Vision Research 115 (2015) 218–230

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Effects of surface reflectance on local second order shape estimation in dynamic scenes

Dicle N. Dövencioğlu [a,b,*], Maarten W.A. Wijntjes [c], Ohad Ben-Shahar [d], Katja Doerschner [a,b,e]

[a] Department of Psychology, Bilkent University, Ankara, Turkey
[b] National Magnetic Resonance Research Center, Bilkent University, Ankara, Turkey
[c] Perceptual Intelligence Lab & Department of Industrial Engineering, Delft University of Technology, Delft, Netherlands
[d] Computer Science Department, Ben Gurion University of the Negev, Beer-Sheva, Israel
[e] Department of Psychology, Justus-Liebig-University Giessen, Giessen, Germany

ARTICLE INFO

Article history:
Received 10 July 2014
Received in revised form 24 December 2014
Available online 30 January 2015

Keywords:
Shape from motion
Shape from specular flow
Surface materials

ABSTRACT

In dynamic scenes, relative motion between the object, the observer, and/or the environment projects as dynamic visual information onto the retina (optic flow) that facilitates 3D shape perception. When the object is diffusely reflective, e.g. a matte painted surface, this optic flow is directly linked to object shape, a property found at the foundations of most traditional shape-from-motion (SfM) schemes. When the object is specular, the corresponding *specular flow* is related to shape *curvature*, a regime change that challenges the visual system to determine concurrently both the shape and the distortions of the (sometimes unknown) environment reflected from its surface. While human observers are able to judge the global 3D shape of most specular objects, shape-from-specular-flow (SFSF) is not veridical. In fact, recent studies have also shown systematic biases in the perceived motion of such objects. Here we focus on the perception of *local shape from specular flow* and compare it to that of matte-textured rotating objects. Observers judged local surface shape by adjusting *a rotation and scale invariant shape index probe*. Compared to shape judgments of static objects we find that object motion decreases intra-observer variability in local shape estimation. Moreover, object motion introduces systematic changes in perceived shape between matte-textured and specular conditions. Taken together, this study provides a new insight toward the contribution of motion and surface material to local shape perception.

© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

The internal representation of the 3D geometry of physical objects is vital for allowing humans to interact with the physical world. From threading a needle, grasping a coffee cup, or carefully stepping along a staircase, the estimation of 3D shape from visual cues covers much of our daily activities. Since the image reflected off an object depends critically on its surface properties, our visual system is faced with a material-shape ambiguity that in practice requires the estimation of both. This ambiguity is particularly severe for specular objects and unknown illumination environments since it is straight forward to manipulate both the shape and the environment to generate any given image (e.g., Fleming, Drorr, & Adelson, 2003). One strategy employed by humans for better observing and reconstructing the 3D shape of objects includes the incorporation of relative motion between the object, the environment, and the observer (e.g., by moving the object or the observer's head or body). Such motion projects as dynamic visual information onto the retina (optic flow) and has been shown to improve or facilitate 3D shape perception (Wallach & O'connell, 1953; Landy et al., 1991; Bradley, Chang, & Andersen, 1998). These studies most often assume that the surface of the object is matte (i.e., diffusely reflecting), a condition that links the optic flow directly to first order shape properties. When the object is specular, however, the corresponding *specular flow* is tightly related to object *curvature* (Koenderink & Van Doorn, 1980; Adato et al., 2010), a property that manifests itself even in partially specular objects (e.g., sweet peppers or other types of fruits). This regime change is further complicated by dependency on the nature of motion, an effect that was analyzed computationally (Adato et al., 2010) and observed perceptually (Hartung & Kersten, 2002; Sakano & Ando, 2008; Doerschner et al., 2011). It is now clear that specular flow behaves very differently from other types of optical flow (e.g. Adato, Zickler, & Ben-Shahar, 2011; Adato & Ben-Shahar, 2011), leaving even its robust estimation an open question.

Previous work by our group has shown that shape-associated estimates for *moving* specular objects are qualitatively different

* Corresponding author at: National Magnetic Resonance Research Center (UMRAM), Bilkent University, Cyberpark, 06800, Ankara, Turkey.

http://dx.doi.org/10.1016/j.visres.2015.01.008
0042-6989/© 2015 The Authors. Published by Elsevier Ltd.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

than those for matte-textured ones. For example, specular flow may bias the perceived object rotation axis (Doerschner et al., 2013) or may produce illusory non-rigid percepts (Doerschner & Kersten, 2007; Doerschner, Kersten, & Schrater, 2011). The perception of local shape from *static* specular reflections has also been studied but are currently inconclusive (e.g., compare the different conclusions in Fleming, Torralba, and Adelson (2004), Savarese, Fei-Fei, and Perona (2004)). In contrast, no systematic exploration of local shape estimation has been done for *moving* specular objects. This paper describes the first step in this direction.

The study of local shape perception of moving objects presents a challenge: Most related studies have used the attitude probe (Koenderink, Van Doorn, & Kappers, 1992), which is often referred to as 'gauge figure probe'. This method is relatively easy to implement in static image stimuli (Wijntjes, 2012), but for rotating objects it becomes problematic. When the stimulus is rotating, the attitude probe has to rotate along, giving away potentially valuable information to the observer. This problem arises because the attitude probe assesses the first order local shape which is not rotation invariant. Therefore, one needs a rotationally invariant shape probe to measure the 3D perception of a rotating stimulus. To meet this challenge we employ a novel, second order shape probe based on the shape index described by Koenderink and van Doorn (1992), Koenderink, van Doorn, and Wagemans (2014).

Using this probe, we address two main questions– can observers estimate the local shape of moving objects and are there any differences in perceived local shape between moving matte-textured and specularly-reflecting objects. The two experiments that follow answer both questions in the affirmative. Taken together, our study provides new insights into the role of motion and surface material in local second order shape perception.

2. Methods

2.1. Overview

We conducted two behavioral experiments to investigate the effect of surface reflectance on perceived local surface shape in dynamic scenes. Object reflectance properties (matte-textured vs. specular) and their motion (rotating vs. static) was varied within and across experimental sessions, respectively. Observers used a rotation invariant shape index probe to make categorical shape judgments of local surface patches marked on the object. We hypothesized that surface reflectance may modulate perceived local shape in dynamic scenes. In the analysis we correlated observers' shape estimates with theoretical surface curvature values and computed several metrics that capture variation in shape estimation. Static conditions served as baseline measures and we expected to find no systematic differences in shape estimation across surface materials for those trials.

In *Experiment 1* we studied local shape perception of simple objects, where the border – as for any smooth and compact shape – serves as boundary condition for local shape inference. At the boundary, the shape is either convex or saddle-shaped, depending on the curvature of the occluding boundary, and this (categorically) unambiguous information may help observers to integrate shape information from outside to inside. However, when the shape is complex and its variations more numerous, this integration process becomes more difficult and prone to errors. Thus, to diminish the potential influence of the occluding boundary on local shape perception, we conducted a second experiment similar to the first but with a more complex object. In general, we expected observers' shape estimates to deviate more from theoretical curvature values towards the interior of such a complex shape – while not necessarily being less consistent in their shape estimates across conditions – compared to judgments made for simple objects. We also expected to find changes in perceived local shape across material conditions. Details of stimuli, experimental setup, procedure and analysis are described next.

2.2. Stimuli

Shape. In *Experiment 1* we used two parametric shapes ('furrow' and 'dimple' as defined in Koenderink and Van Doorn (1980)[1], and used in Nefs, Koenderink, and Kappers (2006)), that were shown to observers from both generic and more accidental viewpoints (Fig. 1A and B). We varied viewpoint in order to obtain shape judgments from different parts of the object and to investigate whether shape estimates for a given location change under different viewing conditions (dots 1 and 3 in Fig. 1A and B, respectively). Simple objects subtended approximately 11 degrees visual angle. To generate the novel complex object in *Experiment 2* (Fig. 1C) we applied randomly determined sinusoidal perturbations to a sphere using 3DS MAX ©Autodesk. This object subtended approximately 9.5 degrees visual angle and, unlike the parametric shapes, it was not symmetric.

Motion. Dynamic trials presented objects rotating back and forth 10 degrees at a rate of 0.2 degrees/frame and 60 frames per second (12 degrees/s) around the vertical axis (Movies can be found at katja.bilkent.edu.tr/localshape).

Reflectance. The 3D shapes were rendered using *environment mapping* (Debevec, 2002). In this technique, the light arriving from all directions at a point in 3D space is captured and stored in a *light probe*, a high dynamic-range spherical image. This image can subsequently be used as an extended light source infinitely far away from the to-be-rendered object. As with any 2D image, the spatial structure of these light probes can be manipulated, though care must be practiced to properly define the desired operations on a sphere. The focus of our study was reflectance-specific motion and not the color or spatial content of the reflected light probe. We therefore desaturated and phase-scrambled (using spherical harmonics) Debevec's *Grace* probe (Debevec, 2002) before using it to render our stimuli.

When objects were *static*, their rendered image was consistent with that of a specularly-reflecting object, e.g., with characteristic compressions at high curvature points. However, the specific manipulations we applied on the light probes resulted in static objects that did not 'look' specular but rather matte and textured. (See Supplementary Fig. 1).

In order to create a *matte-textured* appearance of objects in *dynamic* trials, the specularly-reflected light probe image was "stuck on" to the object (Doerschner et al., 2011), creating a structure from motion consistent with a diffusely- reflecting, textured object. Dynamic *specular* objects elicited characteristic specular flow motion patterns, making them appear shiny when moving (Doerschner et al., 2011).

Shape probe and sampling. We used a novel shape probe to assess local shape perception (Fig. 1D), introduced recently by Koenderink, van Doorn, and Wagemans (2014). The probe consisted of five categorical shape indices corresponding to generic surface patches: 'cap', 'ridge', 'saddle', 'rut', 'cup'. This probe has the advantage over the commonly used gauge figure (Koenderink & van Doorn, 1992) in that it measures local shape (as opposed to surface attitude), it is invariant under rotations and scale (just as the local shape itself), and it corresponds to the three prototypical surface shapes – elliptic (concave and convex), parabolic (concave and convex), and hyperbolic (Do Carmo, 1976).

[1] Note that while we used qualitatively similar shapes, the geometry of our 'furrow' deviates from the one defined by Koenderink and Van Doorn (1980).

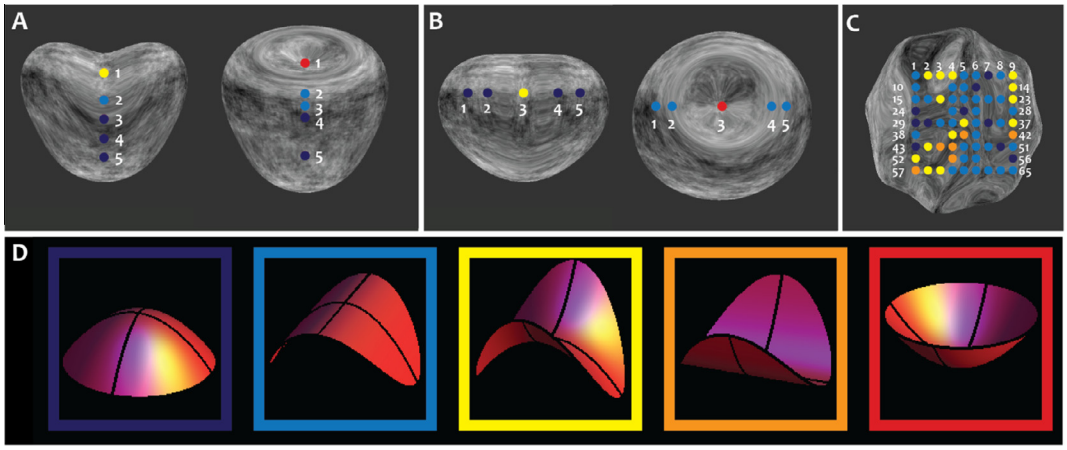

Fig. 1. *Stimuli.* (A) Parametric shapes used in Experiment 1. Both, the "furrow" (left) and the "dimple" (right) are shown from a generic viewpoint. All objects were rendered either with a specular material or a "stuck-on" specular (matte -appearing) texture. In static trials specular and "sticky" stimuli are very similar. Colored dots on the objects and corresponding numbers indicate the shape sample locations. Note that the sampling grid here is just an approximation as actual sample were directly *on* the surface. The color codes the surface curvature ground truth value (see D. below). (B) Depicted are the same shapes as in A, this time from a more accidental viewpoint, which should make shape estimation slightly more challenging. (C) The complex shape used in Experiment 2, and corresponding sample locations (color codes are as in (A) and (B)). (D) Our novel shape probe consisted of five shape categories to choose from (from left to right): "cap", "ridge", "saddle", "rut", "cup".

The surface patch on the object to be judged by the observer was indicated by small red dot that was attached to the object's surface. This dot was modeled as a small sphere with a diameter of.25 degrees visual angle that intersected the object's surface at its equator. In order to cue the area from which observers integrate information around a given sample dot, we briefly (1 s) presented a red circle (radius = 0.5 degrees visual angle, centered on the dot), before the small dot would appear (Fig. 2). In *Experiment 1* this circle was presented prior to every trial. In *Experiment 2* it was shown during instructions prior to the actual experimental trials.

In *Experiment 1*, local shape estimates were obtained from five sampling points lying on a vertical (Fig. 1A) or a horizontal line (Fig. 1B). In *Experiment 2* we used a sampling grid consisting of 65 points, that approximately covered the interior of the complex object (Fig. 1C).

2.3. Apparatus

Stimuli were rendered in real time (OpenGL) and displayed by using custom software written in Psychtoolbox v3.0.11 (Brainard, 1997; Pelli, 1997) on an iMac computer (Apple, Inc., LCD, 21.5 inches, refresh rate: 60 Hz, resolution 1920 × 1080) with a built-in NVIDIA GeForce GT 750 M graphics card. Open GL Shaders were created with the Gratin toolbox (Vergne, Ciaudo, & Barla, 2014). The observer's head position was fixed with a chin rest at 57 cm viewing distance and the experiment took place in a dark room. Responses were indicated via keyboard presses.

2.4. Task and procedure

On every trial of both experiments observers performed a single interval, five alternative forced choice (5AFC) task (see Fig. 2). They were instructed to "choose which shape probe looks most similar to the local surface patch that is marked by the red dot on the object". They were also instructed to make *local* judgments, i.e. to limit their shape judgment to the area indicated by the red circle.

Observers viewed stimuli binocularly. On every trial, a given object was presented at the center of the screen and the shape index was displayed at the bottom of the screen (e.g. Fig. 2). At the beginning of each trial the "selected" shape probe was initialized to "saddle" and observers moved the mouse to toggle a red square between the 5 different shape categories. They pressed the "space" bar to confirm their choice and to proceed to the next trial. The sample location for a given object was chosen randomly from the 5 (*Experiment 1*) or 65 possible locations (*Experiment 2*).

Trials were blocked into static and dynamic sessions. Surface reflectance (matte, specular) was randomized within each session. Observers completed both sessions on the same day. In *Experiment 1*, participants completed 120 trials (2 objects × 2 viewpoints × 5 locations × 2 materials × 3 repetitions) for the static, and 240 trials (2 objects × 2 viewpoints × 5 locations × 2 materials × 6 repetitions) for the dynamic session. For counterbalancing, six participants completed the static session first and six participants completed the dynamic session first. In *Experiment 2*, participants

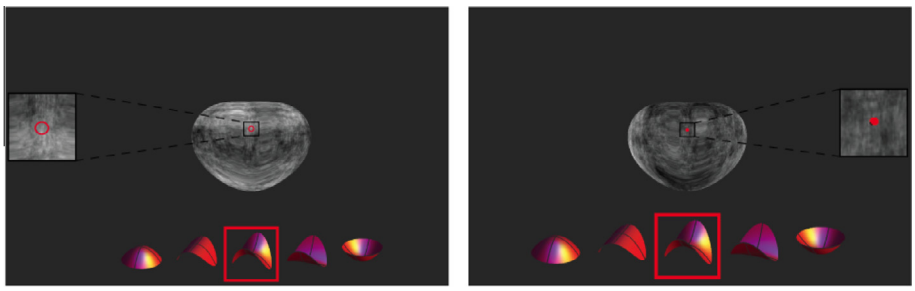

Fig. 2. *Shape probe and integration area.* In order to equate the area from which observers would integrate information for a given sample dot we displayed a red circle for 1 s that was centered on the sample location on a static view of the object (left panel + inset). Subsequently, during the static or dynamic stimulus display, this circle was replaced with the actual sample dot (right panel + inset).

performed 390 trials (65 locations × 2 materials × 3 repetitions) in the static condition and 650 trials (65 locations × 2 materials × 5 repetitions) in the dynamic condition.

As mentioned above, matte and specular objects were very similar in appearance in static trials (also see Supplementary Fig. S1). After confirming that there is no significant difference between these materials, we pooled those trials and analyzed static trials together across materials. However, we did use potential differences in shape judgment between static matte and static specular conditions as a baseline in the *Shape Change* analysis below.

Task Familiarization. In order to assess the effect of familiarity to 2nd order shape estimation, half of the observers practiced the shape index task for 10 trials with a familiar object (OpenGL teapot, with uniform albedo (0.5) and Lambertian reflectance) prior to the experimental trials.

2.5. Observers

Twelve naive observers from Bilkent University volunteered to participate in *Experiment 1*. Three observers from *Experiment 1*, three different naive observers and one author (DND) participated in *Experiment 2*. All observers had normal or corrected-to-normal vision. Experiments were approved by the Bilkent University Ethics Review Board in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki). Observers gave written informed consent prior to the experiment and they were paid 15 Turkish Liras (about 5 Euros) per hour. One observer did not meet the *repetition consistency* criterium (see below) for more than half of the sample points in any of the conditions in *Experiment 1* and was thus was excluded from the analysis.

2.6. Analysis

Response coding and measure. The shape index scale used in our experiments (Fig. 2) consisted of a subset of five indices that where taken from a vertical meridian of the continuous shape index space (Fig. 3, trajectory on the right). Thus, we think of our shape scale as a categorical (discrete) variable and rather than mean response of observers we use the *mode* in our analyses. Indeed, using the mean implicitly assumes a uniform metric on the shape index scale, an unfounded assumption at the perceptual level.

Repetition consistency. If an observer made a different judgment for a given sample location at every repetition, that location would be excluded from further analysis for having low *repetition consistency*. In *Experiment 1* observers had to make consistent judgments in at least 4 out of 6 trials in static (combined across matte and specular conditions), and dynamic conditions. In *Experiment 2*, observers had to have consistent judgments in at least 4 out of 6 trials in the static condition (combined across matte and specular), and in at least 3 out of 5 trials in the dynamic conditions (in order to keep consistency thresholds approximately proportionate to the number of repetitions in static and dynamic trials).

Recall that we used the pool of static trials combined across materials for measuring the *correlation with 3D model curvature*. However, we did separate the static trials by material (matte vs. specular) when analyzing *shape change*, as described below. Thus, to be included in the latter, at least 2 out of 3 responses had to be identical[2].

Correlation with 3D model curvature. This analysis served as an initial assessment of whether observers were able to perform the task and whether their judgments varied with our experimental manipulations of surface material and object motion. To this end, we calculated theoretical *curvature* values (Fig. 4) and the magnitude of curvature (*curvedness*, not shown here) for each 3D model at each sample location using the Graph Theory Toolbox (Peyre, 2008; Cohen-Steiner & Morvan, 2003). We correlated curvature values with observers' shape judgments across the entire object, and separately at object center and border locations. To investigate the effect of curvedness on observers' shape estimates, we sorted the data into high and low curvedness-points, using the median as a cutoff (this cutoff was 0.024 in *Experiment 1* and 0.014 in *Experiment 2*), and repeated the correlation analysis, and we measured the effect of curvedness on *repetition consistency*.

Note, that theoretical curvature values were continuous. To allow for the possibility of perfect correlations with observer settings we rounded the theoretical curvature values to the nearest integer. For readability we refer to theoretical curvature values of our models as *"ground truth"* in the remainder of the paper, but the reader is prompted to practice care in interpreting this as the ground truth of the corresponding real 3D objects that could have given rise to the visual stimuli. Especially in the static condition, where one can create any image by manipulating both the shape and the environment, the definition of ground truth is problematic[3] and no clear statement can be made regarding the ground truth shape of the object that produced the image. As a result, our experimentation, analyses, and conclusions focus primarily on exploring how the perceived local curvature varies with our the experimental conditions, namely the perceived material and the added motion of the object. Said differently, while no rigorous statement may be made about which of these conditions yields the most *veridical* judgments, more conclusive findings *can* be obtained about their effect on perceptual *consistency*. In this sense, *"ground truth"* is just an arbitrary reference point against which we compare observers' perception to infer differences between experimental conditions.

Shape estimation consistency and shape change. We used correlation analysis to obtain an estimate of observers' local shape judgment consistency across variations in motion and surface material. Since identical specular and matte-textured objects have been shown to elicit different perceptions of 3D-shape-related properties such as rigidity or the orientation of the rotations axis (Doerschner & Kersten, 2007; Doerschner et al., 2013), we expected that surface material types would cause differences in shape judgments in the dynamic conditions. Quantifying this difference is not trivial, since it is not clear how the Euclidean distance between any two shape categories would translate to a perceptual shape difference. Here we addressed this issue by developing an index that captures the *change in perceived local shape* between dynamic matte-textured and specular conditions independent of the shape categories. We first define a change index κ as follows:

$$\kappa = 1 - \frac{|M_{specular} - M_{matte}|}{4} \quad (1)$$

where M denotes the mode of the shape estimates, which could take on values of 1 ("cap"), 2 ("ridge"), 3 ("saddle"), 4 ("rut"), or 5 ("cup"). A $\kappa = 1$ would indicate no difference in perceived shape between matte-textured and corresponding specular condition, other values $\kappa < 1$, i.e. $0.75, 0.5, 0.25$ or 0, would indicate some difference in perceived shape. While κ in itself does not avoid the actual metric problem, our subsequent analysis does. We modeled the data as binomial variable:

$$B(p, n) = (p_{\kappa=1})^{n_{\kappa=1}}(1 - p_{\kappa=1})^{n_{\kappa<1}}, \quad (2)$$

[2] Static images of matte and specular objects were very similar. See sections on *Rendering* and *Task and Procedure*.

[3] Theoretically, this may be less of an issue in the dynamic condition, as the theory shows a direct relationship between curvature (or local shape) and the specular flow (e.g., Adato et al., 2010) and thus at least in terms of qualitative curvature values, as used in our experiment, the notion of ground truth might not be as abstruse.

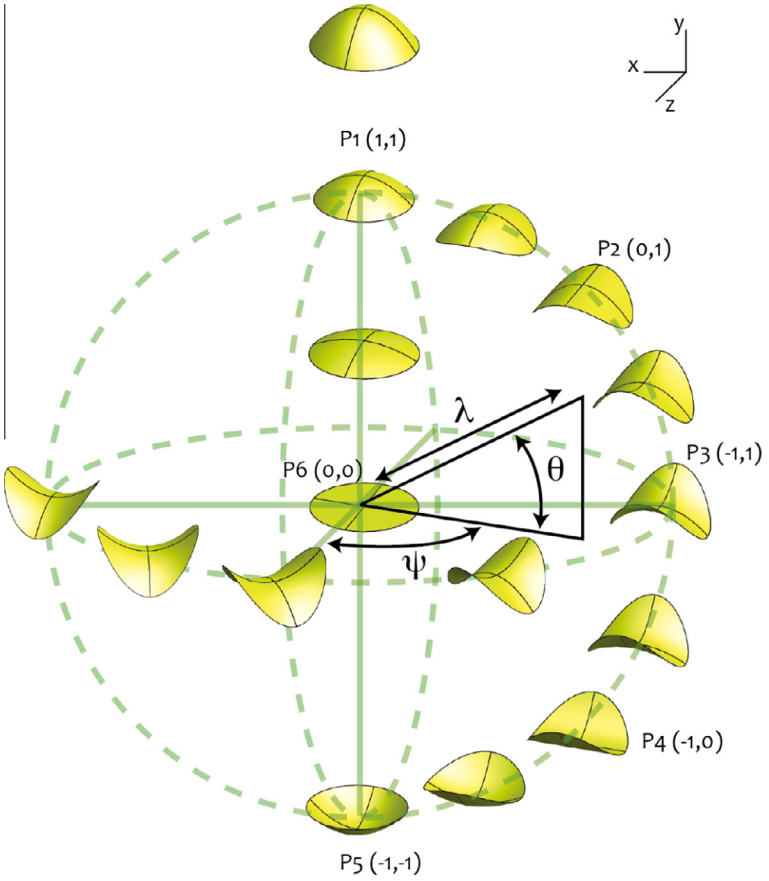

Fig. 3. *Illustration of the local shape space.* Second order curvature is represented by two principal curvature values for each shape, which are essentially derived from two perpendicular lines defining the local surface patch, indicated here by a positive or negative value. On the xy-plane, θ defines the shape index varying from convex elliptic to saddle to concave elliptic, as represented by 9 surface patches from P1 (1,1) to P5(-1,-1). Rotation along the y-axis (orientation) is defined by ψ, it varies between 0 and π as shown with the examples of different orientations of P3(-1,1) aligned across the half-diameter (dashed green line). The magnitude of the curvature (curvedness) varies between 0 (flat surface patch, P6) and infinity as shown by an example shape above P1, and it is defined by λ here (solid green axes). In our shape index probe, we represent the shape index space with 5 quadrics: P1: convex elliptical (cap), P2: convex parabolic (rut), P3: saddle (hyperbolic), P4: concave parabolic (ridge), and P5: concave elliptic (cup). The curvedness of the index was kept constant in this study.

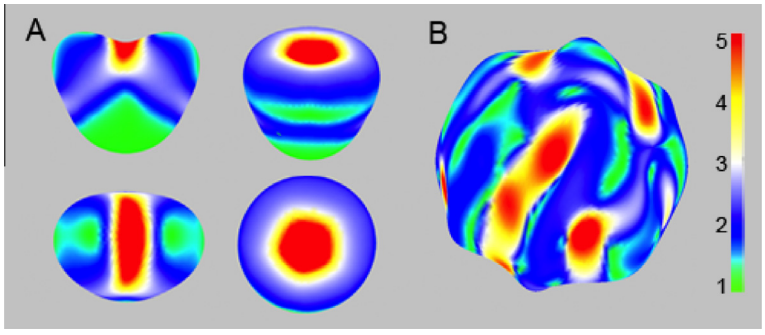

Fig. 4. *Ground truth surface curvature values.* (A) Theoretical curvatures for parametric shapes that were used in Experiment 1. (B) The complex shape of Experiment 2. Note that these values were continuous, unlike observers judgments, which were categorical. Numbers in the legend correspond to shape categories: 'cap'=1, 'ridge'=2, 'saddle'=3, 'rut'=4, 'cup'=5.

where, $p_{\kappa=1}$ is the probability of obtaining no shape-changes, $n_{\kappa=1}$ is the number of observed no shape changes, and $n_{\kappa<1}$ is the number of shape changes. We then estimated $p_{\kappa=1}$ for each, static and dynamic conditions. The question we tried to answer was, whether these two probabilities were significantly different.

In the static condition, any number of changes in perceived shape between the two surface material conditions should have occurred due to chance, since the visual information of corresponding stimuli in matte-textured and specular static conditions was very similar (Doerschner et al., 2011, see Supplementary Fig. 1). Therefore, to test whether the dynamic data can be explained by the static model (i.e. the static $p_{\kappa=1}$) we estimated the likelihood of the dynamic data given the static $p_{\kappa=1}$ (restricted model) and compared it to the likelihood of the dynamic data given the dynamic $p_{\kappa=1}$ (free model).

We then determined which model fits the data better using a parametric bootstrap analysis (Efron & Tibshirani, 1994; Efron & Tibshirani, 1994; Kingdom & Prins, 2010; Kingdom & Prins,

2010). A statistically significant outcome of this analysis implies that we reject the hypothesis that the proportion of perceived shape changes that occurred between matte-textured and specular dynamic trials was due to chance.

All analyses are conducted with custom scripts in Matlab (The MathWorks Inc, 2013) and SPSS (IBM Corp., 2010), and all significance values are reported two-tailed. Specific results for each analysis are presented in the next section.

3. Results

3.1. Overview

In both experiments we find that observers' estimate of local 2nd order curvature correlated positively and significantly with the *ground truth* values, and that correlations tended to be higher for points with larger curvedness values (see Figs. 5 and 6, for simple and complex shapes, respectively). Results further indicated that object motion tended to increase observers' *repetition consistency*. Interestingly, observers tended to be quite consistent with the *ground truth* for local ridges and cups but were rather variable in their estimates for other local shape categories. In both experiments, shape changes between the two types of surface material occurred significantly more often in the dynamic than in the static condition, suggesting that the perception of local shape is modulated by surface material-specific optic flow. Detailed results for each experiment are presented below.

3.2. Experiment 1 – simple shapes

Fig. 5 shows the distribution of frequencies of observers' modes for a given local curvature category across all sample points and objects. Overall, the distributions of modes concentrate around the respective *ground truth* value. For example, for the *ground truth* curvature value of *ridge* (first row, second column in Fig. 5), most observers in dynamic and also static conditions seem to judge the local surface to be a *ridge*, with only a few perceiving it as *cup*, or even as *saddle*. Some interesting biases occur for *caps*, which tend to be perceived also quite frequently as *ridges*, and *saddles*, which often were perceived as *ruts*, especially in dynamic conditions.

In addition to these global patterns, there seemed to be differences between static and dynamic conditions in local shape estimation (e.g. Fig. 5, first row, 3rd column, compare blue and green bars), as well as more subtle differences between the two dynamic conditions (compare green and light green bars in Fig. 5). High curvedness points (Fig. 5, second row) correlated better with *ground truth* values than did low curvedness points (Fig. 5, bottom row).

Overall, we found that the addition of object motion tended to make observers more reliable in 2nd order shape estimation. The rest of this section analyzes these informal observations quantitatively and in details.

3.2.1. Correlations with 3D model – dynamic

Correlations with *ground truth* in the dynamic conditions varied across objects (see Table 1) and materials, with the *furrow* in front

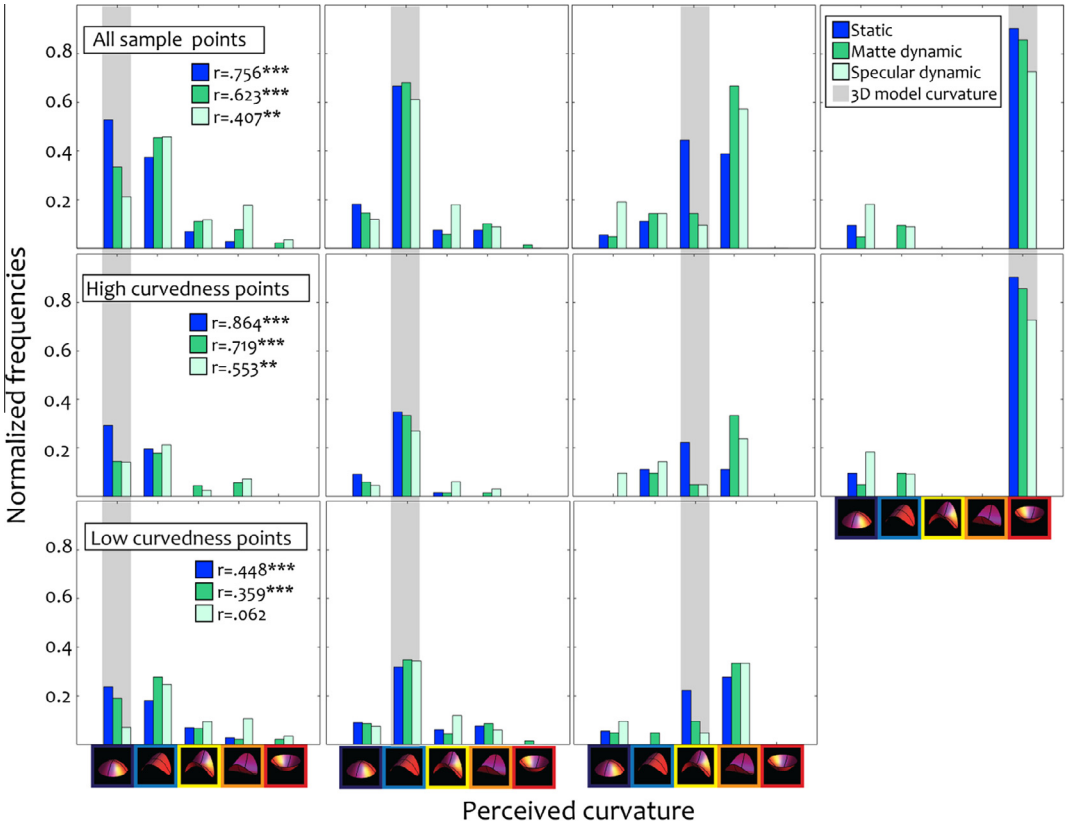

Fig. 5. *Local Shape Judgements in Experiment 1.* Shown are the normalized frequencies of that shape categories were selected (as measured by the mode) for a given *ground truth* category (light gray). Blue denotes frequencies for the static, green for the matte-dynamic and light green for the specular-dynamic conditions. Frequencies were normalized by the total number of responses with high *repetition consistency* in each of the three conditions (static, matte-dynamic, specular-dynamic). The simple objects in Experiment 1 featured only 4 of the 5 possible curvature types (no ruts). Plotted are the data of 11 observers for points with high and low curvedness values combined (upper row), only high curvedness values (middle row) and only low curvedness values (bottom row). There were no low curvedness *cups* among the sample points. As an inset to the plots in the left-most column we present the overall correlations with *ground truth* for each condition. See text for details.

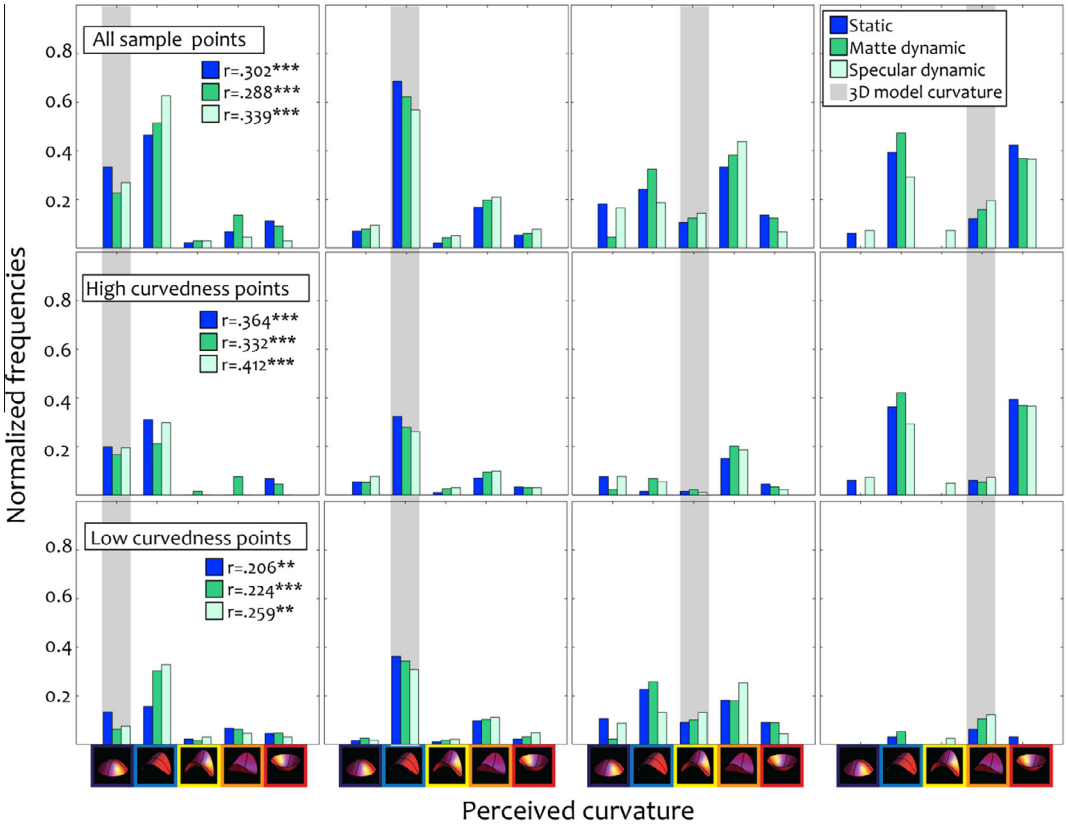

Fig. 6. *Local Shape Judgements in Experiment 2.* Shown are the normalized frequencies of that shape categories were selected (as measured by the mode) for a given *ground truth* category (light gray). Blue denotes frequencies for the static, green for the matte-dynamic and light green for the specular-dynamic conditions. Frequencies were normalized by the total number of responses with high *repetition consistency* in each of the three conditions (static, matte-dynamic, specular-dynamic). The complex object in Experiment 2 featured only 4 of the 5 possible curvature types (no *cups*). Plotted are the data of 7 observers for points with high and low curvedness values combined (upper row), only high curvedness values (middle row) and only low curvedness values (bottom row). As an inset to the plots in the left-most column we present the overall correlations with *ground truth* for each condition. See text for details.

Table 1
Correlations with 3D Model – dynamic conditions. Shown are correlations for each object across all observers.TV and FV stand for top view and front view, respectively. Blue indicates the overall correlation (across objects and observers) of shape judgments with *ground truth*. $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

Material	Objects				
	FurrowTV	*FurrowFV*	*DimpleTV*	*DimpleFV*	*AllObjects*
Matte	0.532^{***}	0.543^{***}	0.763^{***}	0.742^{***}	0.623^{***}
Specular	0.292^{*}	0.116	0.479^{***}	0.726^{***}	0.407^{***}
Static	0.757^{***}	0.517^{***}	0.862^{***}	0.860^{***}	0.756^{***}

view having lowest and *dimple* in top view tending to have the highest correlations with *ground truth*. Overall, these correlations were significantly lower for specular objects ($p < 0.01$, compare first and second rows in Table 1). Moreover, the differences between *ground truth* correlations of high and low curvedness points were significant for both, matte ($p < 0.001$) and specular ($p < 0.001$) conditions (see inset values in Fig. 5 of middle and bottom rows, green and light green symbols). The decrease in *ground truth* correlation of low curvedness points might be in part explained by the absence of the *cup* among the samples - a shape that people tended to be very robust and consistent in estimating (see high curvedness results). In particular, the correlation for dynamic specular objects was severely degraded in this latter analysis.

Repetition consistency was higher in the dynamic conditions, with only 19 out of 220 points for the matte and 25 out of 220 for the specular material not meeting the above described inclusion criterion. For comparison, 43 out of 220 points had to be excluded from the analysis in the static condition. There was no effect of curvedness on *repetition consistency*.

Boundary effects. Next, we determined whether there was difference in *ground truth* correlations between sample points located close to the object's occluding boundary and those closer to its center. To this end we assigned sample points 1 and 5 to the 'border' group, and the remaining ones to 'center', and computed the respective correlations for each object and across all observers.

Local shape judgments were closer to *ground truth* near the object border than those made near the center (Table 2). Differences in correlation coefficients (between center and border) were statistically significant for both, matte ($p < 0.05$) and specular ($p < 0.001$) conditions, with a more pronounced center-border difference for the latter.

3.2.2. Correlations with 3D model – static

Combining shape judgments across materials required us to establish that there was no systematic difference between matte and specular conditions in the static condition (paired sampled *t*-test on modes, $t(194) = -1.209$, $p = .228$). Next, we correlated the modes of individual observers' shape judgements with corresponding *ground truth*. Also in the static condition, correlations varied across objects (Table 1, bottom row). Interestingly, the static *ground truth* correlation was significantly higher than either one of the dynamic conditions ($p < 0.0001$ for specular, $p < 0.014$ for matte). Fig. 5 suggests that this was mostly driven by differences in local shape judgements of *saddles* (Fig. 5, column 1) and *cups* (Fig. 5, column 3) which were judged more frequently as such in

Table 2
Correlations with 3D Model – Boundary effects. Shown are correlations for each object across all observers for samples dots located at the object center and border. Blue indicates the overall correlation (across objects and observers) of shape judgments with *ground truth*. TV and FV stand for top view and front view, respectively, St: Static condition, Ma: Matte Condition, Sp:Specular Condition. $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

	Objects				
	FurrowTV	*FurrowFV*	*DimpleTV*	*DimpleFV*	*Allobjects*
Ma_Center	0.405*	0.313	0.779***	0.258	0.511***
Ma_Border	−0.205	0.675***	−0.354	0.860***	0.768***
Sp_Center	0.102	−0.035	0.437*	−0.040	0.192*
Sp_Border	0.224	0.265	0.301.	0.726***	0.658***
St_Center	0.783***	0.494*	0.866***	0.333	0.671***
St_Border	−0.168	0.771***	−0.407	0.848***	0.845***

the static than in the dynamic conditions. Differences between *ground truth* correlations of high and low curvedness points were significant ($p < 0.0001$, see inset values in Fig. 5 of middle and bottom rows, blue symbols).

Boundary effects. While detailed results are presented in Table 2, the bottom two rows show that as in the dynamic conditions, local shape judgments were significantly closer to the *ground truth* when performed near the object boundary compared to those made near the object center ($p < 0.007$).

3.2.3. Shape estimation consistency

Consistent with Fig. 5, observers were rather stable in their shape judgments across matte and specular-dynamic conditions ($r = 0.714, p < 0.0001$, compare green and light green bars in Fig. 5). Shape estimation consistency was also good across static and dynamic conditions (compare blue and green bars in the same figure). However, the consistency across static and dynamic matte $r = 0.908, p < 0.0001$ was significantly higher ($p < 0.0001$) than the shape consistency across static and dynamic specular conditions $r = 0.647, p < 0.0001$ (compare blue and light green bars).

Consistency across viewpoints. We measured the consistency of local shape perception for the *cup* on the dimple and the *saddle* on the furrow under two different view directions. The total number of consistent observer judgements was comparable for *saddle* and *cup* (Table 3) in static and dynamic matte conditions. However, the *cup* appeared to be judged less frequently consistent across viewpoint changes in the specular-dynamic condition, suggesting that (viewpoint dependent) specular flow altered the perceived shape of the *cup* substantially.

Overall, we find that observers' responses agreed more frequently with the *ground truth* curvature of the *cup*, than for the *saddle* (Table 4, bottom two rows). This was not surprising given that the furrow has been shown to yield ambiguous local shape estimates compared to clear-cut parabolic/hyperbolic boundaries as they occur on the dimple (Compare Figs. 6 and 7 in Nefs, Koenderink, and Kappers (2005)).

Shape change. We expected differences in local shape judgments between matte and specular objects in the dynamic condition. To test this hypothesis, we compared the number of shape-changes in dynamic and static conditions (Supplementary Fig. 6 shows κ for each condition, object, sample dot, and observer). In the static condition we observed 165 no-shape-change trials ($\kappa = 1$) and 42 shape-change trials ($\kappa < 1$), whereas in the dynamic condition we observed 142 no-shape-change trials ($\kappa = 1$) and 59 shape-change trials ($\kappa < 1$). Using parametric bootstrapping we rejected the hypothesis that the dynamic data could have been generated by the static model ($p < 0.0001$). Thus we can conclude that there was a significant change in perceived local shape between dynamic matte and specular conditions.

Effect of Task Familiarisation. We observed significant differences in the *ground truth* correlations between trained and untrained groups for all conditions: static ($p < 0.003$), matte dynamic ($p < 0.004$), and specular dynamic ($p < 0.017$). Table 5 shows that task familiarization resulted in increased *ground truth* correlations.

Table 3
Consistency across viewpoints. Cells show the number of observers, out of 11, who estimated local shape consistently across different viewpoints for *saddle* and *cup*. Note, that this does not mean that observers estimated the *ground truth* correctly – those values are shown in Table 4.

Local shape	Consistency across views		
	Static	Matte dynamic	Specular dynamic
Saddle	8	8	8
Cup	9	9	6

3.3. Experiment 2 – complex shapes

Fig. 6 shows the distribution of frequencies of observers' modes for a given local curvature category across all 65 sample points. Though the data seemed overall more noisy, we find again that the distributions of modes tend to concentrate around the respective *ground truth* (light gray bars in Fig. 6). Interestingly, despite the increased shape complexity, *ridges* were again judged most often as such while *cups* were frequently judged as *ridges*, especially in the dynamic conditions. Moreover, we find that *ruts* were perceived as *cups* or *ridges* – which, given the results of *Experiment 1* might reflect a general bias towards these shape types. As before, the addition of object motion tended to make observers more reliable in their local shape judgements.

3.3.1. Correlations with 3D model

Observers' shape estimates correlated significantly and positively with *ground truth*, though correlations were much lower than for the simple shapes (compare inset values in Figs. 5 and 6). Unlike *Experiment 1*, overall *ground truth* correlations did not differ significantly between any of the conditions, or between high- and low curvedness groups, though the former tended to be higher. Fig. 7 shows data from three representative observers. All individual data can be found in Supplementary Figs. 7 and 8 for dynamic and static trials, respectively.

Repetition consistency was higher in the dynamic conditions, with only 29 out of 455 points for the matte and 32 out of 455 for the specular material not meeting the above described inclusion criterion. For comparison, 126 out of 455 points had to be excluded from the analysis in the static condition. There was no effect of curvedness on *repetition consistency*.

Boundary effects. The *ground truth* correlations for sampling points at the border were higher for all conditions (see Table 6), however, the center-border difference was significant only for the dynamic conditions (specular: $p < 0.0001$ and matte $p < 0.05$).

3.3.2. Shape estimation consistency

As in *Experiment 1*, we first established that there was indeed no systematic difference between matte and specular conditions in the static condition (paired sampled *t*-test on modes, $t(402) = 1.263$, $p = .207$). Though *ground truth* correlations were low, observers were nevertheless rather consistent in their shape judgments across matte- and specular dynamic conditions ($r = 0.541, p < 0.0001$). Shape judgment consistency was also good between static and dynamic conditions for each material category, with $r = 0.722, p < 0.0001$ for matte, and $r = 0.723, p < 0.0001$ for specular conditions, respectively.

Shape change. In the static condition we observed 265 no-shape-change trials ($\kappa = 1$) and 138 shape-change trials ($\kappa < 1$), whereas

Table 4
Agreement with ground truth across viewpoints. Cells shows the number of observers' whose shape judgements where consistent with the *ground truth* for *saddle* and *cup*, respectively. The total number of observers was 11.

Local shape	Correlation with 3D model curvature					
	Static		Matte dynamic		Specular dynamic	
	Top view	*Front view*	Topview	Frontview	Topview	Frontview
Saddle	5	4	2	2	1	1
Cup	9	10	9	10	7	9

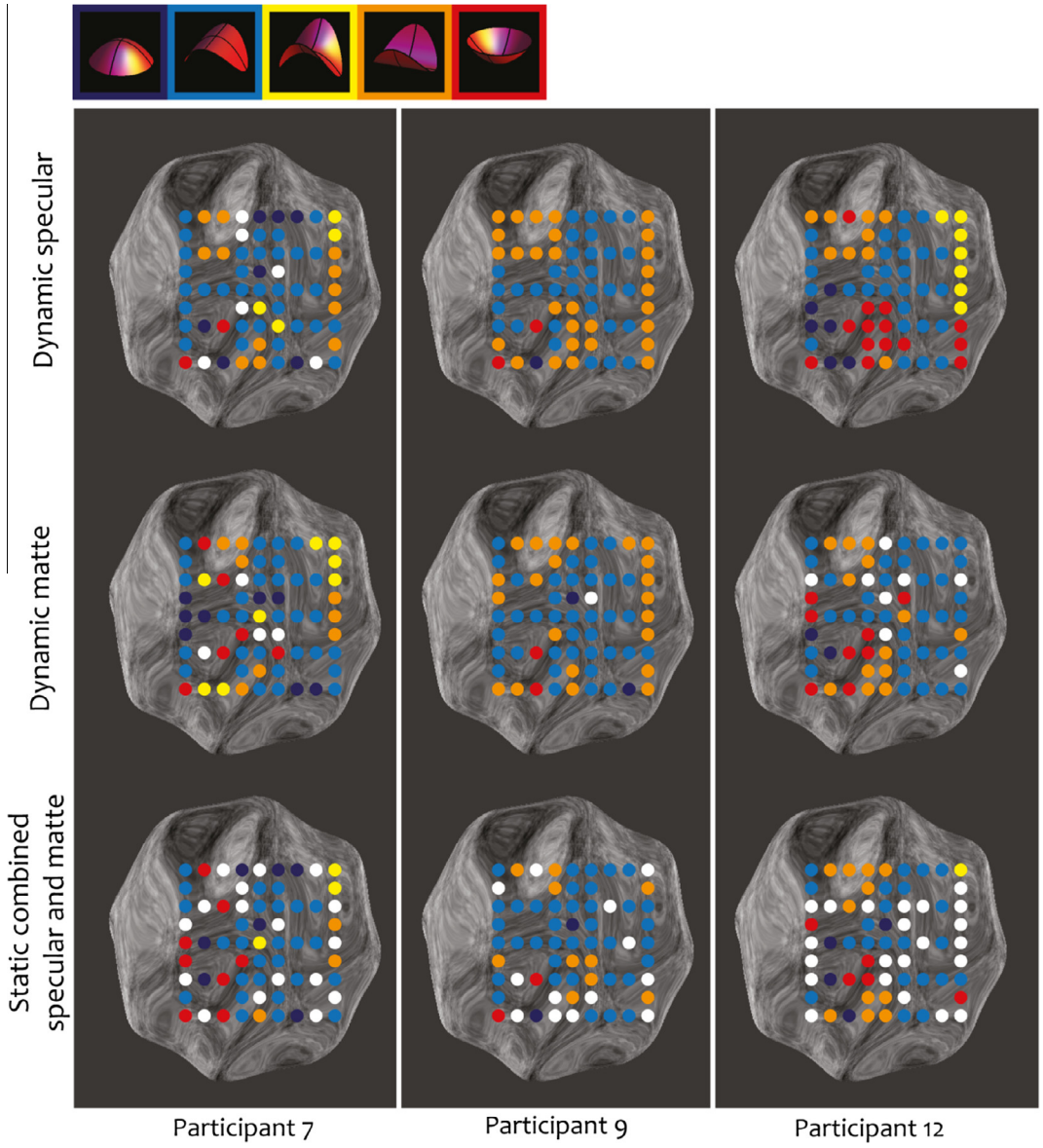

Fig. 7. *Local shape judgments and stimuli.* A more graphical presentation of the modes of observers' shape judgments presented in Fig. 6, superimposed onto the complex shape stimulus. Low repetition consistency locations are marked in white.

in the dynamic condition we observed 261 no-shape-change trials ($\kappa = 1$) and 146 shape-change trials ($\kappa < 1$). This time we could not reject the hypothesis that the dynamic data could have been generated by the static model ($p < 0.0001$). Thus, for the complex object we conclude that there was no significant change in perceived local shape between dynamic matte and specular conditions.

Effect of Task Familiarization. We found that familiarization increased the *ground truth* correlations for static- and matte

Table 5
Effect of Task Familiarization. Shown are correlations across all objects and observers. Observers who were trained on the task tended to have higher *ground truth* correlations. $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

	Condition		
	Static	Matte Dynamic	Specular Dynamic
No Training	0.633^{***}	0.468^{***}	0.246^{*}
Training	0.840^{***}	0.733^{***}	0.539^{***}
Combined	0.755^{***}	0.623^{***}	0.407^{***}

Table 6
Boundary effects for complex shapes. Shown are correlations across all observers for sample dots located at the object center and border. $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

Location	Consistency across views		
	Static	Matte dynamic	Specular dynamic
Center	0.228**	0.213***	0.193**
Border	0.399***	0.394***	0.539***

Table 7
Effect of task familiarization in Experiment 2. Shown are correlations across all observers. $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

	Condition		
	Static	Matte dynamic	Specular dynamic
No training	0.293**	0.155*	0.377***
Training	0.308***	0.389***	0.311***
Combined	0.302***	0.288***	0.339***

dynamic conditions, but not for the specular condition. The difference between trained group and untrained group was significant only for the matte dynamic condition ($p < 0.02$). See Table 7.

Results of both experiments are discussed below.

4. Discussion

4.1. Summary

The goal of this study was to investigate local shape perception of moving objects and how it is influenced by the type of optic flow induced by the surface material (i.e., optic flow vs. specular flow). Using a rotation and scale-invariant shape index probe and stimuli consisting of rotating matte-textured and specular objects we assessed both, how much local shape judgments correlated with curvature values of the 3D model and how this correlation would vary across experimental conditions. These included surface material (matte-textured, specular), object motion (static, rotating), viewpoint changes (only *Experiment 1*), and training. We employed simple-parametric and complex 3D objects, and also tested for each type how distance from the occluding boundary might affect observers' local shape judgments.

Overall, we found good agreement between the perceived local 2nd order shape of static and dynamic objects and the corresponding theoretical curvature values of the 3D-model, though there were differences between simple and complex objects. *Ground truth* correlations for trained observers tended to be higher. Observers were consistent in their shape judgments between static and dynamic conditions, while object motion, a key condition explored in our study, made observers more stable in their shape estimates. Differences between matte- and specular-dynamic conditions manifested itself in a significant proportion of changes in local shape judgments in *Experiment 1* but not *Experiment 2*. Since we had many manipulations in our study we show in Table 8 a brief overview of all the questions asked and the corresponding findings. Taken together, our results suggest that perceived local shape of specular objects is very sensitive to the characteristics of the specular flow and this effect might be modulated by shape complexity. We discuss these results next.

4.2. How is perceived local shape related to the stimulus-generating 3D Model

We defined *ground truth* to be the theoretical curvatures of the 3D model that we used to produce the renderings of our stimuli (2D images). Our results suggest that *ground truth* correlations vary substantially across shapes, surface materials, and even viewing conditions. *Ground truth* correlations for the complex shape in *Experiment 2* were substantially lower than for the simple object in *Experiment 1*. This was in part expected, since the constraints on local surface shape exerted by the object's occluding boundary fall off much more rapidly for complex shapes compared to simple shapes, making the task more challenging. However, a lower *ground truth* correlation does not imply inability to judge local shape or lack of a corresponding perceptual process for this task. What is important is, how *consistent* observers are in their judgment across multiple measurements. And indeed, observers were highly consistent in their shape judgments (repetition consistency 90–94%) – in both experiments. Somewhat surprisingly, though consistent with Savarese, Fei-Fei, and Perona (2004), repetition consistency tended to be lower (about 75–80%) in the static condition in both experiments.

Shape biases. We noticed, when analyzing observers' shape estimates as a function of *ground truth* curvature category, that in both experiments observers were very 'good' at judging *ridges* and *cups*. In fact we saw a tendency to choose these categories even when the *ground truth* was different: For example, *caps* were frequently perceived as *ridges* (*Experiment 1*), and *ruts* as *cups* in *Experiment 2*. *Saddles*, a particularly challenging shape to judge were frequently judged as *ruts* in *Experiment 1*, and as *ridges*, or *ruts Experiment 2*. This is in fact supported by what has been reported by Koenderink, van Doorn, and Wagemans (2014), where observers are uncertain to which shape category to choose for *saddles* almost half the time. The striking pattern here is that a given *ground truth* curvature tended to be most frequently confused with a neighboring category

Table 8
Summary of conditions and results. For simplicity, we refer to the matte and specular dynamic conditions as simply 'matte' and 'specular', respectively. Gray fields highlight observed trends in the data that did not reach statistical significance.

Question	Results	
	Simple shape	Complex shape
Repetition consistency?	Dynamic>static	dynamic>static
Shape biases?	Yes	Yes
	Ridge & *rut*	*Ridge*, *rut* & *cup*
Correlation w. 3D-model?	Higher	Lower
	Static>matte>specular	Specular>static>matte
Changes after training?	Yes	Matte only
Boundary effects?	Yes	Yes, specular only
Effect of curvedness?	High>low	High>low
Shape consistency w. static?	Matte>specular	No difference
Shape consistency		
Between materials?	Higher than complex shape	Lower than simple shape
Shape change?	Dynamic>static	No difference
Viewpoint consistency?	*Cup*: no, *saddle*: yes	N.A.

(e.g. *caps* with *ridges*) but not with those further away on the index e.g. with *saddles* or *ruts*. These patterns might indeed imply that there is a perceptual metric underlying the local shape space - with shapes that are physically closer in this space tending to be perceptually more similar. However, the nature of these patterns should be the subject of a more detailed investigation.

One should note that a puzzling exception to the above mentioned was a tendency of observers in *Experiment 2* to judge *ruts* not only as *cups* (see above) but also as *ridges*, though this might be explained by the general lower curvedness values of the stimulus in *Experiment 2*, i.e. which could have introduced more variability among observers.

Effects of curvedness. Higher curvedness points yielded higher *ground truth* correlations in both experiments, suggesting that such points are more informative for shape inference, both local and global. This is perhaps not surprising, at least in the case of dynamic conditions. Indeed, at the very least, theory indicates that high curvature points result in more stable, slower, and more easily measurable specular flow, while low curvature regions, in particular parabolic patches like *ridges* and *ruts*, entail flow singularities and increased measurement sensitivity (Adato et al., 2010; Adato, Zickler, & Ben-Shahar, 2011).

Our main motivation for studying *ground truth* correlations was to assess whether they would vary with proximity to the object boundary, between views, and most importantly, across surface reflectance type. We next discuss the former two factors within the context of the third, i.e. surface material.

Effects of surface material on 2nd order shape perception. While being as consistent across trials, correlations with ground truth shape were substantially lower in the specular rather than the matte dynamic condition. In an attempt to understand the source of this difference we separated judgments of local surface points near the object boundary from those closer to the object center (in a image projection sense). Intuitively, locations near the occluding boundary of the object should be "easier" to judge since the boundary significantly constrains the perceived shape. Indeed, for matte-textured objects, compared to points close to the boundary, query points far from the boundary had lower (but still significant) correlations with *ground truth*.

Conversely, judgments of local shape far from the occluding boundary were particularly *deviating* from *ground truth* for *specular* shapes, in both experiments. This was true even for the very prominent *cup* seen from a top view on the *dimple*: while the majority of observers responded *cup* in the static and also the dynamic matte condition (a singular exception to the boundary rule just described), this was not the case when the "dimple" was specular.

Interestingly, the *cup* was also judged by most of observers as having a different local shape under the two different views (top and front). This vulnerability to changes in viewpoint might be related to the special properties of shape from specular flow (SFSF): whereas conventional structure from motion (SfM) involves first order information processing (surface attitude), SFSF is based on second order information processing (curvature). Moreover, in a computational sense, the integral curves of the specular flows in the image plane are very sensitive to view point (Adato et al., 2010), and this might explain the altered appearance of local shape of the *cup* across our two viewing conditions.

Though not significant, the correlation with ground truth for complex objects was slightly higher for dynamic specular trials compared to their matte counterparts. This difference might again be attributed to the characteristics of specular flow since despite being dependent on viewpoint, is remarkably robust in signaling singularities (Adato et al., 2007, 2010). Since complex shapes such as those used in our *Experiment 2* posses many such singularities, the complex singular structure of their specular flow may provide potentially useful information for 3D shape.

The fact that we did not find significant differences between material conditions in *Experiment 2* might be due to the overall (expected) increase in task difficulty, which might have been exacerbated by overall lower values of curvedness of query points (see Methods). Moreover, there was substantial inter-observer variability, with some observers having opposite trends in their data (see Supplementary Table 2). This might have 'washed out' potential effects of surface material-specific optic flow.

4.3. Insights regarding specular shape

Dynamic. While the relationship between surface curvature and characteristics of specular flow were already described by Koenderink and Van Doorn (1980) and analyzed in depth in future studies (such as Adato et al., 2010), few have investigated how the dynamics of specular features affect the perceived local curvature, and the little evidence thus far is inconclusive. Hurlbert, Cumming, and Parker (1991) showed that the velocity of a specular feature alters the perceived local 3D curvature. We have shown previously that specular flow can affect the perceived rotation axis of objects (Doerschner et al., 2013), or can lead to a non-rigid appearance (Doerschner & Kersten, 2007; Doerschner et al., 2011)[4]. In contrast, specular highlight motion has been shown to aid *global* shape recognition (Norman, Todd, & Orban, 2004). The discrepancy between these studies might be explained the varying levels of shape complexity in the experiments: while effects in the former were revealed with simple shapes, the latter employed more complex object geometries. Moving complex specular shapes with characteristic high curvature regions will produce optic flow with 'sticky' regions, and thus turn 3D shape recovery *locally* into a classic (matte) SFM problem. While Norman, Todd, and Orban (2004) studied *global* shape recognition, results from the present study strongly suggest that *local shape* of rotating specular objects is perceived to be quite different from that matte-textured ones, especially for simple objects.

Even though we cannot yet explain how this difference manifests itself in the qualitative perception of specular shape, we did find that observers were highly consistent in their shape estimates across *static* and *matte dynamic* conditions. This implies that local shape in these two conditions was perceived to be quite similar. This was in contrast to what we observed for local shape estimates across static and *matte dynamic* conditions, suggesting these are more dissimilar, especially in *Experiment 1*. These differences in perceived local shape of specular objects might also affect overall object recognition. Although we did not include this in the present study, it would be very interesting to examine how local and global perception of shape interact, a study that is part of our forthcoming short term work.

Static. As in the dynamic case, there exists mixed and sometimes conflicting evidence that observers can successfully judge local shape of static specular objects (Fleming, Torralba, & Adelson, 2004; Savarese, Fei-Fei, & Perona, 2004). Our results for static stimuli suggest that it was indeed difficult for observers to judge local shape of specular objects although certain cues, such as orientation fields (Fleming, Torralba, & Adelson, 2004) and occluding boundary were available to the observer. Clearly, it remains possible that local shape judgments using the shape index probe are simply more difficult for observers compared to using a shape attitude probe, which might explain the discrepancy to the findings in Fleming, Torralba, and Adelson (2004). We are currently assessing this possibility more rigorously.

Is SFSF some kind of SFM? We do not know to what extend the neural computations underlying SFSF and SFM overlap. It is very possible that the brain does not "know" the physics of specular

[4] This was also reported by some of our observers in *Experiment 1*.

reflection and interprets specular flow directly. Such interpretation could lead to the perception of a virtual surface, which is different from that generated by conventional SFM, and thus could explain the observed shape change for the specular objects in our experiments, especially *Experiment 1*. One possibility to test whether similar information could be used for extracting 3D shape in matte and specular cases, could be to recover the 3D shape of the dynamic stimuli using a traditional SFM algorithm. In practice, however, this approach is problematic using existing tools. Indeed, in order to solve the correspondence problem, which is necessary for recovering 3D structure from images, typical SfM algorithms require the object to be moving rigidly and to reflect diffusely, and to possess sufficient spatial or appearance structure. Specular flow, however, violates several and sometimes all of these requirements, rendering SFM impossible unless the object is extremely simple (essentially, an ellipsoid) embedded in a sufficiently rich illumination environment. In all other cases, the parabolic lines on the surface give rise to specular flows with direction singularities, infinite speeds, and erratic behavior near the occluding contours (cf. Adato et al., 2010), a collection of properties that complicates the robust measurement of the flow and thus the recovery of shape (Adato, Zickler, & Ben-Shahar, 2011). Perhaps, as was suggested in the computational community, the perceptual system addresses these difficulties by trying to solve for the flow and the shape simultaneously (Adato & Ben-Shahar, 2011). If so, it is conceivable that two such optimization processes are executed concurrently by two neural circuits, one dedicated for specular material hypotheses and the other for matte ones. In such a system, the perceptual outcome of object material and its shape (and possibly its motion or rigidity also) could reflect the winning process, whose occasional failures (i.e., when the "wrong" process wins) explain the visual SFSF anomalies reported in the past (e.g., Doerschner & Kersten, 2007; Doerschner et al., 2011; Doerschner et al., 2013). Both of the possibilities just discussed are part of our short term exploration about the relationship between SfM and SFSF at the perceptual and computational modeling level.

4.4. Concerns about the probe

It might be possible that our probe provided an additional cue in the specular dynamic condition as it had its own geometry (sphere) and color (red). We needed a persistent marker to the surface position, especially in the specular reflectance condition, where the visual pattern could divert the observers' focus of attention from the query location. Thus, we first anchored the observer by briefly presenting a larger red circle around the area where a local shape judgment was to be made, and we subsequently displayed a small point attached to the surface to indicate the center of the area over which the observer to integrate the visual information. Observers were instructed to not use this point itself for their judgments. When the object was specular, there were two types of optic flow information present in the area two be judged: at the center the optic flow generated by the red dot, consistent with diffusely reflecting objects, and in its immediate and extended vicinity specular flow, thus generating a "cue conflict" situation.

While this conflicting information is a potential concern, there are probably two ways it could affect the observers' judgment: Potentially, the probe dot stuck on the surface could have affected the observers' judgment in two ways: First, because the probe was rendered to abandon any reflectance properties, it could have acted as a conflicting cue and pulled results in the specular condition more closely to those in the matte dynamic one. This would make a hypothetically large difference between these two conditions smaller (potential Type II error). Our results indicate significant differences between these two conditions, thus the potential effect might have been big enough.

Conversely, and maybe more worrisome, this possible conflict of cues could have made the curvature estimation task in the specular dynamic condition altogether different and could have introduced a systematic difference between the two conditions, when in fact the perceived local curvature for specular and matte conditions would have been identical if it was not for the dot probe (potential Type I error). We cannot rule out this possibility, though we believe that it was rather the obvious and measurable differences in local optic flow that gave rise to our results.

4.5. Task familiarization

Correlations with *ground truth* tended to be higher for observers who went through the task familiarization phase. In *Experiment 1* we observed the effect of training for all conditions while in *Experiment 2* it was restricted to matte dynamic condition. Keeping in mind that the number of training trials (10 trials) was small compared to the block of experiments (min 320 trials), they proved effective for the clarity of instructions and helped the observers' to categorize the estimated local surface shape. Training did not influence the overall pattern of results, i.e. the observed differences between specular and matte trials were not affected.

4.6. Inter-observer variability

One striking finding in *Experiment 2* is the very strong contrast between intra-observer consistency and substantial inter-observer variability. Large inter-observer variability in shape perception studies has been pointed out before (Wagemans, van Doorn, & Koenderink, 2010; Dövencioğlu et al., 2013) and has, in part, been explained by differences in observers' priors (e.g. light from above Mamassian & Goutcher, 2001; Sun & Perona, 1998; Ramachandran, 1988; or convexity Langer & Bulthoff, 2001)[5]. While this might explain some of the *general* phenomenon[6] (an alternative being that observers misunderstood the task), we believe that the reason for the exaggerated inter-observer variability in the *specular dynamic* condition lies in the nature of specular flow. For example, if the curvature is low, as is the case for some of our sampled locations, the object might in fact appear locally or globally non-rigid, thus making local shape estimation particularly challenging or ambiguous. In such cases some observers may settle for one and other for the alternative interpretation of the flow pattern. Interestingly, these perceptual challenges mirror the inherent computational difficulties in measuring specular flow reliably (Adato, Zickler, & Ben-Shahar, 2011), and those related to the estimation of shape from a single specular flow (Adato et al., 2010; Vasilyev et al., 2011).

5. Conclusion

We have shown that observers are able to make categorical local shape judgments for rotating simple and complex objects and proposed that a second order shape index probe is a useful tool for investigating local shape estimates in dynamic scenes. Moreover, we have demonstrated that perceived local shape of rotating specular objects differs substantially from that of matte-textured objects. We attributed this difference to the special challenges that shape from specular flow poses. Taken together our findings contribute to an understanding of human local shape perception and the role of surface reflectance in this process.

[5] It might also be the case that observer's priors are not stable between trials, which could lead to the low repetition consistency found in the static conditions. This needs further exploration.

[6] Though we did not observe a convexity prior in the data.

Acknowledgments

DD and KD were supported by an EU Marie Curie Initial Training Network "PRISM" (FP7 - PEOPLE-2012-ITN, Grant Agreement: 316746). DD was also supported by a TUBITAK BIDEB 2232 Post-doctoral Reintegration Fellowship (21514107-232.01-9150), and KD by a grant of the Scientific and Technological Research Council of Turkey (TUBITAK 1001 Grant No 112K069) and a Turkish Academy of Sciences (TUBA) GEBIP Award for Young Scientists. OBS work on this project was supported in part by the Israel Science Foundation (ISF individual Grant No. 259/12 and BIKURA Grant 1274/11), by the European Commission in the 7th Framework Programme (CROPS GA No. 246252), and by the Frankel Fund, the ABC Robotics initiative, and the Zlotowski Center for Neuroscience at Ben-Gurion University.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.visres.2015.01.008.

References

Adato, Y., & Ben-Shahar, O. (2011). Specular flow and shape in one shot. In: *British Machine Vision Conference (BMVC)* (pp. 1–11).

Adato, Y., Vasilyev, Y., Ben Shahar, O., & Zickler, T. (2007). Toward a theory of shape from specular flow. In: *International Conference on Computer Vision (ICCV)* (pp. 1–8).

Adato, Y., Zickler, T., & Ben-Shahar, O. (2011). A polar representation of motion and implications for optical flow. In: *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. IEEE* (pp. 1145–1152).

Adato, Y., Vasilyev, Y., Zickler, T., & Ben-Shahar, O. (2010). Shape from specular flow. *IEEE Transactions onPattern Analysis and Machine Intelligence, 32*(11), 2054–2070.

Bradley, D. C., Chang, G. C., & Andersen, R. A. (1998). Encoding of three-dimensional structure-from-motion by primate area mt neurons. *Nature, 392*(6677), 714–717.

Brainard, D. (1997). The psychophysics toolbox. *Spatial Vision, 10*(4), 433–436.

Cohen-Steiner, D., & Morvan, J.-M. (2003). Restricted delaunay triangulations and normal cycle. In *Proceedings of the nineteenth annual symposium on Computational geometry* (pp. 312–321). ACM.

Debevec, P. (2002). Image-based lighting. *IEEE Computer Graphics and Applications*, 26–34.

Do Carmo, M. P. (1976). *Differential geometry of curves and surfaces* (Vol. 2). Prentice-Hall Englewood Cliffs.

Doerschner, K., Kersten, D. (2007b). Perceived rigidity of rotating specular superellipsoids under natural and not-so-natural illuminations. *Journal of Vision* 7 (9), 838–838. URL <http://journalofvision.org/7/9/838/>

Doerschner, K., Fleming, R. W., Yilmaz, O., Schrater, P. R., Hartung, B., & Kersten, D. (2011). Visual motion and the perception of surface material. *Current Biology, 21*(23), 2010–2016.

Doerschner, K., & Kersten, D. (2007). Local 3-D curvature affects perceived rigidity and shininess of rotating specular ellipsoids. *Perception, 36*(Supplement), 106.

Doerschner, K., Kersten, D., & Schrater, P. R. (2011). Rapid classification of specular and diffuse reflection from image velocities. *Pattern Recognition, 44*(9), 1874–1884.

Doerschner, K., Yilmaz, O., Kucukoglu, G., & Fleming, R. W. (2013). Effects of surface reflectance and 3d shape on perceived rotation axis. *Journal of Vision, 13*(11), 8.

Dövencioğlu, D. N., Ban, H., Schofield, A. J., & Welchman, A. E. (2013). Perceptual integration for qualitatively different 3-d cues in the human brain. *Journal of Cognitive Neuroscience, 25*(9), 1527–1541.

Efron, B., & Tibshirani, R. J. (1994). *An introduction to the bootstrap* (Vol. 57). CRC Press.

Fleming, R. W., Dror, R. O., & Adelson, E. H. (2003). Real-world illumination and the perception of surface reflectance properties. *Journal of Vision, 3*(5), 347–368.

Fleming, R. W., Torralba, A., & Adelson, E. H. (2004). Specular reflections and the perception of shape. *Journal of Vision, 4*(9), 798–820.

Hartung, B., & Kersten, D. (2002). Distinguishing shiny from matte. *Journal of Vision, 2*(7). 551-551.

Hurlbert, A., Cumming, B., & Parker, A. (1991). Recognition and perceptual use of specular reflections. *Investigative Ophthalmology and Visual Science, 32*, 105.

IBM Corp. (2010). IBM spss statistics for mac, version 19.0, armonk, ny. <http://www-01.ibm.com/software/analytics/spss/>.

Kingdom, F., Prins, N. (2010). Psychophysics: a practical introduction.

Koenderink, J., van Doorn, A., Wagemans, J. (2014). Local shape of pictorial relief. *i-Perception* 5 (3), 188.

Koenderink, J., & Van Doorn, A. (1980). Photometric invariants related to solid shape. *Optica Acta, 27*(7), 981–996.

Koenderink, J. J., & van Doorn, A. J. (1992). Surface shape and curvature scales. *Image and Vision Computing, 10*(8), 557–564.

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

Landy, M. S., Dosher, B. A., Sperling, G., & Perkins, M. E. (1991). The kinetic depth effect and optic flowii. first-and second-order motion. *Vision Research, 31*(5), 859–876.

Langer, M. S., & Bulthoff, H. H. (2001). A prior for global convexity in local shape-from-shading. *Perception-London, 30*(4), 403–410.

Mamassian, P., & Goutcher, R. (2001). Prior knowledge on the illumination position. *Cognition, 81*(1), B1–B9.

Nefs, H. T., Koenderink, J. J., & Kappers, A. M. (2005). The influence of illumination direction on the pictorial reliefs of lambertian surfaces. *Perception-London, 34*(3), 275–288.

Nefs, H. T., Koenderink, J. J., & Kappers, A. M. (2006). Shape-from-shading for matte and glossy objects. *Acta Psychologica, 121*(3), 297–316.

Norman, J., Todd, J., & Orban, G. (2004). Perception of three-dimensional shape from specular highlights, deformations of shading, and other types of visual information. *Psychological Science, 15*(8), 565.

Pelli, D. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision, 10*(4), 437–442.

Peyre, G. (2008). Graph theory toolbox. <http://www.mathworks.com/matlabcentral/fileexchange/5355-toolbox-graph>

Ramachandran, V.S. (1988). Perception of shape from shading. *Nature*.

Sakano, Y., Ando, H. (2008). Effects of self-motion on gloss perception. *Perception* 37.

Savarese, S., Fei-Fei, L., Perona, P. (2004). What do reflections tell us about the shape of a mirror? In: *Proceedings of the 1st Symposium on Applied Perception in Graphics and Visualization*. ACM, pp. 115–118.

Sun, J., & Perona, P. (1998). Where is the sun? *Nature Neuroscience, 1*(3), 183–184.

The MathWorks Inc. (2013). Version 8.2.0.701 (r2013b), natick, Massachusetts. <http://www.mathworks.com/products/matlab/>.

Vasilyev, Y., Zickler, T., Gortler, S., & Ben-Shahar, O. (2011). Shape from specular flow: Is one flow enough? In: *Computer Vision and Pattern Recognition (CVPR)* (pp. 2561–2568).

Vergne, R., Ciaudo, G., & Barla, P., 2014. Gratin – a programmable node-based system for gpu-friendly applications. <http://gratin.gforge.inria.fr/>

Wagemans, J., Van Doorn, A. J., & Koenderink, J. J. (2010). The shading cue in context. *i-Perception, 1*(3), 159.

Wallach, H., & O'connell, D. (1953). The kinetic depth effect. *Journal of Experimental Psychology, 45*(4), 205.

Wijntjes, M. W. (2012). Probing pictorial relief: from experimental design to surface reconstruction. *Behavior Research Methods, 44*(1), 135–143.