What do reflections tell us about the shape of a mirror?

Silvio Savarese*   Li Fei-Fei   Pietro Perona

California Institute of Technology

Abstract

Three-dimensional shape may be perceived from static images. Contours, shading, texture gradients, perspective and occlusion are well-studied cues to this percept. When looking at a picture of a specular object, such as a silver vase, one additional cue is potentially available: a deformed picture of the reflected environment is seen in the surface of the object and the amount and type of deformation depend on its shape. Can specular reflections be used as a visual cue for shape perception? Our experiments show that our subjects are very poor at judging the shape of mirror surfaces in absence of other visual cues. However, for a considerable subset of the stimuli, subjects are highly consistent in their (most often wrong) perception. This observation leads us to the hypothesis that our subjects rather than 'computing' a percept from each image based on geometrical considerations, may be associating a shape to each pattern in a stereotypical way, akin to pattern-matching. This behavior is reasonable since, as suggested by our ideal observer analysis, the information available from specular reflections is ambiguous when the surrounding world is (partially) unknown.

Keywords: shape perception, specular reflections, mirror surfaces, shape from texture

1   Introduction

A sense of three dimensional shape may be perceived by looking at a 2D image of an object. This perception arises due to different cues such as contour, shading, perspective and texture. When viewing a picture of a specular object, such as a silver vase, one additional piece of information is represented by the reflection of the environment. A deformed picture of the environment is seen in the surface of the object and the amount and type of deformation depend on its shape. Can specularities be used as a visual cue for shape perception, especially in absence of other cues?

There have been extensive research in the past years to understand the geometrical relationship between shape and specular reflections and many computational models were proposed. Pioneering work was carried out by [Zisserman et al. 1989] and [Blake and Brelstaff 1988] who tackled the problem under the hypothesis of viewer motion. [Oren and Nayar 1997] analyzed 3D surface profiles travelled by virtual features. More recently, in [Savarese et al. 2004] we recovered the local shape of an unknown smooth specular surface through the reflection of a known scene patch. In our approach a surface reflects a scene patch onto the image plane of a calibrated camera. For instance, the scene can be a grid of intersecting lines. Thus, a mapping from the scene grid to the reflected grid in the image is defined. Such mapping not only changes the *orientation* of
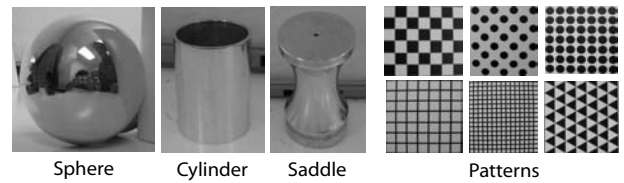
Figure 1: **Left:** the 3 shapes used in the experiment. **Right:** the 6 patterns used in the experiment
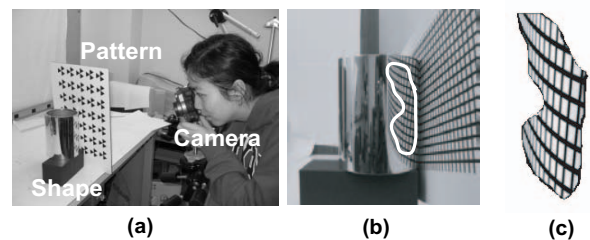

Figure 2: **(a)**. The setup: a camera takes a picture of a mirror shape reflecting a pattern. **(b)** Picture of the mirror shape reflecting the pattern: a patch of the reflection of the pattern is cropped from the picture. **(c)** The cropped patch does not include any visual cues besides a portion of the reflected pattern.

the grid lines but also *stretches* the grid step, modifying the local scale of the pattern. We proved that local shape of the surface can be recovered if and only if i) the scene is calibrated, ii) at least two orientations and local scale measurement (or equivalent information) are available at the same point in the image. Our results can be considered as an ideal observer theory for local shape reconstruction from specular reflection.

What happens when the surrounding world is unknown? The ideal observer analysis tells us that the problem is under-constrained and many solutions are possible. Thus, it may be useful to study the problem psychophysically and answer questions of the sort: how do humans cope with the problem of perception of the shape of specular surfaces given that most of time we do not have the exact geometrical information about the world? How useful are specular reflections when other visual cues are absents?

Shape perception from visual cues has been largely discussed in the vision literature. Models of visual processes, however, have tended to regard specular reflections as noise disturbing the underlying object geometry. The fact that highlights carry useful information has been often neglected. Exceptions are the studies by [Koenderink and Doorn 1980; Beck and Prazdny 1981; Todd et al. 1997; Lu et al. 1999; Nayar and Oren 1995]. In particular, [Blake and Bülthoff 1990] showed that the human visual system estimates shape and quality of a shiny object when the highlights are viewed stereoscopically. [Fleming et al. 2003b] explored whether statistics of real world illumination could be helpful in recognizing optical reflectance properties of shiny materials. Surprisingly, the reflection of a surrounding scene as a cue for shape perception has hardly been explored psychophysically (exceptions are recent works by [Fleming et al. 2003a; Savarese et al. 2003]). Through the present paper
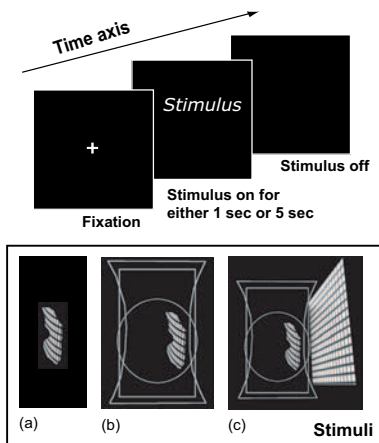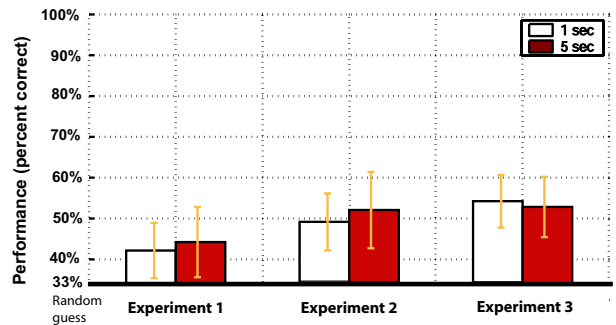
Figure 4: Performances summary of Experiment 1, 2 and 3. 33% means chance level performance. Note the average performances of all 10 subjects are only slightly better than chance.

Figure 3: **Top:** Each trial of the experiment is carried out as follows: A fixation cross is presented at the beginning of the trial for 240 msec. A stimulus (see **Bottom** panel) is then presented for either 1 sec or 5 sec. Subjects are instructed to respond the shape of the mirror as fast and as accurately as possible by pressing one of the 3 designated keys. **Bottom (a,b,c)** Examples of stimuli presented during experiment 1,2 and 3 respectively.

we would like to give some insight into this problem and address the following questions: i) to what extent the human visual system can use specular reflections in the absence of other visual cues; ii) what are the underlying computational strategies for this task?

The rest of the paper is organized as follows: in Sec. 2 we describe the experimental setup and method. In Sec. 3 we present our results. In Sec. 4 we analyze and discuss the results. We conclude the paper with final remarks in Sec. 5

2 Methods

Our goal is to study whether the human visual system uses specular reflection as a cue for shape perception. We asked 10 naive human subjects to discriminate images of mirror surfaces of qualitatively different shapes: a sphere, a cylinder and a saddle (i.e. with positive, zero and negative Gaussian, respectively. Fig. 1, left). Such shapes reflected the same scene with distinctly different distortions.

There were three consecutive experiments in our study. In the first experiment, the stimuli were 144 photographs (Fig. 2, (a)) of large patches of each mirror surface. Each patch contained the reflection of one of the six regular patterns (Fig. 1, right). The patterns were shown to the subjects in the familiarization phase before the actual experiment. Each patch was obtained by vignetting one of the photographs. We used irregularly shaped boundaries in order to eliminate occluding boundary information (Fig. 2, (b) and (c)). Each patch comprehended $25 - 35\%$ of the overall observable surface area. This guaranteed that the geometrical information of the surface shape provided by each stimulus was sufficiently rich. An example of stimulus is in Fig. 3, (a). Each stimulus was viewed monocularly and centrally in a standard computer monitor and subtends an average visual angle of 20°. Each trial was made up of 3 main steps as explained in Fig. 3, top. We presented each stimulus with two different presentation times, 1*sec* or 5*sec*.

In the second experiment, we used the same set of stimuli as in experiment 1, but with an additional piece of information: three overlapped silhouettes of the shapes were presented together with the cropped patch (Fig. 3, (b)). Thus, subjects were provided the approximate position of the patch within the shapes. In the third experiment, we used the same set of stimuli as in experiment 2.
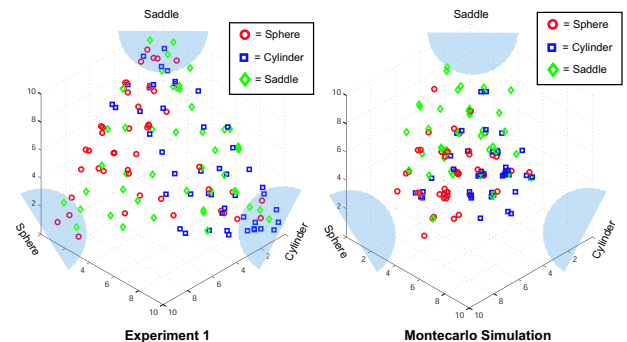
Figure 5: **Left.** Distribution of 144 stimuli for experiment 1 and presentation time 1sec. Each point corresponds to one of the 144 stimuli. Its position depends on the response patterns generated by 10 subjects. First, second and third coordinate of the point are the number of subjects who consider the stimulus as the specular reflection yielded by a sphere, cylinder and saddle respectively. For instance, a stimulus chosen by 10 subjects as the specular reflection yielded by a sphere, corresponds to a point with coordinates $(10,0,0)$. Regions around the 3 vertices $(10,0,0)$, $(0,10,0)$ and $(0,0,10)$ contain stimuli that exhibit high agreement among subjects. Those regions are in blue. The color code indicates the *actual* shape of the surface yielding that particular stimulus. **Right.** Monte-Carlo simulation to the 144 stimuli showing that if different subjects' responses were uncorrelated we would observe a completely different distribution.

Again an additional piece of information was given: the position of the pattern producing the reflection was shown (together with the cropped patch and three overlapped silhouettes, Fig. 3, (c)). Hence subjects were presented with the full geometrical structure, the only unknown being the actual shape which is yielding the patch.

We ran the three experiments in a consecutive order for each subject. This was done to rule out the possibility that subjects may have had memory and taken advantage of additional information provided by experiments 2 and 3. During each experiment, all 288 stimuli (144 images $\times 2$ different presentation times) were presented in random order, counter-balanced across all 10 subjects. A familiarization phase preceded the experiments, during which subjects were shown the 6 patterns. In addition, subjects were given a short period to practice keyboard responses for the three different mirror shapes (in the absence of specular reflection).

3 Results

The results of the 3 experiments are summarized in Fig. 4. Performances (in percent correct) are shown for experiments 1, 2 and 3. This result indicates that i) our subjects are only slightly better than

chance in discriminating the three mirror shapes; ii) performances corresponding to presentation time of 1 second and 5 seconds are not statistically different (t-test: p value > 0.05), suggesting the processing of such task is fast; iii) performances slightly improve from experiment 1 to experiment 2 (t-test: p value < 0.05). Thus, the knowledge of the location of the patch within the shape silhouettes seems to be helpful. We found that there is no substantial improvement of performances throughout experiment 1. This observation rules out the possibility that the improvement from experiment 1 to 2 is due to practice. Performances are, however, not significantly different between experiment 2 and experiment 3, suggesting that a richer geometrical structure has little effect in the way shape is recognized from reflection. This supplemental information also contributes to rule out the possibility that subjects might treat irregularly shaped outlines as occluding boundaries (thus, as a intrinsic properties of the object). It might be interesting to also repeat these experiments by showing subjects the stimuli through an irregular shaped hole, as suggested by one of the reviewers.

While subjects are almost at chance in discriminating shape differences from reflection, we find that responses are highly correlated. In other words, subjects tend to make decisions in a very consistent manner. We may visualize this observation through Fig. 5 (details are in the caption of Fig. 5). Let us call *high consistency regions* the volumes of radius $R = 10/3 = 3.3$, centered at each of the vertices $(10, 0, 0)$, $(0, 10, 0)$ and $(0, 0, 10)$ of the diagram. By inspection of Fig. 5, it is clear that stimuli tend to cluster toward the three vertices, revealing that there is high agreement amongst subjects' answers. We can contrast this with an uncorrelated Monte Carlo simulation of the responses. We take the observed average performance of experiment 1 (44%), and generate a simulated response pattern similar to the one in Fig. 5 Left panel. On average, only 0.05% of the 144 stimuli would fall into the 3 *high consistency regions*, indicated by the light blue semi-spheres (Fig. 5, Right panel). On the contrary, 29% of our actual stimuli are contained in the high consistency regions, revealing strong agreement in subjects' answers, both correctly and incorrectly. We obtain similar results for experiment 2 and 3. Fig. 6 shows examples of stimuli contained in the 3 consistency regions as well as in the regions in between.

4 Analysis

The conclusions drawn in Sec. 3 lead us to consider the nature of the mechanism that may underly this perception. If we consider the stimuli which give consistent shape perception (Fig. 6, first three rows) we see patterns of deformations that are similar within each category and different across categories. Either one or two patterns seem to account for the perception of each shape (see Fig. 7). Stimuli for which there is no consistent perception (Fig. 6, last row) may be associated to combinations of the 'pure' patterns of Fig. 7. One possible interpretation of this finding is that our subjects, rather than 'computing' a percept from each image based on geometrical considerations, may be associating a shape to each pattern in a stereotypical way, akin to pattern-matching. This behavior is reasonable since, as suggested by our analysis [Savarese et al. 2004], the information that is available from our stimuli is ambiguous.

If shape perception is the result of pattern-matching, it would be reasonable to believe that the 'fundamental' patterns of specular surfaces have been learned from experiences, and correspond to highly likely scene-surface configurations. A mirror cylinder standing on a horizontal textured plane would be one such example. However, we find that many of the stimuli for which there is a consistent and wrong perception correspond to such situations. Therefore we consider this an unlikely explanation. Another possibility is that our subjects are analyzing our stimuli as if they were images of curved textured surfaces rather than of curved mirrors. Notice that

in Fig. 3, (c) and Fig. 6 our stimuli often do not appear to be mirror-like. A qualitative analysis of the patterns in Fig. 7 is consistent with this interpretation, if one assumes a textured surface where the principal lines of the texture are roughly aligned with the principal curvatures of the surface. This leads to the intriguing hypothesis that shape perception from specularities and texture deformations [Zaidi and Li 2002; Todd and Akerstrom 1987; Knill 1998; Kim et al. 2004; Malik and Rosenholtz 1997; Blake et al. 1993] seem to share a common basis.

If the texture based pattern-matching is a strategy that the visual system uses to perceive shape from reflection, what would this tell us about the underlying computation pathways that is used to perform such tasks? Moreover, can we infer from this that shape perception from texture and from mirrors share some common computation or strategies? How is such sharing of resources mapped out? Where do they merge and where do they branch off? All these questions are highly important and crucial to further our understandings of the visual system. In addition, it might shed more light on how to engineer more powerful and efficient models for shape perception.

A natural question is whether such performance would be different if one uses reflections of real world scenes rather than regular synthetic patterns. The reason is twofold. First, our patterns were simple geometrical forms. This could have led subjects to misinterpret our stimuli as textures rather than actual specular reflections. Second, the human visual system may exploit prior knowledge on regularities of real world structure (i.e. statistics of real world images). In fact, by identifying in specular reflections some degree of deviation from such statistics, the human visual system may become capable to successfully interpret the shape of specular surfaces. We would like to address such questions in the future.

We conclude this section by comparing our results with those in [Fleming et al. 2003a]. They found that the shape of mirror surfaces is readily perceived, even when the only cue is specular reflections. They argue that specular reflections exhibit a different pattern of compression than surface texture. This feature would allow the human visual system to discriminate between these two cases. Hence the pattern of compression would represent a cue for shape from specularities. We interpret this apparent opposite results by pointing out that a number of different cues might be integrated by the human visual system in their experiment. For instance, specular reflections show different behavior depending on whether they appear on smoothly curved regions or in proximity of occluding contours and in regions of high curvature. In each case specular reflection would provide the visual system with different information about the geometry of the surface. This is particularly true when the specular object has complex topology. Therefore these indirect shape cues might also have contributed in the perception of shape in their experiments. Clearly much additional work is needed to reconcile these findings and understand the contributions of these different factors, and in turn the underlying computational strategies.

5 Conclusion

We have studied how humans perceive shape from mirror reflections. We use six regular synthetic patterns and their mirror reflections in three different shapes: sphere, cylinder and saddle. Our results indicate that mirror reflection is only a very weak cue to shape perception in the absence of other visual cues. Closer look at the response pattern of our ten subjects reveals that there might be a set of underlying *templates* that subjects use for shape perception from reflections. These templates follow closely the deformation patterns of texture on shapes, rather than reflection in shapes. Future work may include further investigation of the scope of this hypothesis using natural scene reflections.

9/10 sp　8/10 sp　8/10 sp　8/10 sa　7/10 sa　7/10 sa　7/10 sp

Best agreement for sphere

10/10 sa　10/10 cy　10/10 cy　9/10 sa　9/10 cy　9/10 cy　8/10 cy　8/10 sa　7/10 sp　7/10 sp

Best agreement for cylinder

10/10 sa　9/10 s　9/10 sa　9/10 sa　8/10 s　8/10 cy　8/10 cy

Best agreement for saddle

Sphere - Cylinder　Cylinder - Saddle　Sphere - Saddle　Sphere - Cylinder - Saddle

(6,4,**0**) (6,4,**0**) (4,5,**1**)　(0,5,**5**) (0,5,**5**) (0,6,**4**)　(**6**,0,4) (**5**,1,4) (**4**,0,6)　(3,4,**3**) (4,3,**3**) (2,3,**5**)
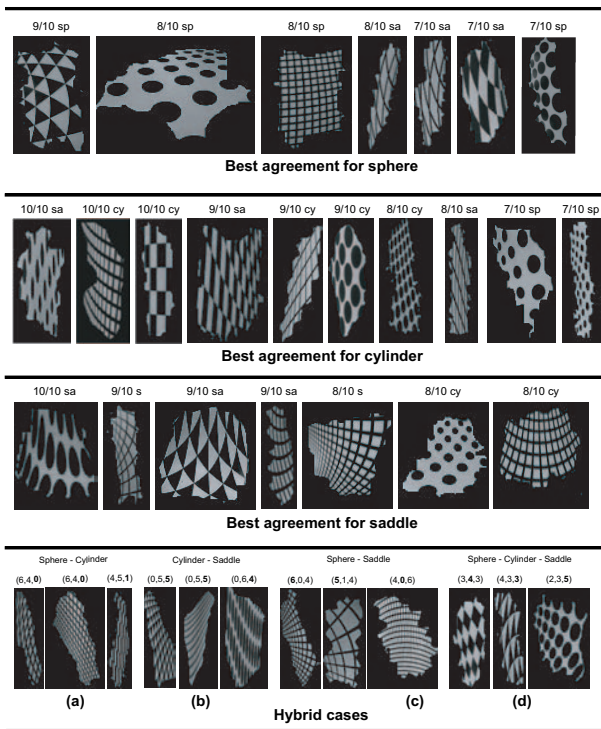
(a)　　(b)　　(c)　　(d)

Hybrid cases

Figure 6: **First panel (from the top):** Examples of stimuli most frequently perceived as from the sphere shape (i.e. examples from the *high consistency region* of sphere in Fig. 5). The patches are sorted according to the score. The text above each stimulus (e.g 8/10 sa) indicates the score fraction (e.g. 8 subjects out of 10 agreed responded as a sphere) followed by the label of the actual shape of the surface attached to that patch (e.g. saddle). **Second, Third:** Examples of stimuli most frequently perceived as from the cylinder and saddle respectively. **Fourth-(a):** Example of stimuli contained in the region between sphere and cylinder (see also Fig. 5). These stimuli are consistently classified as either sphere or cylinder. Each stimulus is labelled by a vector as described in the caption of Fig. 5. E.g. $(6,4,\mathbf{0})$ means that this stimulus is chosen as from sphere by 6 subjects, cylinder by 4, and saddle by none. Bold font of 0 indicates the actual shape of which the stimulus is reflected, in this case a saddle. The order of shape for the responses is (sphere, cylinder, saddle). **Fourth-(b,c,d):** Examples of stimuli (see also Fig. 5) contained in the region between cylinder and saddle, sphere and saddle, and central region respectively, the latter having obtained the lowest agreement among subjects.

References

BECK, J., AND PRAZDNY, S. 1981. Highlights and the perception of glossiness. In *Perception and Psychoph.*, vol. 30, 407–410.

BLAKE, A., AND BRELSTAFF, G. 1988. Geometry from specularities. In *Proc. of Int Conf. of Computer Vision*, 394–403.

BLAKE, A., AND BÜLTHOFF, H. 1990. Does the brain know the physics of specular reflection? In *Nature*, vol. 343, 165–168.

BLAKE, A., BÜLTHOFF, H., AND SHEINBERG, D. 1993. Shape from texture: ideal observers and human psychophysics. In *Vision Research*, vol. 33, 1723–37.

FLEMING, R., TORRALBA, A., AND ADELSON, E. H. 2003. How image statistics drive shape-from-texture and shape-from-specularities. In *Proc. of Third Annual Meeting of the VSS - http://journalofvision.org*.

FLEMING, R. W., DROR, R. O., AND ADELSON, E. H. 2003. How do humans determine reflectance properties under unknown illumination? In *Journal of Vision*, vol. 3, 347–368.

KIM, S., HAGH-SHENAS, H., AND INTERRANTE, V. 2004. Conveying shape with texture: experimental investigations of texture's effects on shape categorization judgments. In *IEEE Transactions on Visualization and Computer Graphics*.
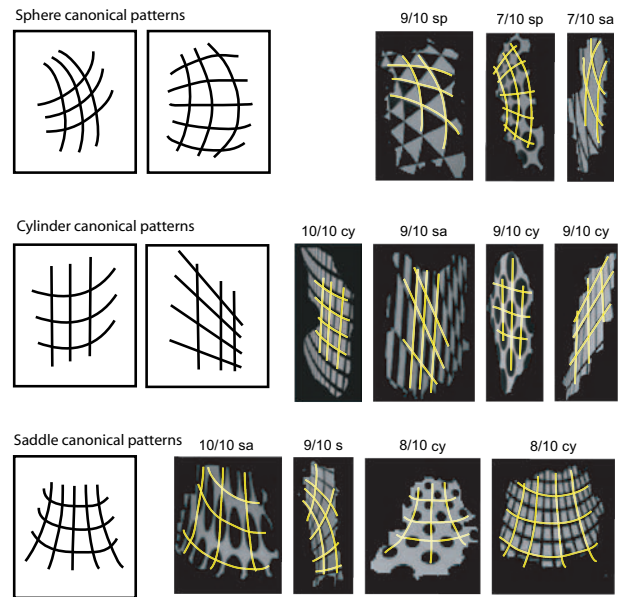
Sphere canonical patterns　　　9/10 sp　7/10 sp　7/10 sa

Cylinder canonical patterns　　10/10 cy　9/10 sa　9/10 cy　9/10 cy

Saddle canonical patterns　　10/10 sa　9/10 s　8/10 cy　8/10 cy

Figure 7: **Canonical patterns. Top:** Stimuli that are most frequently perceived as from sphere shape reflection seem to exhibit features that can be connected in either of the two template patterns on the left side of the panel. **Middle and Bottom:** Similar conclusion can be drawn for stimuli that are most frequently perceived as from cylinder and saddle shapes, respectively.

KNILL, D. 1998. Surface orientation from texture: Ideal observers, generic observers and the information content of texture cues. In *Vision Research*, vol. 38, 1655–1682.

KOENDERINK, J., AND DOORN, A. V. 1980. Photometric invariants related to solid shape. In *Acta Optica*, vol. 27, 981–996.

LU, R., KOENDERINK, J., AND KAPPERS, A. M. 1999. Surface roughness from highlight structure. In *Applied Optics*, vol. 38, 2886–2894.

MALIK, J., AND ROSENHOLTZ, R. 1997. Surface orientation from texture: Isotropy or homogeneity (or both?). In *Vision Research*, vol. 37, 2283–2293.

NAYAR, S., AND OREN, M. 1995. Visual appearance of matte surfaces. In *Science*, vol. 267, 1153–1156.

OREN, M., AND NAYAR, S. 1997. A theory of specular surface geometry. In *Trans. Int. Journal of Computer Vision*, 105–124.

SAVARESE, S., FEI-FEI, L., AND PERONA, P. 2003. Can we see the shape of a mirror? In *Proc. of 3rd Annual Meeting of VSS - http://journalofvision.org*.

SAVARESE, S., CHEN, M., AND PERONA, P. 2004. Recovering local shape of a mirror surface from reflection of a regular grid. In *Proc. of European Conference of Computer Vision*.

TODD, J. T., AND AKERSTROM, R. A. 1987. Perception of three-dimensional form from patterns of optical texture. In *Perception and Psychophysics*, vol. 13, 242–255.

TODD, J. T., NORMAN, J. F., AND KOENDERINK, J. 1997. Effects of texture, illumination, and surface reflectance on stereoscopic shape perception. In *Perception*, vol. 26, 807–822.

ZAIDI, Q., AND LI, A. 2002. Limitations on shape information provided by texture cues. In *Vision Research*, vol. 47, 815–835.

ZISSERMAN, A., GIBLIN, P., AND BLAKE, A. 1989. The information available to a moving observer from specularities. In *Image and Video Computing*, vol. 7, 38–42.