

Surface Perception of Planar Abstractions

JAMES MCCRAE, University of Toronto
 NILOY J. MITRA, University College London
 KARAN SINGH, University of Toronto

Various algorithms have been proposed to create planar abstractions of 3D models, but there has been no systematic effort to evaluate the effectiveness of such abstractions in terms of perception of the abstracted surfaces. In this work, we perform a large crowd-sourced study involving approximately 70k samples to evaluate how well users can orient gauges on planar abstractions of commonly occurring models. We test four styles of planar abstractions against ground truth surface representations, and analyze the data to discover a wide variety of correlations between task error and measurements relating to surface-specific properties such as curvature, local thickness and medial axis distance, and abstraction-specific properties. We use these discovered correlations to create linear models to predict error in surface understanding at a given point, for both surface representations and planar abstractions. Our predictive models reveal the geometric causes most responsible for error, and we demonstrate their potential use to build upon existing planar abstraction techniques in order to improve perception of the abstracted surface.

Categories and Subject Descriptors: I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling—*Curve, surface, solid, and object representations*

General Terms: Design, Experimentation, Human Factors, Measurement, Performance

Additional Key Words and Phrases: Crowd-sourced user study, perception, planar abstraction, surface representation

ACM Reference Format:

McCrae, J., Mitra, N. J., and Singh, K. 2013. Surface perception of planar abstractions. *ACM Trans. Appl. Percept.* 10, 3, Article 14 (August 2013), 20 pages.
 DOI: <http://dx.doi.org/10.1145/2501853>

1. INTRODUCTION

Shape information conveyed and perceived by viewers of a 3D shape abstraction is an important measure of its efficacy as a shape proxy for both art and engineering applications. Shape abstractions have been employed in art for centuries and more recently in scientific pursuits for shape design, visualization, and fabrication. Various algorithms have been proposed to produce such abstractions, evaluated either using geometric measures or visually by a few individuals. While such evaluations may be sufficient for judging recognition tasks, little is known about the accuracy of perception of the abstracted surface. Since the end-goal of most abstractions is in fact to be a perceptual stand-in for the original shape, a principled study of perception of 3D shape abstractions remains an ambitious, unavoidable, and worthy goal.

The authors thank GRAND for financial assistance.

Author's address: J. McCrae; email: mccrae@dgp.toronto.edu.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from Publications Dept., ACM, Inc., 2 Penn Plaza, Suite 701, New York, NY 10121-0701 USA, fax +1 (212) 869-0481, or permissions@acm.org.

© 2013 ACM 1544-3558/2013/08-ART14 \$15.00

DOI: <http://dx.doi.org/10.1145/2501853>

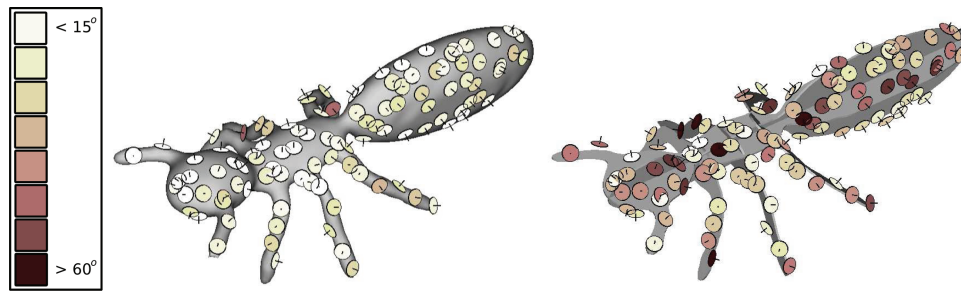


Fig. 1. We conduct a large crowd-sourced user-study to evaluate human perception of models represented by their shaded surfaces (left) and planar abstractions (right) using an established orientation task. Participants manipulate gauges scattered across the visible surface of models to define perceived normals at each point. Overlaid are error visualizations for a fixed view rotation task for the model *Ant* (error legend at left).

Different 3D shape abstractions serve different applications: 3D curves [Mehra et al. 2009], which capture the metaphor of wire armatures of the shape, provide an effective approach to shape creation via sketching, or geometric resurfacing; or 3D volumetric primitives that capture the volume of a closed shape have been explored in the context of artistic collages and applications in finite element remeshing. Another popular shape abstraction deals with the use of planar polygons to capture shape, either on the surface [Cohen-Steiner et al. 2004], or its interior. For the latter, axis-aligned regular or adaptive [Hildebrand et al. 2012] sections, radial sections and abstractions using a minimal arrangement of planar sections [McCrae et al. 2011] have been proposed. Such planar 3D abstractions have a number of advantages: they are easy to render as planar facets, simple to author using typical 2D pen or mouse input, and straightforward to fabricate using cutting. In this article, we focus on the perception of 3D shape represented by a variety of planar abstractions.

Perception of surfaces has been studied before: In a seminal effort, Koenderink et al. [1992] study how well subjects adjust a local gauge figure to perceptually *fit* a photographed surface. Cole et al. [2009] recently adapt the work to perform a large crowd-sourced study to evaluate accuracy of orientation estimates as perceived from line drawings. In our context, since planar abstractions are intended to be viewed in 3D, either once manufactured, or virtually via view manipulation, we need to investigate further. Specifically, we design an experiment to evaluate perceptual accuracy for orientation tasks, accounting for the perceptual impact of 3D viewpoint changes.

In this article, we study surface rendering along with 4 styles of planar abstraction across 6 models: 4 models from the Princeton Shape Benchmark (PSB) (a mix of organic and man-made objects) and 2 cylindrical objects created to establish baseline performance and test specific geometric conditions. For each planar abstraction, we use a small number of slices, one that minimally ensures a geometrically meaningful abstraction. For each combination of abstraction and model, we test for surface perception, that is, tasks relating to how well subjects can orient gauges on the abstracted models. Specifically, we designed two tasks: (i) orientation within a fixed view, and (ii) orientation where viewpoint rotation is allowed. To test these tasks, we designed and deployed a large crowd-sourced user study using the Amazon Mechanical Turk (AMT) framework. We collected approximately 70k gauge samples across all study conditions (task and representation), consisting of a total of 1161 runs of the study by 178 unique participants.

Finally, based on analysis of the study data, we reveal numerous strong correlations with various geometric properties and task error. Utilizing these, we describe a process to build linear models for each condition to predict error. These models require only a set of measurements at a given surface point and provide a predicted error. For some predictive models, our selection of predictors resulted

in an estimated prediction error of nearly 50% of the base error rate. The parameters to our models provide insight into the geometric causes of error and the models themselves can be used to evaluate perception error of abstractions, and potentially improve them.

2. RELATED WORK

In light of a recent flurry of research in creating 3D planar abstractions [Willis et al. 2010; Li et al. 2011; McCrae et al. 2011; Schwartzburg and Pauly 2011; Hildebrand et al. 2012], we present the first comparative study on the surface perception of 3D shape abstractions. Our work is inspired by the research of Cole et al. [2009], which provided a much needed perceptual basis and validation for a body of research on 2D line drawings of 3D objects. Line drawings of 3D shape however, are largely view-dependent and their perceived surfaces are appropriately studied using static 2D images [Cole et al. 2009]. We instead explore the perception of view-independent planar 3D abstractions, designed for 3D viewing.

3D shape abstraction. For decades now, designers have used networks of curves to effectively model 3D shapes. Recent research has shown that such curve abstractions can be effectively computed from 3D objects [Mehra et al. 2009] and that design curves are perceptually consistent [Shao et al. 2012]. Design curves, such as cross-section curves [Shao et al. 2012], are in fact often planar. Furthermore, planar surfaces and planes of reflective symmetry [Mitra et al. 2006; Simari et al. 2006] are common to both natural and man-made objects. An important thread of research thus relates planar abstractions and approximations of 3D shape.

Perhaps the earliest use of planar polygon abstraction has been to simplify 3D surface geometry [Cignoni et al. 1997]. Approaches based on variational shape analysis have been proposed [Cohen-Steiner et al. 2004; Zhang et al. 2012], which effectively optimize the geometric difference between simplified planar pieces and the original surface. Techniques such as billboard clouds [D  coret et al. 2003] or polypostors [Kavan et al. 2008] similarly use geometric or view estimates to project surface geometry onto a small number of planar polygons for lightweight rendering.

Planar abstractions ranging from pop-up sculptures [Li et al. 2011], radial [Willis et al. 2010], orthogonal [Schwartzburg and Pauly 2011], or axis-aligned *crdbrd* [Hildebrand et al. 2012] planar sections have been proposed, with a focus on 3D manufacturing. We validate this body of research by providing a comparative perceptual evaluation of radial, orthogonal (XYZ), and axis-aligned (*crdbrd*) abstractions.

McCrae et al. [2011] showed that most 3D objects could be abstracted by humans using a small number of planar sections and presented an algorithm to construct such a set of *minimal planar sections* (MPS). They also provided evidence that abstractions were as recognizable as the original objects. In contrast, we study the comparative fine-grained 3D surface perception of the MPS, XYZ, radial, and *crdbrd* abstractions.

Perceptual psychology and shaded surfaces. A significant body of perceptual psychology literature addresses human understanding of 3D shape from pictures [Koenderink et al. 1992]. The surface information conveyed by contours [Stevens 1981; Knill 1992] have been studied in various pictorial contexts. Significantly, an orthogonal grid of XY lines projected upon a rendered 3D height-field has been shown to enhance surface perception over a shaded surface [Sweet and Ware 2004]. Traditionally, these perceptual studies have been carried out from static viewpoints providing us with a rich understanding of shape from pictures but with little insight to our perception of an imagined 3D surface induced by a 3D abstraction.

Since our abstractions are shaded planar constructs, we also refer to prior art on surface perception of shaded objects [Ramachandran 1988]. Numerous studies have been conducted on the perception of shape under a variety of view, material and illumination conditions [Norman et al. 1995; Caniard

and Fleming 2007; O'Shea et al. 2008; Secord et al. 2011; Faisman and Langer 2012], sometimes with contradictory evidence, such as the whether specular highlights aid or hamper surfaces perceived under Lambertian shading [Faisman and Langer 2012]. Once again, these studies are either performed from static viewpoints or at most provide automatic view oscillation about a fixed viewpoint as a motion cue to shape perception. We take the findings of this body of research into account when designing the viewing, lighting and material parameters for our study in Section 3.

Finally, we acknowledge a corpus of research in shape understanding of line drawings and other artistic illustrations of shape [Saito and Takahashi 1990; Phillips et al. 2003; Winnemöller et al. 2007; Cole et al. 2009, 2012] and attempt to provide similar answers to the surface perception of view-independent 3D planar abstractions.

3. USER STUDY

Our goal is to study the effectiveness of various planar abstractions for representing 3D shapes. Broadly, we investigate the following questions: (i) how do planar abstractions compare with the original 3D models in terms of human perception of the underlying shape; (ii) how well can humans orient normals on the surface of the abstracted shape; (iii) how consistently are the same planar abstractions perceived by different humans; (iv) how do geometric features relate to the error distribution across planar abstractions.

We performed several user studies to answer the above questions. Designing the user study comprised the following stages: choosing among candidate planar abstractions, designing specific tasks for the subjects to perform, gathering inputs from a large pool of subjects, and detecting and removing outliers before analyzing the data.

3.1 Candidate Planar Abstractions

In related user studies, Cole et al. [2009] compare accuracy of perceived orientations based on human and machine generated line drawings from fixed viewpoints, while Secord et al. [2011] develop a perceptual model for viewpoint preference for 3D models based on the view preferences of a large number of subjects. We also test human ability to orient normals on the perceived surface, both for shaded surfaces and their corresponding planar abstractions. What further separates our work from previous work is that we also test the condition of allowing view manipulation for the orientation task, since the planar abstractions are 3D objects themselves and are meant to be viewed from a number of different directions.

As ground truth, we used 3D surface models shaded with Lambertian reflectance to test human ability to perceive the surface. For comparison, we also selected four different planar abstractions (see Figure 2): (i) minimal planar slices (MPS) [McCrae et al. 2011]; (ii) regular XYZ-aligned planar slices; (iii) radially arranged planar slices with a manually selected centroid and the model-up vector as the axis direction; and (iv) crdbrd [Hildebrand et al. 2012]. In each case, we selected a suitable number of planes such that each abstraction was compact, yet of sufficient subjective quality to represent the original surface. The range of styles of planar abstraction in our study captures the large variation of the number of planes used. For example, taking the mean over all 6 models used in our studies, MPS used 5.0 slices, crdbrd used 11.3 slices, radial used 18.0 slices, and XYZ used 40.6 slices.

Planar abstractions in general have been observed [McCrae et al. 2011; Hildebrand et al. 2012] to represent some classes of 3D models better than others. For example, man-made objects being regular and comprising of large planar faces are better abstracted with planar slices as compared to a highly articulated organic shape, such as a spider. We can expect such variability within the genre of planar abstractions itself, such as XYZ slices being better suited to boxlike objects than radial slices. We thus

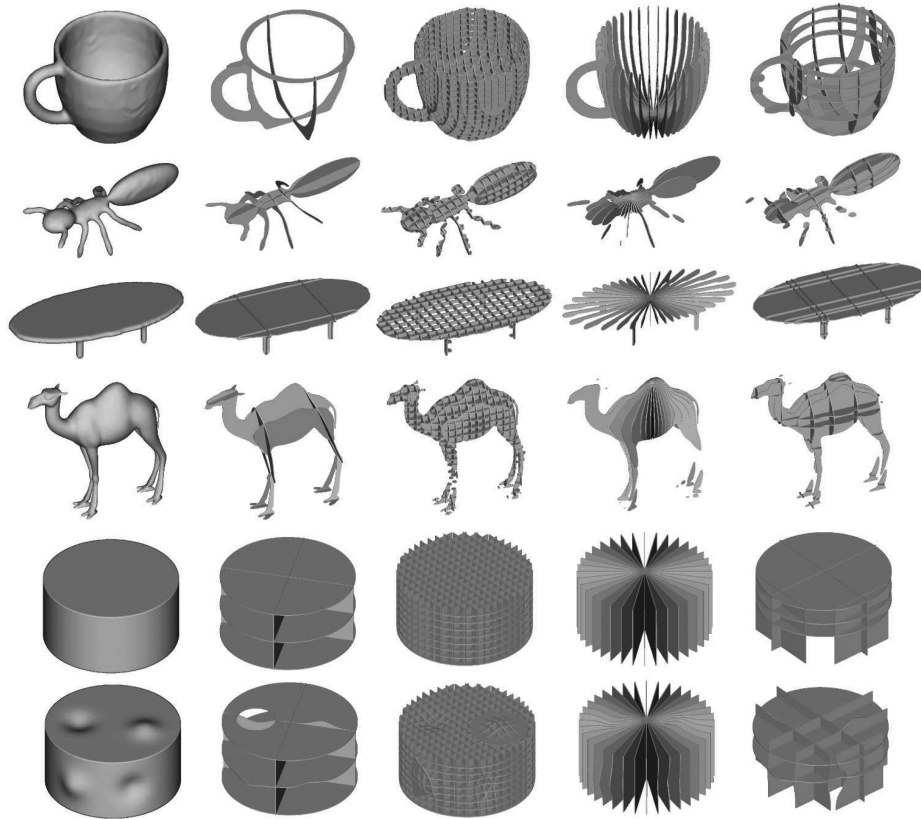


Fig. 2. The models and various representations used in our study. From top to bottom: *Cup*, *Ant*, *Table*, *Camel*, *Cylinder*, *Bumpy*. From left to right: surface, MPS (minimal planar sections), XYZ (axis-aligned slices), radial (regular radial slices) and crdbrd. Models *Cup*, *Ant*, *Table*, *Camel* from the Watertight Track of SHREC 2007 [Giorgi et al. 2007] and used with permission.

chose a range of models of varying geometric complexity to check the representative qualities of the different planar abstractions.

3.2 Methodology

Both tasks in the study involve orienting normals at sampled surface points, perceived from a presented representation (a shaded surface model or planar abstraction). We evaluated two tasks: orientation with a fixed viewpoint, and orientation while allowing the viewpoint to rotate around what is shown.

User interface. Participants were first shown a webpage that contained a link to download the study application, as well as instructions which are applicable to all conditions (task and representation). These first instructions explain the basic goals of the participant: how to rotate the gauges, how to proceed to the next trial, and how to submit the generated study data. This also lets the participant decide if this is a study they are interested in.

Running the downloaded study application, participants first enter their unique ID provided by the AMT framework. We map this unique ID to the study conditions for the participant, and additional instructions are displayed specific to these conditions. For a participant given the viewpoint rotation task, instructions on how to rotate the view are provided. All participants are shown examples of bad

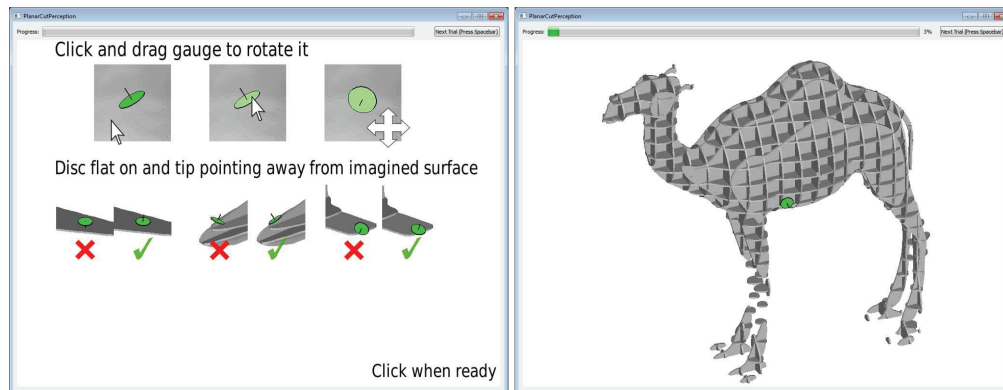


Fig. 3. (Left) An instruction screen shown initially for every run, with information specific to the experiment condition for the participant, shown is the fixed view MPS condition. (Right) The study interface shows a representation at center, here the XYZ style planar abstraction of model *Camel*. For the *fixed view* task the viewpoint is stationary for each model, but for the *rotated view* task the user can rotate the viewpoint around the model by left-dragging.

and good gauge orientations, but whether a surface or planar abstraction is shown in the examples is based on condition. For the conditions where planar abstractions are shown, the instructions state that while performing the task, they should be visualizing or imagining the surface being represented. Figure 3 shows the typical interface to our studies along with sample instruction images shown to the subjects (see also supplementary material).

Gauges. Similar to Cole et al. [2009], we show participants a series of gauges (i.e., an oriented disc at given points) in random order. Participants are asked to rotate each gauge to the correct (perceived) orientation before moving on to the next gauge. Previous investigations of perception of pictures and sketches [Koenderink et al. 1992; Cole et al. 2009] use Halton sequences to define candidate gauge positions in screen space. Like these studies, we use Halton sequences to define 100 quasi-random gauge positions for each of the 6 models in the study. In addition, we hand author 5 additional gauge positions, accounting for regions where the spatial sampling is sparse, or if an interesting surface feature lacks gauge coverage (such as a bump). Thus each model in the study has a total of 105 unique positions where gauges are placed and ground-truth surface normals are known.

The gauges are drawn as small circular discs, with a single line segment orthogonal to the disc indicating the normal. Gauges are colored green to make them clearly visible. We use a fixed gauge size of 1.8% of the diameter of the models (see Figure 3 (right)). Occlusion between the gauge and model would unintentionally provide hints about orientation, so to avoid this gauges are effectively drawn “on top” of the rendering of the presented model representation.

View and viewpoint rotation. For each model, we manually pick non-accidental default views based on the principles from Secord et al. [2011]. For rendering, the gauges and model share the same orthographic projection.

Specific to the rotating view task, when a new model is presented users observe a 2 second animation where the viewpoint spins 90 degrees about the up vector and comes to rest at the predefined view-point. Importantly, when the user initiates a viewpoint rotation, gauges are not rendered at any time throughout the manipulation. The reason behind this is that viewpoint rotation should only aid in the spatial understanding of the presented representation and not the orientation task itself. If the gauge were to remain visible as a floating 3D object as the view rotates, information is being provided about a 3D point on a surface which is not explicitly presented by a surface abstraction (i.e., we are revealing

more perceptual information than what is represented by the abstraction). By hiding the gauges during viewpoint rotation, we prevent any extra information relevant to the orientation task from being unintentionally provided. When the user ends the viewpoint rotation, a 0.75 second animation rotates the view back to the predefined viewpoint rather than immediately resetting it.

We also experimented with two other candidate tasks in a pilot study: a gauge *positioning* and *combined* task (both rotating and positioning). For the former, participants would left-drag a gauge to translate it along the fixed axis it was oriented. To do this effectively, a viewpoint rotation is often necessary, to view the gauge at angle distant from the translation axis. For conditions where the surface is shown, the positioning task becomes trivial—simply choose a view where the surface point is a silhouette and place the gauge on it. Even for the abstraction conditions, the direction the gauge is oriented in the position task reveals information about a local normal which is not explicitly represented by the abstraction. The idea of the combined task aimed to overcome these problems by not revealing any information explicitly represented, but in our pilot experiment we discovered the combined task to be too complex to explain to participants, and that it was difficult for them to produce meaningful results.

Shading and lighting. Models are rendered with Lambertian reflectance for shading, with the light coming from behind the camera at infinity. There are several reasons for this: we choose shaded models since that is how 3D abstractions would appear if manufactured; we do not add lines or contours to the shading so as not to confuse with the percept produced by line drawings that are an abstraction in themselves; we choose Lambertian shading since the effect of specular highlights on surface perception is unclear [Faisman and Langer 2012], cast shadows are avoided since they can interfere with surface perception [O’Shea et al. 2008]; while O’Shea et al. [2008] suggest that an elevation of 20–30° above the viewpoint for best perceptual results, we observed little visual difference between that and light from behind the view, and further expect users to manipulate the view under some conditions, where the light direction will move with viewpoint. If the light direction is kept stationary, surface points not facing the predefined viewpoint would be uniformly dark, leading to less-complete shape understanding. In addition, surface normals at points near silhouette edges are more difficult to perceive because shading is uniformly dark there. Since the purpose of a viewpoint rotation condition is to explore effects of improved 3D shape understanding, our choice of a light direction which changes with viewpoint supports this. (Equivalently, one can also think of the viewpoint as fixed with the light direction fixed from behind at infinity; only the presented representation rotates.)

Study trials. In each trial the participant manipulates one of the gauges. We randomly set the initial orientation of the gauge to a point on the front-facing hemisphere. We selected six models in total, four from the Princeton Shape Benchmark (*Cup*, *Ant*, *Table* and *Camel*) which are a mix of man-made and organic shapes, and two manually created shapes (*Cylinder* and *Bumpy*). For each of the models presented in a random sequence, the participant performs 10 trials (5 of which are duplicates) totaling 60 trials in each study run. The participant is allowed to run the study up to a maximum of 10 times, but always subject to the same condition (task and shape representation) which is based on their unique ID, in order to avoid learning effects.

At the end of a study, the participant is informed if the performance was acceptable and the results can be submitted, or the performance was unacceptable and the results were discarded (see Section 3.4).

Tasks. In our study, participants were assigned one of the following two tasks.

- (i) *Fixed view task.* The user prescribes orientations by rotating gauges on fixed positions on shaded models or their abstractions, viewed from prespecified viewpoints. The participant can left-drag gauges in order to rotate them. This task is performed with static viewpoints to keep it simple and connect our study to an established task protocol [Cole et al. 2009].

- (ii) *Rotated view task.* Each presented model is first rotated 90 degrees in an animation lasting 2 seconds. At any time, the participant can then rotate the viewpoint by left-dragging anywhere on the screen, but off the gauge. Otherwise, this task is the same as the fixed view task.

Error measures. Gauge positions on the surface are determined by inverting the projection from each screen-space position, casting a ray outward and using the closest intersection point on the surface. The measured error for a gauge is the angle between the gauge normal and the true surface normal at that point, always expressed in degrees. Initial gauge orientations are set randomly to point within the front-facing hemisphere.

3.3 Data Collection

We deployed our studies to a large number of users using the Amazon Mechanical Turk (AMT) framework (see also [Cole et al. 2009; Secord et al. 2011]). While this allowed us to efficiently gather large number of inputs from an uncontrolled pool of participants, we had to detect and discard outlier samples generated by insincere users (see Section 3.4).

The scale of each model has been normalized to fit within a unit-diameter sphere. For each model, we randomly choose a sequence of 5 gauge positions from the 105 prepared. Each of the 5 gauges is presented twice, but each has its own randomized initial configuration. All gauge positions are taken at surface positions which are both front-facing and unoccluded relative to the predefined viewpoint for the model.

In order to avoid cross-condition learning effects, participants are required to enter a unique ID provided by AMT which is used to define the experiment conditions (task and representation presented). Participants are encouraged to run the study up to a maximum of 10 times, and conditions remain constant throughout multiple runs (e.g., a participant selected to be shown a planar abstraction as the model representation would never later see the surface representation). If a participant opts to perform multiple runs of the study, we keep track of the gauges they have previously been assigned, so that they are presented new gauge positions each run. We manually verify that the ID entered by each participant into the study application matches the ID of the data submission on AMT.

Our complete study consists of 2 tasks: fixed view and rotating view. It also consists of 5 model representations: the surface and four styles of planar abstraction (MPS, XYZ, radial and crdbrd). In a single run of the study, a participant manipulates 60 gauges in total, 10 for each of 6 models. 178 unique participants performed a total of 1161 runs of the study, fairly evenly covering the 10 combinations of task and model representation (each condition having between 84 and 152 runs). We gathered a total of approximately 70,000 gauge samples which were used in the subsequent analysis. For each run of the study a participant was awarded \$0.35.

3.4 Data Verification

Participants coming from the AMT may only be interested in maximizing money earned in the shortest time and with the least effort. On the other hand, for 8 of our 10 conditions participants were required to imagine or fill in details of a surface only from a planar abstraction, leading to errors despite best intentions. It is such errors we are most interested in. Hence, we design a filter that tolerates errors arising from (honest) intent.

We adapt a solution similar to those proposed in other crowd-sourced user studies [Cole et al. 2009], namely having the participants repeat the same task. Specifically, participants set 10 gauges for each model, but there are in fact 5 *gauge pairs*, where gauges in a pair share a unique position. Each gauge within a pair is set with a different initial gauge orientation. For each pair, we evaluate if the two gauge settings roughly agree (i.e., their orientations are similar). Since gauges are initialized with

random orientations, participants who are trying in earnest and have a specific mental “goal” at a specific gauge position will tend to perform consistently, but not if they are carelessly manipulating the gauges.

A participant’s work is deemed unacceptable if (i) less than 70% of gauge pairs are such that the gauges in each pair are within 30 degrees of each other; or (ii) if the standard deviation of the angle between pairs is less than 5 degrees. The second condition detects insincere participants who do not perform the task they have been instructed (e.g., they align all gauges in the view direction to pass the first condition, but the inputs are meaningless) (see also Cole et al. [2009]).

On task comprehension. One might point out that when a participant is presented with only the planar abstraction, from their perspective the hidden surface is *more abstract* than the planar abstraction. We maintain that our planar abstractions are more abstract than surface representations, despite presentation of only the planar abstraction. Objectively, surface representations reveal object shape explicitly, while planar abstractions allow room for interpretation of shape. Subjectively, our reasoning is that surface representations are of greater geometric similarity to how most objects in reality present themselves, which all participants will be acquainted with.

One might also point out that when a participant is presented with only the planar abstraction, this may produce difficulty with task comprehension—the participant may align gauges to the planar sections directly rather than to an imagined surface. Further, one may consider the observation that a participant’s estimated normals matching nearest planar section normals is evidence that they misunderstand the task. However, how does one reliably tell the difference between (i) the participant’s intent of orienting the gauge directly onto a planar section, or (ii) the participant fully comprehends the task yet orients onto an (incorrectly) imagined surface (which happens to align with the nearest planar section)? Because of this fundamental difficulty, we did not remove or filter participant data where we may suspect the task was not fully understood. In the design of our study, we were careful in our validation of participant data to allow inclusion of any sources of error, as long as results were consistent. If we did not do so, we would potentially introduce bias and arrive at incorrect conclusions. The best we can do is to strive to make the instructions clear so that participants understand the task. On both AMT and in our study application, each time the program is run, the initially presented instructions state (for the planar abstraction conditions) that gauges should be oriented on the imagined surface. Visual examples where gauges are slanted/tilted away from the nearest planar section normal are also provided (see Figure 3).

4. ANALYSIS OF RESULTS

We first perform some general analysis of the data, detecting additional outliers and computing mean error for each model and condition. We also consider measurement of task error only at gauges positioned on flat surface regions in order to consider a baseline for task performance. We detail where we discovered correlations for a number of non-geometric measurements including: user and gauge persistence, gauge consistency, trial duration and number of camera rotations. (Note that all correlations we present were tested for statistical significance: $p < 0.05$.)

We then consider a number of geometric measurements one can evaluate at any gauge position, and discover in many cases the correlations between these measurements and task error. Many of these measurements are based on the geometry of surfaces: curvature, local thickness, medial axis distance, centroid distance and view-normal angle difference. Two others are unique to planar abstractions, these are: abstraction distance, and abstraction angle difference.

We detail the meaning and how we computed all of these measurements in the following subsections. The results of correlation tests between all of these measurements and task error are presented in Table III.

Table I.
Mean errors for the fixed view task (white rows) and rotated view task (grey rows). Units are in degrees.

| | surface | MPS | XYZ | radial | crdbrd |
|----------|---------|------|------|--------|--------|
| Cup | 20.7 | 50.4 | 31.5 | 62.7 | 54.7 |
| | 22.5 | 46.7 | 23.9 | 44.4 | 55.2 |
| Ant | 24.6 | 40.8 | 32.5 | 44.9 | 48.0 |
| | 24.5 | 39.9 | 30.8 | 38.0 | 47.9 |
| Table | 13.2 | 14.8 | 17.4 | 49.9 | 16.8 |
| | 13.1 | 15.0 | 13.4 | 38.5 | 26.0 |
| Camel | 26.1 | 35.9 | 28.5 | 42.2 | 40.9 |
| | 25.5 | 34.3 | 25.7 | 34.6 | 44.9 |
| Cylinder | 17.7 | 36.8 | 21.1 | 58.3 | 25.7 |
| | 16.2 | 35.2 | 15.6 | 41.5 | 31.4 |
| Bumpy | 19.2 | 38.3 | 20.8 | 59.2 | 42.0 |
| | 18.8 | 35.5 | 16.3 | 42.3 | 40.9 |
| Average | 20.2 | 36.2 | 24.3 | 52.8 | 38.0 |
| | 20.1 | 34.4 | 21.0 | 39.9 | 41.0 |

4.1 Initial Analysis

Outliers. Despite taking measures to verify that participants are well-intentioned through demonstrating consistent performance, we still encountered outliers. Looking at mean error per participant, we discovered one that had a mean error exceeding 120 degrees taken over 10 runs of the study. Such participants fall well outside the normal range, but are few. Our method to detect these outliers is: for each participant, we compare their mean error with the group of other participants with the same experiment condition (task and model representation). If the participant's mean error falls outside of three standard deviations of the group mean, the participant is deemed an outlier and their data is discarded. Using this method we classified 4 participants out of 182 as outliers, leaving the results of 178 unique participants to use in our analysis.

Error. Table I shows the mean task error for each combination of model and condition in our study. Comparing the performance between tasks, we observed a difference in mean error within about 3° in most cases. The exception to this observation was the radial representation, where the ability to rotate the viewpoint seemed to result in a significant improvement of about 13° . However for all other representations, showing and allowing viewpoint rotation does not seem to significantly reduce average error; we find this result to be of interest.

Comparing the performance between representations by averaging tasks, as expected we observed the surface representation produces the lowest error. Among the planar abstractions, the best style was XYZ. This result was expected given that this abstraction style uses the most planar sections of any abstraction. However, the worst style was radial (despite using 18 planar sections in each abstraction, more than both the MPS and crdbrd styles).

On bas-relief ambiguity. Bas-relief ambiguity [Belhumeur et al. 1999] exists when a surface shaded with Lambertian reflectance is viewed under orthographic projection. The generalized bas-relief (GBR) transformation, which consists of a scaling and shearing along the depth axis, captures the family of plausible surfaces which all produce the exact same rendering (note that a corresponding change in surface albedo is required for the shading to appear unchanged following the transformation).

In previous work by Cole et al. [2009], given a participant's gauge orientations for a specific view, the parameters of the GBR transform are optimized so that the true surface normals at the gauge

Table II.

Mean error taken only for gauge positions at flat surface regions, for the fixed view task (white rows) and rotated view task (grey rows). Units are in degrees.

| | surface | MPS | XYZ | radial | crdbrd |
|----------|---------|------|------|--------|--------|
| Table | 10.6 | 10.9 | 12.7 | 48.8 | 12.5 |
| | 9.2 | 10.5 | 8.3 | 37.1 | 22.2 |
| Cylinder | 15.5 | 10.6 | 14.4 | 48.2 | 11.6 |
| | 13.3 | 9.9 | 7.9 | 34.9 | 19.1 |
| Average | 13.1 | 10.7 | 13.5 | 48.5 | 12.0 |
| | 11.2 | 10.2 | 8.1 | 36.0 | 20.7 |

positions most closely match the participant's samples. We also experimented with this, implementing an algorithm to compute optimal GBR transform parameters via gradient descent for each of 6 independent sets of 10 participant samples. We discovered that participant error is aggressively reduced, the mean error for 10 samples decreasing typically by at least 5°.

The bas-relief ambiguity only applies to participants with a fixed view condition, as viewpoint rotation will reveal surface variation along the predefined viewpoint's depth axis. Examining the mean error between conditions (see Table I), we observe roughly consistent performance (excluding radial) to within a few degrees. For example, for the surface representation, mean errors were 20.2° and 20.1° for the fixed and rotated view conditions.

Based on this observation, we decided not to filter our data to account for the ambiguity present in the fixed view conditions. If we had done this, mean errors for fixed view conditions would be significantly lower than the equivalent rotated view condition. This would not be meaningful, as the fundamental difference between these two conditions is that viewpoint rotation provides more geometric information to participants; this should only serve to reduce error or have no effect, but should certainly not cause a significant increase.

Error at flat surface regions. Two models in our study, *Table* and *Cylinder* have large flat regions which can be used to establish a sense of baseline error for task performance. Since the geometry of flat regions are easiest to perceive, we can speculate that the error observed may be due largely to the participant's inability to perform the task accurately (that is, they may perceive the correct normal, but have limited ability in manipulating the gauge to an intended orientation). We include these findings because they may be of interest, especially for comparing error with all other gauges used in the study.

The results are shown in Table II. As expected, the mean error for flat regions alone is generally consistently lower. From these findings, we believe that the baseline error to be somewhere between 10 and 13°, excluding the radial representation and an interesting finding where the rotated view crdbrd condition resulted in 8.7° more error than the fixed view crdbrd condition. We attribute the increased error of the rotated view crdbrd task to a difference in ability between the groups of participants selected for the conditions (17 participants for the fixed view crdbrd condition, 18 participants for the rotated view crdbrd condition).

On comparing abstractions. It is worthwhile to note that a direct comparison between the errors resulting from different styles of planar abstraction may not be meaningful as two abstractions may vary in other important ways, and may be created with any other requirement in mind than task error minimization. For instance, some styles such as XYZ use many planar sections, while others such as MPS and crdbrd are more frugal. Another consideration may be the surface area (or amount of material, in a physical sense) used. Some abstractions have a configuration that may or may not be *realizable*: meaning the abstraction is physically assemblable by creating slots and sliding them

Table III.

A table showing Pearson's correlation coefficient r for correlations between task error and a variety of other measurements. Rows show correlations for a specific condition, columns show correlations for a specific measurement with error. Red entries indicate any correlation which was very low ($|r| < 0.1$) or not statistically significant ($p \geq 0.05$). The three groups of measurements from left to right are: general, surface-specific and abstraction-specific.

| Condition | user persistence | gauge persistence | gauge consistency | trial duration | number of view rotations | absolute κ_1 | absolute κ_2 | Gaussian curvature | mean curvature | local thickness | medial axis distance | centroid distance | view-norm angle difference | abstraction distance | abstraction angle difference |
|-----------------|------------------|-------------------|-------------------|----------------|--------------------------|---------------------|---------------------|--------------------|----------------|-----------------|----------------------|-------------------|----------------------------|----------------------|------------------------------|
| Fixed surface | 0.95 | 0.17 | 0.51 | | N/A | 0.51 | 0.42 | 0.44 | 0.49 | | -0.13 | | -0.17 | N/A | N/A |
| MPS | | 0.31 | 0.51 | | N/A | 0.42 | 0.26 | 0.19 | 0.42 | 0.49 | -0.20 | 0.23 | | 0.39 | 0.72 |
| XYZ | 0.47 | 0.21 | 0.59 | | N/A | 0.54 | 0.42 | 0.46 | 0.53 | | -0.24 | 0.11 | -0.14 | 0.13 | -0.26 |
| radial | | 0.25 | | 0.15 | N/A | | -0.14 | -0.16 | | 0.39 | -0.10 | 0.28 | 0.21 | 0.20 | 0.30 |
| crdbrd | 0.52 | 0.33 | 0.49 | | N/A | 0.46 | 0.41 | 0.33 | 0.46 | 0.14 | -0.16 | 0.15 | | 0.17 | 0.43 |
| Rotated surface | 0.57 | 0.25 | 0.59 | | | 0.60 | 0.50 | 0.53 | 0.57 | | -0.21 | | -0.19 | N/A | N/A |
| MPS | | 0.31 | 0.50 | 0.17 | | 0.40 | 0.25 | 0.18 | 0.39 | 0.46 | -0.20 | 0.21 | -0.13 | 0.37 | 0.72 |
| XYZ | 0.63 | 0.22 | 0.57 | | | 0.64 | 0.50 | 0.56 | 0.62 | | -0.28 | | -0.22 | | -0.35 |
| radial | | 0.18 | 0.45 | | | 0.12 | | | 0.11 | 0.22 | -0.16 | 0.21 | 0.15 | 0.19 | 0.17 |
| crdbrd | | 0.15 | 0.42 | | | 0.43 | 0.37 | 0.31 | 0.42 | | -0.20 | 0.10 | | | 0.43 |

■ ≥ 0.5
■ ≥ 0.3
■ ≥ 0.1
■ ≥ 0.0
■ N/A not applicable

together [Hildebrand et al. 2012]. Thus a direct error-based comparison between one abstraction style and another to determine which is “better” may not be a fair one as it ignores properties such as these which may be of importance.

4.2 Participant-Specific Correlations

Persistence. Participants are given co-located gauge pairs to set as part of our approach to validation. The measurement of the difference in angle between the two samples for the gauge pair we call the *persistence* measurement. Mean persistence can be computed for a participant, or for a gauge. In general, we expect persistence measurements to be below 30° since the main criteria for accepting a participant's results is that persistence is below 30° at least 70% of the time. Despite this, persistence measurements sometimes do exceed 30° and the variation in persistence may lead to some interesting correlations.

Testing for correlation between persistence and error averaged over each participant, we discovered some strong correlations of statistical significance for the surface and XYZ conditions, and also the fixed view crdbrd condition. We also tested correlation between mean persistence and error for gauges, but found correlations were generally small. Correlation results for both measurements are shown in Table III.

Consistency. We explore the notion of *consistency* of performance at a gauge position. To do this, we measure the standard deviation of task error for samples taken at the gauge position. We tested for correlation between consistency and error and discovered that generally, there was strong correlation. Since consistency and error share a clearly evident linear dependence in most cases, for any given gauge position where mean error of the samples is low, the samples will be consistent, and vice versa. For the surface, MPS and XYZ abstraction styles, consistency and error are strongly correlated for both tasks.

Trial duration. In our study, we measured the duration in milliseconds it takes for the participant to complete the orientation task for each presented gauge. We tested for correlation between mean trial duration at the gauge and mean error at the gauge in order to determine if there was any relationship. When testing for correlation between mean trial duration and mean error at a gauge position, we discarded those samples where the trial duration exceeded 15 seconds; in our data there were even rare trials where the trial duration exceeded five minutes (likely a participant taking a break mid-run).

We predicted that an inverse correlation between trial duration and mean error might exist. Upon testing we were surprised to find that generally, there was no correlation between trial duration and error (see Table III).

Number of view rotations. Only applicable to the rotated view task, we measured the total number of view rotations each participant performed and tested for correlation with their mean error. We discovered no correlations of statistical significance for any of the five applicable conditions. We also tested normalizing the number of view rotations by the number of runs of the study the participant completed, in order to account for the variation, but still discovered no correlations of statistical significance (see Table III).

4.3 Surface-Specific Correlations

For the following correlations, we aim to maximize the coefficient of determination R^2 . To do this, for each measurement we considered the class of *power transformations* which include logarithm, square root, reciprocal, etc., as instances. Specifically, we found that curvature-related measurements had correlation maximized when taking the *fourth root*, and the local thickness measurement had maximum correlation when *squared*.

Curvature. We tested for correlation between curvature at a gauge position on the surface and the mean error at the gauge position. To evaluate curvature at surface positions, we used the method of Rusinkiewicz [2004]. We consider four different measurements: absolute principal curvatures κ_1 and κ_2 , Gaussian curvature and mean curvature. As expected, we found strong correlations with error across many conditions, the correlation coefficients can be seen in Table III.

Local thickness. We were curious about a relationship existing between a notion of *thickness* at a surface point and task error there. To estimate local thickness, we implemented a ray collision test: we emit a ray from the gauge position in a direction opposite to the surface normal, intersecting it with the surface, and use the nearest intersection point whose distance is nonzero. We use the distance to the nearest intersection point as our measure of local thickness. Testing for correlation between local thickness and mean error, we discovered positive correlations for the MPS and radial styles of abstraction for both fixed and rotated view tasks, as shown in Table III.

Medial axis distance. To compute an approximate medial axis for each model, we implemented the approach of Palagyi [2002]. To compute our medial axis distance measurement for a gauge position, we use the smallest distance to a point on the medial axis. We tested for correlation between the medial axis distance measurement and the mean gauge error. We expected there may be an inverse correlation between them and for all conditions, we discovered a small inverse correlation. These results are shown in Table III.

Centroid distance. The centroid is computed for each model, and we evaluate the distance between the centroid and each surface point. Checking correlation results in Table III, we found statistically significant small positive correlations for the majority of conditions, suggesting that error increases at points further from the centroid of the surface.

View-normal angle difference. We measured the angle between the view direction and the (back-facing) surface normal at each position. As shown in Table III, we found statistically significant small negative correlations. These correlations indicate the tendency for error to increase as the surface normal becomes parallel with the view direction. The exceptions to this general observation are the conditions with the radial representation—since the radial abstraction style has the most “geometric coverage” at the center of models and less where the surface’s silhouette will be from the given view, the view-normal angle difference is high.

4.4 Abstraction-Specific Correlations

Abstraction distance. For each gauge position, we computed the distance to the nearest point on the planar abstraction. We call this measurement the *abstraction distance*. Styles of planar abstraction which use the fewest planes in their model representation (such as MPS or crdbrd) we expect to exhibit the highest variance for this measurement. We tested for correlation between abstraction distance and mean gauge error for the conditions where abstractions are used. We discovered that for the MPS style abstraction that the abstraction distance most correlated with error, with a few other small correlations as shown in Table III.

Abstraction angle difference. As with abstraction distance, our method to compute the abstraction angle difference measurement first involves finding the nearest point on the abstraction to the gauge position. The nearest point lies on a planar section which has a constant surface normal, we call this the *abstraction normal*. To compute the *abstraction angle difference* measurement, we compute the angular difference between the abstraction normal and the true surface normal at the gauge position. We tested for correlation between abstraction angle difference and mean error, and discovered significant correlations of varying strength for all 8 conditions. Of these, the correlations for the MPS style were especially strong and medium strength correlations were also found for the crdbrd conditions. Conversely, for XYZ and radial abstractions, because abstraction angle differences are often very high (distributions have a mean closer to 90 degrees and smaller variance), correlations with error are not strong. The correlation coefficients are shown in the last column of Table III.

5. PREDICTING ERROR

The task faced is one of *supervised learning*, where measurements from a number of independent variables or *predictors* are used to predict the dependent outcome variable or *response*. We use the study data for *learning*, that is, we create a predictive model from it. In our study, the predictors are the measurements that can be made at surface points (such as curvature or medial axis distance), and the response is the task error measured at that point.

5.1 Linear Models

Given the variety of strong linear correlations discovered in the data and presented in Section 4, we chose to use a linear predictive model. Linear models tend to be more stable than nearest neighbor models, but are more biased. We use the least squares error criterion to measure model fitness.

Expressed as a function $f : \mathbb{R}^p \rightarrow \mathbb{R}$, linear models take the following form:

$$f(\mathbf{x}) = \hat{\beta}_0 + \sum_{j=1}^p \hat{\beta}_j h_j(\mathbf{x}),$$

where p is the number of predictor variables, \mathbf{x} is a p -vector of input predictor measurements, $h_j : \mathbb{R}^p \rightarrow \mathbb{R}$ is predictor j 's transformation function and $\hat{\beta}$ is a $p + 1$ -vector of linear parameters for the model ($\hat{\beta}_0$ is the *intercept* or *bias* term).

5.2 Regularization

We perform *regularization*—determining which predictors significantly contribute to the predictive model and eliminating those that do not. Regularization prevents over-fitting and improves generalization by simplifying the model.

We use the LASSO method [Tibshirani 1996] to select predictors. The LASSO method works by penalizing the \mathcal{L}_1 -norm of $\hat{\beta}$. A parameter λ can be used to adjust the amount of penalization, tradi-

Table IV.

A table revealing parameters for each condition's linear model, and cross-validated error measurements. The linear model parameters $\hat{\beta}_0$ to $\hat{\beta}_{10}$ correspond to: (0) intercept, (1) fourth-root absolute principal curvature κ_1 , (2) fourth-root absolute principal curvature κ_2 , (3) fourth-root absolute Gaussian curvature, (4) fourth-root absolute mean curvature, (5) medial axis distance, (6) squared local thickness, (7) centroid distance, (8) view direction-normal angle difference, (9) abstraction distance and (10) abstraction angle difference. Cross-validated errors include mean absolute predicted error (MAPE) and mean absolute base error (MABE). The final column expresses the percentage improvement resulting from model predictors, using the ratio of MAPE to MABE.

| Condition | $\hat{\beta}_0$ | $\hat{\beta}_1$ | $\hat{\beta}_2$ | $\hat{\beta}_3$ | $\hat{\beta}_4$ | $\hat{\beta}_5$ | $\hat{\beta}_6$ | $\hat{\beta}_7$ | $\hat{\beta}_8$ | $\hat{\beta}_9$ | $\hat{\beta}_{10}$ | MAPE | MABE | Improved |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|--------------------|------|------|----------|
| Fixed surface | 13.8 | 4.69 | . | . | . | . | . | . | . | — | — | 6.43 | 7.64 | 16% |
| MPS | 9.07 | 4.75 | . | . | . | . | 40.9 | . | 0.02 | 126 | 0.27 | 10.2 | 19.5 | 47% |
| XYZ | 16.0 | 6.79 | . | . | . | . | . | . | . | . | . | 9.08 | 11.0 | 18% |
| radial | 27.0 | . | . | . | . | −37.3 | 29.5 | 14.3 | 0.12 | 194 | 0.15 | 11.6 | 14.3 | 19% |
| crdbrd | 9.75 | . | 4.53 | . | 7.83 | . | 12.0 | . | 0.02 | 209 | 0.19 | 14.1 | 19.2 | 27% |
| Rotated surface | 11.9 | 5.69 | . | 0.26 | . | . | . | . | . | — | — | 5.81 | 7.74 | 25% |
| MPS | 10.25 | 3.59 | . | . | . | . | 33.3 | . | . | 111 | 0.29 | 10.5 | 19.0 | 45% |
| XYZ | 10.7 | 7.73 | . | 0.40 | . | . | . | . | . | . | −0.01 | 7.36 | 10.7 | 32% |
| radial | 19.5 | 1.91 | . | . | . | −44.3 | 10.6 | 6.36 | 0.10 | 177 | 0.11 | 9.98 | 11.3 | 12% |
| crdbrd | 9.23 | 7.90 | 8.08 | −3.01 | . | −9.44 | 2.72 | . | 0.14 | 139 | 0.18 | 12.8 | 16.6 | 23% |

tionally a value of λ that is within one standard error of the minimal mean squared error is selected; this choice is known to be *parsimonious* and builds the model with as few predictors as possible, while keeping the mean squared error of the fit low.

5.3 Validation

We perform k -fold cross-validation using $k = 10$, which is known as a good compromise between bias and variance (e.g., leave-one-out cross-validation, where $k = n$, exhibits low bias but can have high variance). The n input samples are split up into 10 groups or *folds* of approximately equal size. We train the model with data from 9 folds and use the last one for testing. This process is repeated 10 times, so that each fold gets used for testing once. We are careful to retrain the model each time using only the folds selected for training.

The *estimated prediction errors* for the model are evaluated as the average of the mean absolute errors over all 10 folds. We can compare this value to the *base error*, which is the error resulting from using the mean of the responses in each training set as the predictive model. The ratio between estimated prediction and base errors reveals how much the predictors collectively improve each linear model. Table IV shows cross-validated error statistics for models fit to each condition.

We used the R statistical computing software [R Core Team 2012] and the *glmnet* package [Friedman et al. 2010] to perform the regularization and validation steps.

5.4 Models and Performance

The purpose of our models is to predict the value of the response: the task error. Given a surface representation, the 8 relevant predictors are: curvature (κ_1 , κ_2 , Gaussian and mean), local thickness, medial axis distance, centroid distance and view-normal angle difference. Given a planar abstraction based on a (hidden) surface representation, there are two additional predictors: abstraction distance and abstraction angle difference. Note that all of our predictors capture some geometric property of the representation at each surface point; we do not incorporate nongeometric measurements from the first group of columns in Table III (e.g., trial duration, measures of persistence, etc.).

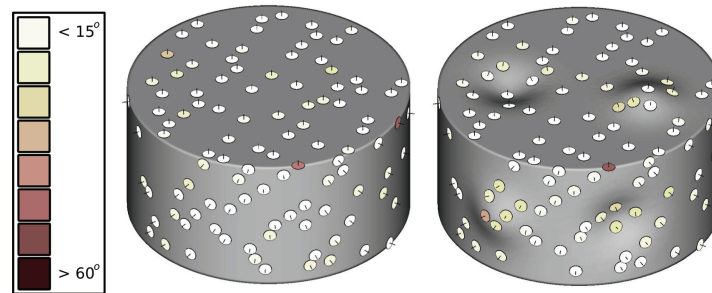


Fig. 4. A visualization of task error for the *Cylinder* and *Bumpy* models and fixed view surface condition. For these surfaces, error is generally low for locally flat and locally cylindrical regions, but not for locally convex or concave regions.

Predictor selection. Table IV shows all $\hat{\beta}$ parameters for our linear models built for each of the 10 conditions. Regularization causes numerous predictors to be excluded from the model in many cases—the corresponding $\hat{\beta}$ value of the predictor is set to zero—and this is indicated with an “.” entry. (Note that although ideal, $\hat{\beta}$ values cannot be used for relative comparison of predictor importance, even if predictor scales are standardized. However, the inclusion of the predictor in the model does ensure it to be of some significance in characterizing task error for the condition, which leads to some insights.)

The absolute principal curvature κ_1 was the most frequently selected predictor across the conditions. Our intuition was that high curvature—that is, high local variation of the surface normal—would result in higher task error. The presence of the κ_1 predictor in most of the predictive models generated is evidence in support of that claim.

In contrast, the principal curvature κ_2 predictor was not of nearly as much use in our predictive models. We expected κ_2 to be of greater significance: consider the error visualizations presented in Figure 4. In these examples, both locally flat and cylindrical ($\kappa_2 = 0$) surface regions generally have low error, but the convex and concave ($\kappa_2 \neq 0$) surface regions do not. Therefore, one may expect κ_2 would be an ideal predictor since it varies with error. However, since $\kappa_2 \neq 0$ only when $\kappa_1 \neq 0$, and κ_1 correlates well with error when curvature is especially high, κ_1 is the more useful predictor.

A few other predictors we selected also did not find much use. For instance, the centroid distance predictor was only of use in the radial abstraction style, where error appears to positively correlate with distance from a central axis formed at the intersection of the planar sections (see Figure 6). The mean curvature predictor was hardly utilized—since we have predictors for each individual principal curvature, there is little advantage to be had in combining them through addition to form a new predictor, as linear models do this already.

The abstraction-specific predictors, distance and angle difference, found use in all predictive models for abstractions except for the XYZ style. Since the XYZ style results in a “dense” geometric abstraction where surface-to-abstraction distances are relatively very small and abstraction angle differences are generally 90 degrees, these predictors will have little if any correlation with error as a result. We believe these predictors are increasingly useful if the surface measurements can increase in variability—which happens naturally as the number of planar sections in the abstraction decreases.

Performance. We shift our perspective from predictors to the individual models, and examine their estimated prediction errors relative to base error. Figure 5 shows scatterplots for the predicted error vs observed error for all conditions in our study, and the final columns of Table IV show mean absolute predicted error and mean absolute base error for each model. The final column expresses the percentage improvement over base error—a result of the predictors selected.

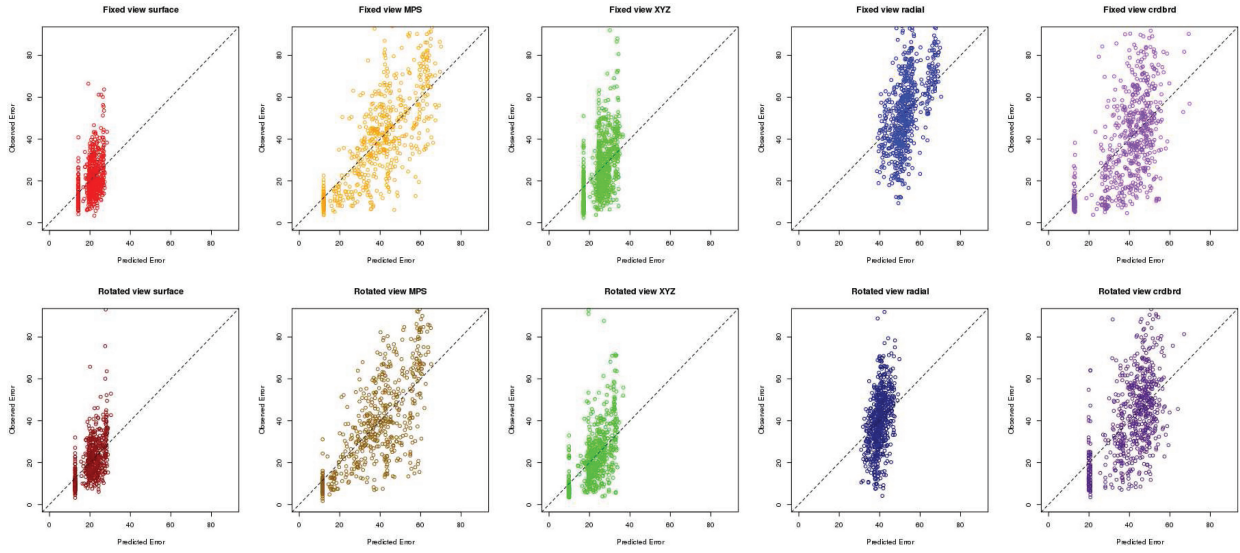


Fig. 5. Scatterplots for each condition comparing predicted and observed task error. Each point represents an observation at a gauge position, the distance of the point from the dashed diagonal line indicates error of the linear model in prediction. The vertical distribution of points indicates the task error, while the horizontal distribution of points reveals the significance of the predictors used in the model—high horizontal variation also tends to correspond with increased improvement over base error rate. These scatterplots reveal that predictors yield the greatest improvement over base error for the MPS models, as the point distributions both tend most toward the diagonal line and exhibit high horizontal variation. For comparison, although surface model point distributions are even closer to the diagonal line and this indicates lower estimated prediction error, the lack of horizontal variation reveals improvements over base error are not as great.

The least improvement in error was 12%, by the model for the rotated view radial condition. This result is to be expected, as none of the predictors we selected had a medium or higher correlation ($|r| > 0.3$) for this condition. Note that while there is a 19% improvement for the fixed view radial condition, the mean errors are significantly higher as well and provide greater room for improvement. In the scatterplot for the rotated view radial condition in Figure 5, the distribution of points is very narrow horizontally, indicating that our predictive model performs comparably to simply using mean error as the prediction.

The greatest improvements in error were 47% and 45% for the fixed and rotated view MPS conditions, which are substantial. On average, we estimate our predictive models will be approximately 10° from the true error. Considering the variation of error in data collected for the MPS abstraction, and that the base error is approximately 19° , our models make relatively good predictions and capture the geometric sources of error. Since the MPS abstraction style also had the highest base errors at approximately 19° , there was the greatest potential for improvement.

5.5 Improving Abstractions

We ran a separate substudy to exemplify the potential of our predictive models to improve surface perception by modifying the process of planar abstraction creation. Without much difficulty, we were able to incorporate the positively-valued abstraction distance predictor of the MPS predictive models into the MPS algorithm of McCrae et al. [2011], so planes that minimize the average abstraction distance would be considered in the selection process (see Figure 7). Running a second smaller

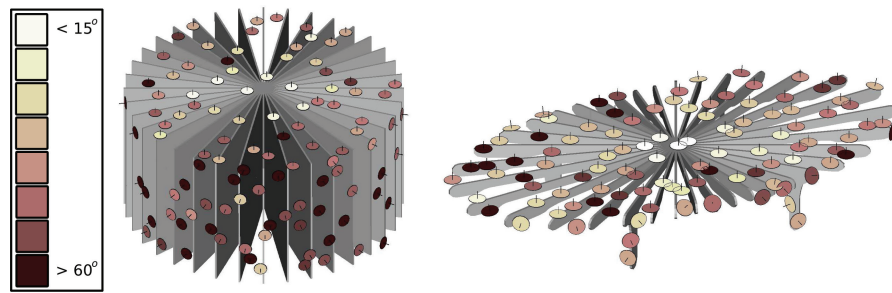


Fig. 6. A visualization of task error for the *Cylinder* and *Table* models and fixed view radial condition. Note the positive correlation between error and distance from the central axis formed by this abstraction style.

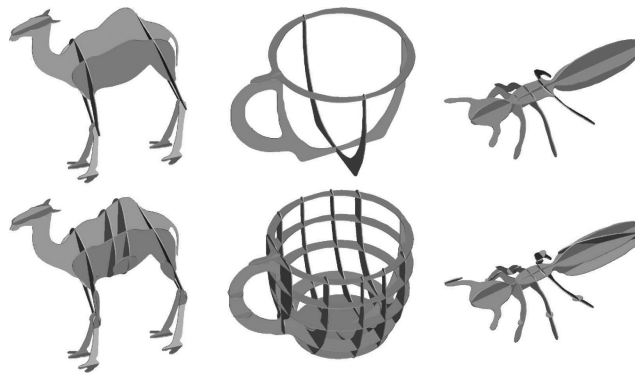


Fig. 7. (Top) The MPS style abstraction of models *Camel*, *Cup* and *Ant*. (Bottom) Our “modified MPS” abstractions, which additionally seek to minimize the abstraction distance predictor.

crowd-sourced study on AMT for these 3 models, we discovered significant reductions in mean task error: *Cup* by 52%, *Camel* by 18% and *Ant* by 14%.

6. CONCLUSION

We have conducted a large-scale study involving gauge orientation to define perceived surface normals, which expands upon previous methods by incorporating conditions where viewpoint rotation is encouraged. We considered a set of geometric measurements—many of which exhibited strong linear correlation with task error—and used these as predictors in linear models. The chosen parameters to the linear models themselves provide insight into the geometric causes of task error, and the models can be used as an automatic means to evaluate perception at a given point (for either surfaces, or for planar abstractions). Finally, we demonstrate in a second substudy that the predictive models can be used to augment existing planar abstraction creation algorithms in order to improve perception of the abstracted surface.

ACKNOWLEDGMENTS

We thank Vangelis Kalogerakis and W. John Braun for helpful correspondence, especially concerning our approach to regularization of model predictors. We thank Kristian Hildebrand and Dongming Yan for assisting us by providing some of the geometric representations. We thank Daniela Giorgi and AIM@SHAPE for providing meshes from the Watertight Track of SHREC 2007.

REFERENCES

- BELHUMEUR, P. N., KRIEGMAN, D. J., AND YUILLE, A. L. 1999. The bas-relief ambiguity. *Int. J. Comput. Vision* 35, 1, 33–44.
- CANIARD, F. AND FLEMING, R. W. 2007. Distortion in 3D shape estimation with changes in illumination. In *Proceedings of the Symposium on Applied Perception in Graphics and Visualization*. 99–105.
- CIGNONI, P., MONTANI, C., AND SCOPIGNO, R. 1997. A comparison of mesh simplification algorithms. *Comput. Graphics* 22, 37–54.
- COHEN-STEINER, D., ALLIEZ, P., AND DESBRUN, M. 2004. Variational shape approximation. In *Proceedings of the ACM SIGGRAPH International Conference on Computer Graphics and Interactive Techniques*. 905–914.
- COLE, F., GOLOVINSKIY, A., LIMPAECHER, A., BARROS, H. S., FINKELSTEIN, A., FUNKHOUSER, T., AND RUSINKIEWICZ, S. 2012. Where do people draw lines? *Comm. ACM* 55, 1.
- COLE, F., SANIK, K., DECARLO, D., FINKELSTEIN, A., FUNKHOUSER, T., RUSINKIEWICZ, S., AND SINGH, M. 2009. How well do line drawings depict shape? *ACM Trans. Graph* 28, 3.
- DÉCORET, X., DURAND, F., SILLION, F. X., AND DORSEY, J. 2003. Billboard clouds for extreme model simplification. *ACM Trans. Graph* 22, 3, 689–696.
- FAISMAN, A. AND LANGER, M. S. 2012. Perception of qualitative shape from diffuse and specular reflections. In *Proceedings of the ACM Symposium on Applied Perception*. 123–123.
- FRIEDMAN, J., HASTIE, T., AND TIBSHIRANI, R. 2010. Regularization paths for generalized linear models via coordinate descent. *J. Statist. Softw.* 33, 1, 1–22.
- GIORGI, D., BIASOTTI, S., AND PARABOSCHI, L. 2007. SHREC: SHape REtrieval contest: Watertight models track.
- HILDEBRAND, K., BICKEL, B., AND ALEXA, M. 2012. crdbd : Shape fabrication by sliding planar slices. *CGF Eurographics* 31, 2.
- KAVAN, L., DOBBYN, S., COLLINS, S., ŽÁRA, J., AND O’SULLIVAN, C. 2008. Polypostors: 2D polygonal impostors for 3D crowds. In *Proceedings of the Symposium on Interactive 3D Graphics and Games*. 149–155.
- KNILL, D. C. 1992. Perception of surface contours and surface shape: from computation to psychophysics. *J. Opt. Soc. Am.* 9, 9, 1449–1464.
- KOENDERINK, J. J., DOORN, A. J. V., AND KAPPERS, A. M. L. 1992. Surface perception in pictures. *Percept. Psychophysics*, 487–496.
- LI, X.-Y., JU, T., GU, Y., AND HU, S.-M. 2011. A geometric study of v-style pop-ups: theories and algorithms. *ACM Trans. Graph.* 30, 6, 98:1–98:10.
- MCCRAE, J., SINGH, K., AND MITRA, N. J. 2011. Slices: a shape-proxy based on planar sections. *ACM Trans. Graph.* 30, 6, 168:1–168:12.
- MEHRA, R., ZHOU, Q., LONG, J., SHEFFER, A., GOOCH, A., AND MITRA, N. J. 2009. Abstraction of man-made shapes. *ACM Trans. Graph.* 28, 5, 137:1–137:10.
- MITRA, N. J., GUIBAS, L. J., AND PAULY, M. 2006. Partial and approximate symmetry detection for 3D geometry. *ACM Trans. Graph.* 25, 3, 560–568.
- NORMAN, D., TODD, J. T., AND PHILLIPS, F. 1995. The perception of surface orientation from multiple sources of optical information. *Percept. Psychophysics*, 629–636.
- O’SHEA, J. P., BANKS, M. S., AND AGRAWALA, M. 2008. The assumed light direction for perceiving shape from shading. In *Proceedings of the Conference on Applied Perception in Graphics and Visualization*. 135–142.
- PALÁGYI, K. 2002. A 3-subiteration 3d thinning algorithm for extracting medial surfaces. *Pattern Recogn. Lett.* 23, 6, 663–675.
- PHILLIPS, F., TODD, J., KOENDERINK, J., AND KAPPERS, A. 2003. Perceptual representation of visible surfaces. *Percept. Psychophysics* 65, 747–762.
- R CORE TEAM. 2012. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- RAMACHANDRAN, V. 1988. Perceived shape from shading. *Sci. Amer.* 76–83.
- RUSINKIEWICZ, S. 2004. Estimating curvatures and their derivatives on triangle meshes. In *Proceedings of the 2nd International Symposium on 3D Data Processing, Visualization and Transmission*.
- SAITO, T. AND TAKAHASHI, T. 1990. Comprehensible rendering of 3-d shapes. In *Proceedings of the 17th Annual Conference on Computer Graphics and Interactive Techniques*.
- SCHWARTZBURG, Y. AND PAULY, M. 2011. Design and optimization of orthogonally intersecting planar surfaces. In *Proceedings of the Design Modeling Symposium*.
- SECORD, A., LU, J., FINKELSTEIN, A., SINGH, M., AND NEALEN, A. 2011. Perceptual models of viewpoint preference. *ACM Trans. Graph.* 30, 5.
- SHAO, C., BOUSSEAU, A., SHEFFER, A., AND SINGH, K. 2012. Crossshade: shading concept sketches using cross-section curves. *ACM Trans. Graph.* 31, 4, 45:1–45:11.

- SIMARI, P., KALOGERAKIS, E., AND SINGH, K. 2006. Folding meshes: Hierarchical mesh segmentation based on planar symmetry. In *Proceedings of the Eurographics Symposium on Geometry Processing*.
- STEVENS, K. A. 1981. The visual interpretation of surface contours. *Artif. Intell.* 17, 47–73.
- SWEET, G. AND WARE, C. 2004. View direction, surface orientation and texture orientation for perception of surface shape. In *Proceedings of Graphics Interface 2004*. 97–106.
- TIBSHIRANI, R. 1996. Regression shrinkage and selection via the lasso. *J. Roy. Statist. Soc. Ser. B* 58, 1, 267–288.
- WILLIS, K. D., LIN, J., MITANI, J., AND IGARASHI, T. 2010. Spatial sketch: Bridging between movement and fabrication. In *Proceedings of the 4th International Conference on Tangible, Embedded, and Embodied Interaction*. 5–12.
- WINNEMÖLLER, H., FENG, D., GOOCH, B., AND SUZUKI, S. 2007. Using NPR to evaluate perceptual shape cues in dynamic environments. In *Proceedings of the 5th International Symposium on Non-Photorealistic Animation and Rendering*. 85–92.
- ZHANG, J., ZHENG, J., WU, C., AND CAI, J. 2012. Variational mesh decomposition. *ACM Trans. Graph.* 31, 3, Article 21.

Received June 2013; accepted July 2013