions div. rectangle3	17
discriptive cross sections	5	proportions div. ellipse	18
axis and grids	6	proportions mult. rectangle	19
scaffolds	7	hinging and rotating elements	20
scaffolds: lines to VP	8	ticks	21
scaffolds: square for an ellipse	9	hatching	22
scaffolds: tangents	10	text	23
surfacing: cross section	11	outline	24
surfacing: cross section scaffold	12	shadow construction	25

3.1 Stroke labeling

Table 3 lists the labels we used to annotate strokes of our sketches.

4 POSE ESTIMATION

We use a standard pose estimation algorithm to recover camera parameters from sparse correspondences [Hartley and Zisserman 2000, Sec. 7.2: Algorithm 7.1]. The algorithm computes a camera projection matrix in three main steps – coordinate normalization, initialization using a linear method, and non-linear least-squares optimization. While Direct Linear Transform (DLT) is often used for the initialization [Hartley and Zisserman 2000, Sec. 4.1: Algorithm 4.1], we rely instead on a RANndom SAmple Consensus algorithm (RANSAC, [Hartley and Zisserman 2000, Sec. 4.8 Algorithm 4.6]) to be robust to outliers, which in our case occur when the drawing contains substantial inaccuracies that prevent all the point correspondences to be well explained by a single camera model. For inlier filtering, we use a fixed distance threshold of 0.4 in normalized coordinates. We then use the estimated camera parameters as an initialization for a Levenberg-Marquardt optimization that minimizes the image-space reprojection error of the inlier points. Note that we center each 3D model so that its center of mass is at the origin prior to running the algorithm.

We use the algorithm to compute a general camera matrix with 11 parameters, as well as a more restricted 9-parameters camera model with zero skew and equal horizontal and vertical fields-of-view. In practice, we first add a soft penalty term to the Levenberg-Marquardt optimization to favor small skew s and equal horizontal and vertical fields-of-view f_x and f_y: $w\left(s^2+(f_x-f_y)^2\right)$, where we set the weight w to 10 [Hartley and Zisserman 2000, Sec. 7.3]. We then use the values obtained with this first optimization to initialize a second optimization where we fix skew to zero and solve for a single field-of-view initialized to $f = (f_x + f_y)/2$. We will provide a Matlab script to reproduce this pose estimation on our data.

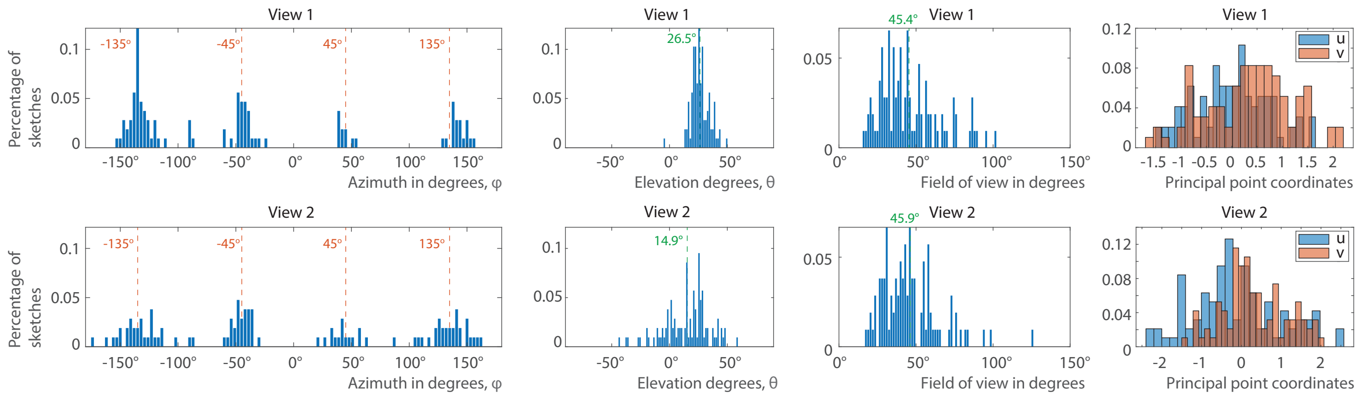


Fig. 1. Histograms of camera parameters. While participants followed our instructions to adopt a 3/4 viewpoint in the first sketch (top, 1st column), they adopted more diverse viewpoints for the second sketch (bottom, 1st column), with wider spread around the $\pm\pi/4$ and $\pm 3\pi/4$ azimuth angles. Some even adopted views from bellow, which correspond to negative elevation angles (bottom, 2nd column). In contrast, the distribution of fields-of-view remains similar in the two viewpoints (3rd column). Finally, in both viewpoints there are many sketches with non-centered principal point (4th column).

5 SHAPE COMPLEXITY

We asked participants to rate each shape they drew in terms of sketching complexity, on a 5-point Likert scale – 1 for *complex to sketch* and 5 for *easy to sketch* (Table 4, 2nd and 6th columns). ***Bumps*** was rated the most complex shape, about which *Professional 2* wrote: "*The top surface is complex and hard to define in a line sketch because it has no hard edges*". ***Mixer*** is the second most complex shape, in agreement with our motivation for selecting it. ***Flange*** was also judged difficult, possibly because it contains many ellipses that are hard to sketch well, especially on a pen tablet. Interestingly, the visually simple ***Wobble surface*** was ranked to be much more complex than the visually complex ***Vacuum cleaner***. *Professional 1* commented that the objects cover a good range of complexity and geometry, in line with the range of scores we obtained. The average sketching time varies between less than eight minutes to more than 20 minutes (Table 4, 3rd and 7th columns), and the average number of strokes between less than 140 and more than 350 (Table 4, 4th and 8th columns).

Table 4. Shapes are ordered according to their average sketching complexity score (1 = complex, 5 = easy). The 3rd and 7th columns provide the average time on both views in minutes spent by participants to draw the shape. The 4th and 8th columns provide to the average number of strokes. In both cases the values exceeding 98% percentile on all sketches were excluded prior to averaging.

Shape	Score	Time	№Str.	Shape	Score	Time	№Str.
Bumps	1.86	17.65	310.07	***Mouse***	3.38	11.15	165.69
Mixer	2.00	17.01	354	***Vacuum cleaner***	3.50	13.49	285.6
Flange	2.29	18.76	226.64	***Hairdryer***	3.63	12.45	222.06
Wobble surface	2.83	7.44	145.87	***Waffle iron***	3.86	16.41	196.29
Tubes	2.86	19.09	226.56	***Shampoo bottle***	4.00	12.76	185.8
Potato chip	3.00	10.6	136.13	***House***	4.10	9.26	137.48

The rated shape complexity correlates weakly with both the number of strokes used and the time spent (see inset for Spearman correlation and p-values). The time spent moderately correlates with the number of strokes, Spearman correlation is 0.522 with p-value of $1e-13$.

	Complexity	
	r_s	p_s
Num.~str.	-0.277	2e-04
Time	-0.325	1e-05

Table 5 summarizes the average complexity of the task asked to participants. While pairs of students were assigned different shapes, the estimated complexity is around 3.2 for all, close to the average complexity of all shapes drawn by professionals (3.1).

Table 5. The estimated complexity of the task for each of the participants, based on the average complexity score of each shape.

Student 1	3.61	*Student 2,9*	2.91	*Student 3*	3.20
Student 4,5	3.05	*Student 6,8*	3.48	*Student 7*	3.44
Professionals	3.11	Mean	3.19	Standard deviation	0.21

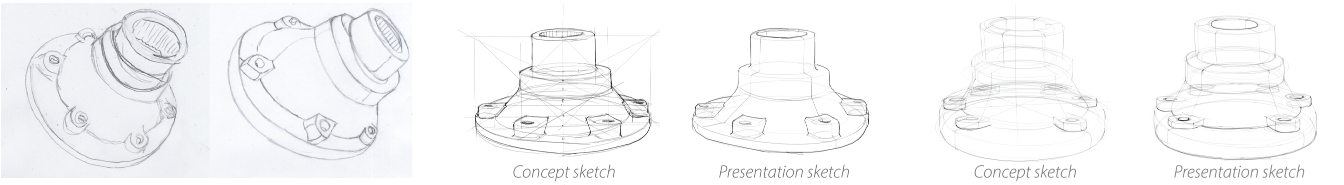
For each participant, we computed the average time and number of strokes per sketch (Table 6). Both quantities vary widely within the two groups of participants, and do not correlate with participant experience, rather indicating individual sketching strategies.

6 COMPARISON TO COLE ET AL.

Our analysis of re-projection error reveals that designers draw accurate representations of an object given three orthographic views. This accuracy contrasts with distortions observed in the drawings collected by Cole et al. [2008], where participants were given a rendering from the same viewpoint as the one they used for drawing. Figure 2 provides a visual comparison between representative drawings from the two datasets.

7 APPLICATIONS

In this section we provide additional proof-of-concept applications of our dataset.



a. Non-designers [Cole et al.] b. Professional 2 c. Professional 3

Fig. 2. Comparison between drawings from the study by Cole et al. [2008] (a) and concept and presentation sketches produced by designers with increasing expertise (b,c). Our designers drew more accurate drawings and employed a wider variety of lines. In particular, concept sketches contain lines to construct symmetric shapes. Our design drawings also often include hidden creases (b,c). Both presentation drawings contain descriptive cross-sections (b,c), which also appear in one of the drawings from Cole et al. (a, right). See also Figure 6 for sketches of the *bumps* shape that also appeared in Cole et al.'s study.

Table 6. Average time in minutes and number of stokes spent on each sketch, for each of the participants, excluding the values exceeding 98% percentiles.

Student	Time	№Str.	Professional	Time	№Str.
Student 1	8.32	274.13	*Professional 1*	4.11	108.96
Student 2	19.30	331.75	*Professional 2*	2.85	136.17
Student 3	24.06	223.63	*Professional 3*	5.97	93.38
Student 4	7.95	110.75	*Professional 4*	50.04	198.14
Student 5	11.13	113.38	*Professional 5*	5.78	330.40
Student 6	4.12	385.29	*Professional 6*	24.09	370.32
Student 7	11.58	196.40			
Student 8	7.06	88.75			
Student 9	9.47	200.50			

7.1 Evaluation of assumptions made by existing systems

One of the main motivation of our work is to provide data to test sketch-based modeling systems. However, the sketches we collected are composed of many overlapping strokes forming over-sketched lines. While great progress has recently been made on line drawing vectorization [Bessmeltsev and Solomon 2018; Favreau et al. 2016; Liu et al. 2018], further research is needed to automatically convert our data into clean, well-connected networks of vectorial curves suitable for benchmarking existing systems. Nevertheless, our dataset can be used to evaluate whether the lines that such systems expect are indeed present in product design sketches.

Figure 3 illustrates such an evaluation for *True2Form* [Xu et al. 2014], an algorithm that lifts a curve network to 3D by imposing regularity constraints. For instance, this algorithm requires that every surface patch contains at least one pair of intersecting cross-sections, each such intersection providing one orthogonality constraint. We used our dataset to evaluate whether designers include required section lines, using as a reference a curve network of a ***Mixer*** shown in the original *True2Form* paper. We compared this network to our sketches of the same object, counting for each input cross-section line its presence in our sketches. We used the presentation sketches for this evaluation since they contain more descriptive cross-section lines than the original sketches. In this example, the cross-sections used as input to *True2Form* were present in 70% of our sketches (Figure 3 top-left). In particular, all participants drew the section lines that also denote the global symmetry of the shape. However, many participants omitted the horizontal section on the bowl (Figure 3, red) and on the top part of the mixer (Figure 3, orange), which *True2Form* require to obtain a well-connected network with enough orthogonality constraints. We hope that our dataset will similarly inform researchers about the usage of different types of lines in real-world product design sketches.

7.2 Automatic stroke classification

As a third application, we use our labeled data to train a classifier that predicts whether a stroke represents a visible descriptive line, or a hidden or construction line. Our motivation for those two classes is that visible descriptive lines are the ones most present in presentation drawings, and should as such be of special interest for sketch-filtering algorithms.

We use a non-linear SVM with a Gaussian kernel function, which we train to predict a stroke label given a set of stroke features – *length*, *speed*, *time*, *pressure*, *mean curvature*. We normalize the *length*, *speed* and *pressure* within each sketch by dividing their values by the 95%-percentile values over all strokes in that sketch. Finally, for a given training set of sketches, we randomly select 1000 strokes to form the training set for the classifier.

We first trained one classifier per designer, using a subset of objects for testing and all other objects for training. We used the ***Wobble surface*** as a test object for students, and the ***Wobble surface***, ***Bumps*** and ***Tubes*** as test objects for professionals. Table 7 shows the classification accuracy when each of the five features is used separately, and with several feature combinations. The highest accuracy, 73.5% on average, is achieved when all of the features are used (last column highlighted in bold). *Speed* is the most discriminative feature, achieving an accuracy of 69.7% when used alone, followed by time (66.1%) and pressure (64.6%).

We then trained a single classifier, using a subset of objects and designers for testing, and all others for training. We again used the ***Wobble surface***, ***Bumps*** and ***Tubes*** as test objects, and *Student 2*, *Student 5*, *Student 9*, *Professional 3* and *Professional 5* as test designers. We selected this partitioning to cover a wide range of experience and sketching strategies in both train and test sets, while also assigning to the test set the students who drew the test objects. Table 8 reports the accuracy of the classifier for different subsets of designers and objects. The last two columns show that the classifier generalizes well to new designers and new shapes, with a median accuracy of

Table 7. Percentages of correctly labeled lines, for different combinations of features. Here we trained one classifier per designer.

	Speed (S)		Time (T)		Pressure (P)		Length (L)		Curvature (C)		S.T.P.		S.T.P.C		S.T.P.L.		S.T.P.L.C	
	train	test	train	test	train	test	train	test	train	test	train	test	train	test	train	test	train	test
Mean	69.1	69.7	75.6	66.1	66.4	64.6	68.2	63.5	62.8	60.7	81.7	73.4	81.6	73.1	84.7	72.2	84.1	**73.5**
Median	68.9	67.9	74.8	65.3	67.5	64.9	69.8	63.1	63.1	61.1	81.0	75.9	82.2	71.8	85.1	73.0	86.6	**73.9**
Std	6.2	9.3	3.7	7.3	4.6	7.4	4.8	8.9	5.5	7.6	5.7	6.7	3.9	6.2	5.5	5.1	5.7	**3.9**

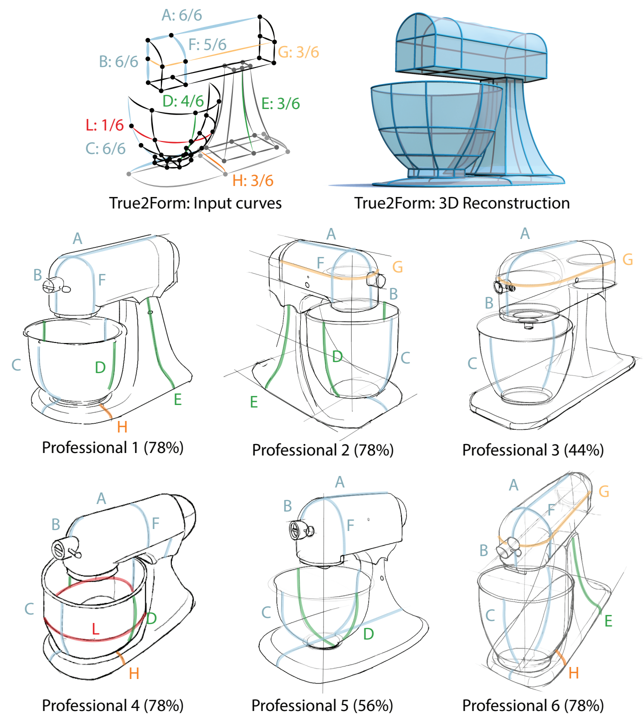


Fig. 3. Comparison between one of the inputs shown in True2Form [Xu et al. 2014] (top left, input curves and 3D reconstruction from the original paper) and 6 sketches from our dataset of the ***Mixer*** model drawn from a similar viewpoint. Since True2Form [Xu et al. 2014] exploits cross-section lines to lift a curve network to 3D, we count for each of the cross-sections of the curve network its presence in our sketches. Each line is highlighted in each of the sketches where it appears, and its overall usage is reported as a fraction in the top left image. In addition, we indicate underneath each sketch the percentage of lines present. While cross-sections that denote symmetries or smooth ridges are present in most sketches (blue and green labels, lines $A-F$), other cross-sections only appear in a few sketches even though they are necessary to provide enough constraints to the algorithm (red and orange labels, lines $G-L$).

76.2% on sketches of objects and participants not present in the training set. Note that our classifier analyses each stroke independently, while accounting for spatial relationships may increase accuracy further [Schneider and Tuytelaars 2016].

Finally, Figure 4 provides a visual comparison between the stroke classification produced by the classifier and ground truth, for two sketches from two different participants. While the classifier identified most scaffold lines, it mis-classified a few descriptive lines, especially on the simple ***Wobble surface***.

Table 8. Percentages of correctly labeled lines trained on all features. The training is performed on a subset of designers and objects and tested on a complementary sets of designers and objects.

	S.T.P.L.C.			
	Train designers		Test designers	
	Train obj.	Test obj.	Train obj.	Test obj.
Mean	75.2	70.7	76.6	73.6
Median	75.6	75.2	76.8	76.2
Std	10.2	14.4	8.2	7.7

Table 9. Angular distances between the predicted normal maps and ground-truth normal maps. MO corresponds to the errors computed without accounting for global rotations, while MI corresponds to the full metric described in the main document.

Line rendering	D1		D2		D3		D4	
	MO	MI	MO	MI	MO	MI	MO	MI
Unifrom	34.3	**23.2**	35.5	**23.1**	34.6	**21.3**	32.5	**21.2**
Original	34.8	**23.6**	33.8	**22.3**	33.5	**21.7**	32.4	**21.6**

REFERENCES

Mikhail Bessmeltsev and Justin Solomon. 2018. Vectorization of Line Drawings via PolyVector Fields. *CoRR* (2018).

Forrester Cole, Aleksey Golovinskiy, Alex Limpaecher, Heather Stoddart Barros, Adam Finkelstein, Thomas Funkhouser, and Szymon Rusinkiewicz. 2008. Where do people draw lines?. In *ACM Trans. Graph. (Proc. SIGGRAPH)*, Vol. 27. ACM, 88.

Jean-Dominique Favreau, Florent Lafarge, and Adrien Bousseau. 2016. Fidelity vs. Simplicity: a Global Approach to Line Drawing Vectorization. *ACM Trans. Graph. (Proc. SIGGRAPH)* (2016).

Richard Hartley and Andrew Zisserman. 2000. Multiple View Geometry in Computer Vision Second Edition. *Cambridge University Press* (2000).

James W. Hennessey, Han Liu, Holger Winnemöller, Mira Dontcheva, and Niloy J. Mitra. 2017. How2Sketch: Generating Easy-to-follow Tutorials for Sketching 3D Objects. , Article 8 (2017), 11 pages.

Jan J. Koenderink, Andrea J. Van Doorn, and Astrid M. L. Kappers. 1992. Surface perception in pictures. *Perception & Psychophysics* 52, 5 (1992), 487–496.

Chenxi Liu, Enrique Rosales, and Alla Sheffer. 2018. StrokeAggregator: Consolidating Raw Sketches into Artist-Intended Curve Drawings. *ACM Trans. Graph. (Proc. SIGGRAPH)* 37, 4 (2018).

Rosália G. Schneider and Tinne Tuytelaars. 2016. Example-Based Sketch Segmentation and Labeling Using CRFs. *ACM Trans. Graph.* 35, 5 (July 2016), 9.

Baoxuan Xu, William Chang, Alla Sheffer, Adrien Bousseau, James McCrae, and Karan Singh. 2014. True2Form: 3D curve networks from 2D sketches via selective regularization. In *ACM Trans. Graph. (Proc. SIGGRAPH)*, Vol. 33.

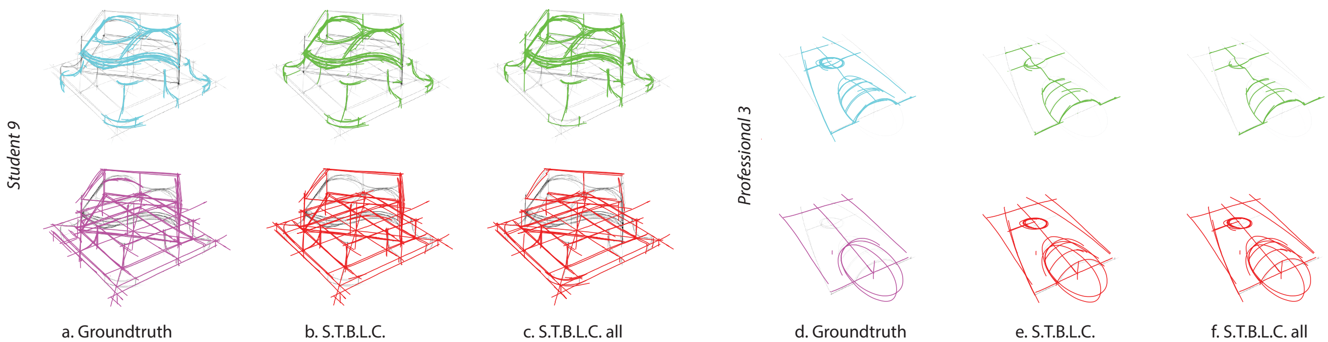


Fig. 4. Visualization of stroke classification for two designers and two test objects. Groundtruth labeling (a,d), classification on test objects when the training is performed on sketches of the same designer (b,e), and classification on test objects when the training is performed on sketches of other designers (c,f).