

# Dissertation Proposal

Gizem Küçükoğlu

New York University

Advisor: Michael S. Landy

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Background</b>	<b>2</b>
<b>3</b>	<b>Project Descriptions</b>	<b>3</b>
3.1	Is gloss perception context dependent? . . . . .	3
3.1.1	General Methods . . . . .	4
3.1.2	Apparatus . . . . .	6
3.1.3	Data Analysis . . . . .	6
3.1.4	Experiment 1 . . . . .	6
3.1.5	Experiment 2 . . . . .	9
3.1.6	Experiment 3 . . . . .	11
3.1.7	General Discussion . . . . .	12
3.1.8	Control Experiment . . . . .	13
3.1.9	Future Directions . . . . .	14
3.2	Are gloss and shape jointly estimated? . . . . .	15
3.2.1	Physical Stimuli . . . . .	15
3.2.2	BRDF Fitting . . . . .	16
3.2.3	Rendered Stimuli . . . . .	18
3.2.4	Experimental Design . . . . .	18
3.2.5	Project Goals . . . . .	18
3.3	Do additional cues influence material perception? . . . . .	19
3.3.1	Stimuli . . . . .	20
3.3.2	Experimental Design . . . . .	20
3.3.3	Project Goals . . . . .	21
<b>4</b>	<b>Bibliography</b>	<b>23</b>

# 1 Introduction

Material perception research is the study of how we perceive what things are made out of. Even though we can extract information about materials using all five senses, the focus of this work is the problem of extracting information about materials from visual information. Surface gloss, a specific material characteristic, allows us to perceptually distinguish many materials, yet it is not completely understood what cues contribute to estimating surface gloss. The problem of recovering gloss from visual information is under-constrained and complex because the highlight field of a glossy object depends on an object’s surface material, its 3D shape and the illumination.

When an object reflects the world around itself, the features of this reflected environment create a highlight field on the object’s surface. The goal of this work is to understand whether the visual system estimates gloss only using the image of the highlight field independent of the 3D shape and the illumination, or together with shape and the illumination, attempting to stay “gloss constant” when the 3D structure and the illumination changes.

# 2 Literature Survey

A good overview of the past and recent developments in research on gloss perception is provided by Chadwick and Kentridge (2015) and a broader overview of material perception research is provided by Fleming (2014) and Anderson (2011).

Two broad classes of theory have been proposed to explain how the highlights on a glossy surface are interpreted. The first approach suggests that simple image statistics are used to estimate gloss independently from shape and illumination (Motoyoshi, Nishida, Sharan, & Adelson, 2007) or by using specific properties of the surface highlights such as the highlight’s size, contrast and sharpness (Fleming, 2012; Marlow & Anderson, 2013; Marlow, Kim, & Anderson, 2012). This approach of using “image cues” to estimate gloss reduces the problem down to image statistics to explain how shape (Ho, Landy, & Maloney, 2008; Olkkonen & Brainard, 2010; 2011) and illumination (Fleming, Dror, & Adelson, 2003; Olkkonen & Brainard, 2010; 2011) affect gloss. However, this approach failed to predict perceived gloss from image cues such as luminance skew, particularly when the highlights on a glossy surface are inconsistent with the 3D structure of the surface (Anderson & Kim, 2009; Marlow, Kim, & Anderson, 2011).

The second approach is inverse optics, which suggests that the visual system models the physical properties of the scene such as the light sources, 3D structure of shapes and

the reflectance parameters of the surfaces by using the visual image. The advantage of this approach is that it results in a detailed model of the scene and it can also provide gloss constancy when illumination and shape changes. However, this is a complex and an underconstrained problem to solve. To estimate and discount the illumination, the visual system needs to estimate and discount the reflectance. This is impossible to solve as different pairs of illumination and surface reflectance can create the same image. For example, a chrome sphere in a “blurry” environment might look the same as a pearly sphere in a “crisp” environment. This is why many studies in inverse optics make prior assumptions about the light sources (e.g. light coming from above), the environment and the surface reflectance properties.

Our work’s goal is to understand which cues are used to estimate gloss and how our visual system uses these cues to solve for gloss in a scene. Some of this work is based on the findings of Ho, Landy and Maloney (2008) on the interaction between gloss, shape and illumination and Fleming et al. (2003) on the interaction between illumination conditions and perceived gloss.

### 3 Project Descriptions

#### 3.1 Is gloss perception context dependent?



Figure 1: *Two copper spheres reflecting a cloudy environment (left) and a high contrast forest environment (right).*

Objects reflect the environment that they are in (Figure 1). This means that an image presented by a material not only depends on the reflectance properties of the surface, but also on the illumination condition. Yet, in both of the spheres shown in Figure 1 something about the sphere’s appearance stays constant: the material still looks like copper. If we are gloss constant, the visual system can estimate the glossiness of a material reliably while compensating for changes in illumination. This project aims to

understand what information we use to compensate for changes in illumination when estimating surface gloss.

Previous work by Fleming, Dror and Adelson (2003) tested this by rendering spheres with identical gloss parameters under simple and complex (natural) illumination conditions. They present these objects on arbitrary checkerboard backgrounds, which do not provide any information about the true illumination condition of these scenes. They demonstrate that people can't compensate for changes in illumination and therefore, show poor gloss constancy. However, we predict that the reason observers show poor gloss constancy in this task is due to not having any information about the illumination context; they can only rely on priors over illumination, gloss and shape.

So far, we have conducted three experiments to understand the effect of context on gloss constancy. In Experiment 1 we asked whether contextual information about the light field has an effect on perceived gloss. In Experiment 2 we tested whether observers use simple image statistics such as luminance and contrast to compensate for the changes in the light field. Lastly, in Experiment 3 we investigated the effect of structure of the light field on gloss constancy.

### 3.1.1 General Methods

Stimuli used in all experiments were spheres warped with Perlin noise. We use these ‘blobby’ objects as they have novel shapes and therefore subjects can't do image matching using the image features due to the shape. The shapes of the objects are randomized both within and across trials, so any effects of specific shapes should average out.

We placed the blobby objects in a natural light field, which is a spherical environment where the object reflects light coming from all directions. We rendered the stimuli using Octane software. We used 4 light fields from the SYNS dataset (<https://syns.soton.ac.uk/>), which were cloudy coastal, cloudy residential, sunny beach and sunny forest (Figure 2). The reason we chose the sunny and cloudy light fields is because they differ in their contrast, which creates differences in the highlights on the object's surface, as shown in Figure 2. Sunny light fields have higher contrast than cloudy light fields and therefore, will create higher contrast highlights on the surface of a glossy object. On the other hand, a cloudy light field is lower contrast as the light is scattered in many directions, which creates more diffuse and blurry highlights on the surface of a glossy object.



Figure 2: A blobby object presented in all four light fields. From left to right: cloudy coastal, sunny beach, sunny forest and cloudy residential.

The surface glossiness was manipulated by varying the specular reflectance (proportion of light reflected specularly) and the specular roughness parameters. Instead of using the actual parameter values for reference, we are going to define the gloss levels of our stimuli in the range [1, 7] where gloss level 1 is the least glossy and gloss level 7 is the glossiest. In all experiments the standard stimulus had two gloss levels 3 and 5 and these were always fixed. The comparison stimulus had gloss levels ranging from 1 to 7. In each trial subjects compared the glossiness of a standard stimulus and a comparison stimulus.

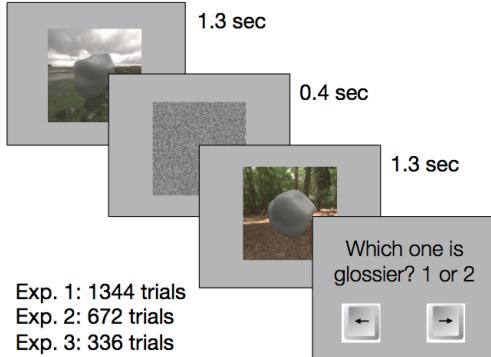


Figure 3: Experimental task. In a given trial subjects saw one stimulus followed by a noise mask and the second stimulus. Their task was to choose which object is glossier. They indicated their choice with a key press. No feedback was provided.

The experimental task was the same for all three experiments. Subjects viewed the first stimulus for 1.3 seconds and then a noise mask for 0.4 seconds and then the second stimulus for 1.3 seconds. In each stimulus a blobby object was shown with the light field context visible in the background. The subject's task was to choose which object; the first or the second was glossier (Figure 3). They indicated their response with a key press using the arrow keys, where left arrow key represented the choice for “first object is glossier” and the right arrow key indicated the choice for “second object is glossier”. In all experiments the order of all the trials and the presentation order of two stimuli were

randomly interleaved. We did not provide any feedback to the subjects about their responses.

We had seven subjects for Experiment 1, eleven subjects for Experiment 2 and ten subjects for Experiment 3. All subjects had normal or corrected to normal vision.

### 3.1.2 Apparatus

The display used for the experiment was a Dell P780 monitor with a resolution 1024x768. Prior to the experiment the display was calibrated for any non-uniformity in luminance. The experiment was implemented using MATLAB and Psychtoolbox Version 3 (Brainard, 1997; Kleiner, Brainard, Pelli, & Ingling, 2007; Pelli, 1997). The experiment ran on a Mac Pro with 2.66 GHz Dual Core Intel Xeon equipped with NVIDIA GeForce 7300GT graphics card. Subjects were seated away 57 cm away from the display with their head resting on a chinrest throughout the experiment. The experiment room was completely dark and the subjects viewed the stimuli monocularly using an eye-patch to avoid binocular depth cues.

### 3.1.3 Data Analysis

For each experiment we pooled the data of each subject, across all comparisons for fixed gloss levels 3 and 5 separately, and calculated the proportion of choosing comparison stimulus as glossier to total number of trials for every condition. For every subject we did psychometric fits using a cumulative normal function with the PALAMEDES toolbox for MATLAB (Prins & Kingdom, 2009). We used a maximum likelihood estimation to estimate parameters of threshold ( $\alpha$ ) and slope ( $\beta$ ) of the psychometric function which varied between 0 and 1 where 0.5 represents chance level performance.. We calculated the point of subjective equality (PSE) shift for each subject for all conditions for both gloss levels 3 and 5 and averaged the PSE shifts across all subjects.

### 3.1.4 Experiment 1

#### Methods

In Experiment 1, we provided explicit information about the illumination condition where objects were presented in either a congruent or an incongruent context. In every trial subjects were shown a standard and a comparison stimulus. To prepare the comparison stimulus, first we rendered a blobby object in the cloudy coastal environment. Then we cropped this object from the light field it was rendered in and

pasted it in all four light fields. We repeated this for all of our blobby objects. As a result, the object was presented in an incongruent light field context in the three conditions (sunny beach, sunny forest, cloudy residential) and was presented in the same, congruent light field (cloudy coastal) in one condition. We repeated the same steps for a comparison stimulus that was initially rendered in the sunny forest light field. For the standard stimulus, we always presented the object in the congruent context; an object rendered and shown against the same light field where the light field was either the cloudy coastal or the sunny forest.

## Predictions

When an object is reflecting a cloudy light field, we expect it to have lower contrast and blurrier highlights. A similar image effect would occur if the light field is sunny but the object surface is less specular and therefore spreads the specular reflections causing blurrier highlights. However, in these experiments, we are interested in the effect of changing the illumination on the highlight field.

We expect that if the comparison stimulus is rendered in a cloudy light field, we expect this object to have lower contrast and blurrier highlights. When this object is cropped and pasted into a sunny light field, the object should look even less glossy. This is due to the mismatch in the contrast of the scene and the highlights on the object's surface. When comparing the glossiness of this less glossy looking comparison stimulus to a congruent standard stimulus, you need to pick a higher gloss level comparison stimulus as the perceptual match. In other words, we would expect an increase; hence a rightward shift in the PSE in perceived glossiness. The opposite is also true, when a comparison stimulus is rendered in a sunny context but presented in an incongruent cloudy context. Such a comparison stimulus is going to appear even glossier. To match its glossiness to a congruent standard stimulus with a fixed gloss level, you need to pick a lower gloss level comparison stimulus. Therefore, we would predict a leftward shift or a decrease in the PSEs.

## Results

Data from Experiment 1 show that for conditions in which the standard stimulus was rendered in a cloudy light field, the predicted PSE shift was not present in all the conditions. We predicted positive PSE shifts for the conditions where the comparison stimulus was presented in an incongruent sunny context. We found a significant positive PSE shift ( $t(6)=2.69$ ,  $p<0.05$ ) only in the condition where the incongruent context was the sunny forest and the fixed gloss level was equal to 5. We also found a strong, significant evidence for the effect of context on perceived gloss in a condition we did not

expect to see any PSE shift. This was the condition where the incongruent context was another cloudy environment (cloudy residential) and we found a significant positive PSE shift for both gloss levels 3 and 5 ( $t(6)=5.17$ ,  $p<0.01$ ;  $t(6)=4.76$ ,  $p<0.01$ ). In other conditions there was no significant PSE shift (Figure 4).

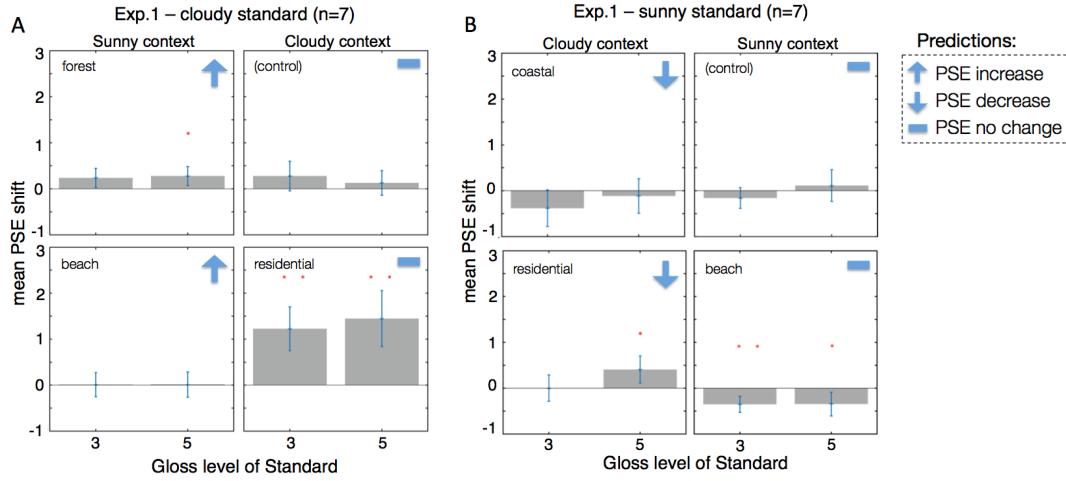


Figure 4: *Results of Experiment 1.* The plots show the mean PSE shift across all subjects ( $n=7$ ) for both gloss levels 3 and 5. The blue arrows show predicted PSE shift directions. (A) In these conditions the comparison stimulus is rendered in cloudy context and pasted into all four contexts (top left corner of each plot). The standard stimulus is always presented in the congruent cloudy coastal context. (B) In these conditions the comparison stimulus is rendered in a sunny context and pasted into all four contexts. The standard stimulus is always presented in the congruent sunny forest context. For both conditions “(control)” represents the comparison stimulus that was presented in the congruent context. The other three contexts were incongruent.

The results for the conditions in which the standard stimulus was rendered in a sunny light field did not show the predicted PSE shifts (Figure 4). Instead, we found unexpected significant PSE shifts in some of the other conditions. We found a significant positive PSE shift ( $t(6)=2.76$ ,  $p<0.05$ ) in the condition where the incongruent context was cloudy residential and the gloss level was 5. This was an opposite shift compared to the expected decrease in the mean PSE. We also found unexpected significant negative PSE shifts when the incongruent context was another sunny context (sunny beach) for both gloss levels 3 and 5 ( $t(6)=-4.00$ ,  $p<0.01$ ;  $t(6)=-2.68$ ,  $p<0.05$ ). Even though we did not find the predicted PSE shifts, the presence of significant shifts suggest that there is

an effect of context on perceived gloss. This means that there is an effect of providing explicit information about the illumination in making gloss judgments.

## Discussion

Even though some of the predicted effects are not seen in the results, the fact that there are significant PSE shifts in perceived gloss suggests that the context has an effect on gloss judgments. This means that subjects are using the visible context when they make judgments about the glossiness of an object presented in that context. However, it is not evident from these results what information the subjects might be using to make gloss judgments. We found an unexpected significant effect in the condition where subjects compared objects presented in cloudy coastal and cloudy residential contexts. This was a surprising finding, as we expected the largest effects when subjects were making comparisons between cloudy and sunny contexts.

In order to understand the underlying cause of this effect, one simple thing to check was the statistical properties of these two cloudy light fields. When we looked at the mean luminance of all four light fields we saw that the cloudy residential light field had the highest mean luminance. It was also three times as bright than the cloudy coastal light field. This leads to the hypothesis that, subjects might be using simple image statistics such as mean luminance and contrast to make gloss judgments. In order to test this, we conducted Experiment 2 where we manipulated the luminance and contrast of the context objects were presented in.

### 3.1.5 Experiment 2

## Methods

In Experiment 2, we again provided explicit information about the illumination condition. Objects were presented in either a congruent or an incongruent context but, different from Experiment 1, the incongruent context was created by manipulating the luminance and contrast of the original context. In Experiment 2, we created the standard stimulus by first rendering an object in the sunny forest context and then cropping and pasting this object into higher/lower luminance and higher/lower contrast manipulated versions of the original sunny forest light field. As a result, the standard stimulus was always presented against an incongruent context. For the comparison stimulus, we rendered an object in the sunny forest and cropped and pasted it into another sunny forest. We did this to keep the general color scheme of the light field consistent. Note that in this experiment, unlike Experiment 1, it was the standard

stimulus that was presented in an incongruent context and the comparison stimulus was always presented in the congruent context.

## Predictions

In Experiment 2 we predicted that when the standard stimulus was presented in an incongruent context with higher luminance or higher contrast, it should appear less glossy. Therefore, the subject is expected to choose a less glossy comparison stimulus to match the standard stimulus, thus we expect a decrease in PSEs. This is again due to the expected mismatch in the contrast/luminance of the context and the contrast/luminance of the highlights on the object surface. For example, when the mean luminance of an environment is low, we would expect to see dimmer highlights on the object's surface. So when we crop and paste this object into a high mean luminance context, it should appear even less glossy. The similar approach applies to mean contrast as well. Similarly, the opposite prediction is expected when the incongruent context is lower in luminance or contrast. In this case, the object should appear even more glossy and we predict an increase in PSEs for perceived gloss.

## Results

The results for Experiment 2 do not show a significant effect of the incongruent context in most conditions. We found a significant positive PSE shift ( $t(10)=2.55$ ,  $p<0.05$ ) in low contrast incongruent condition for gloss level 3. We did not find significant effects in other conditions (Figure 5).

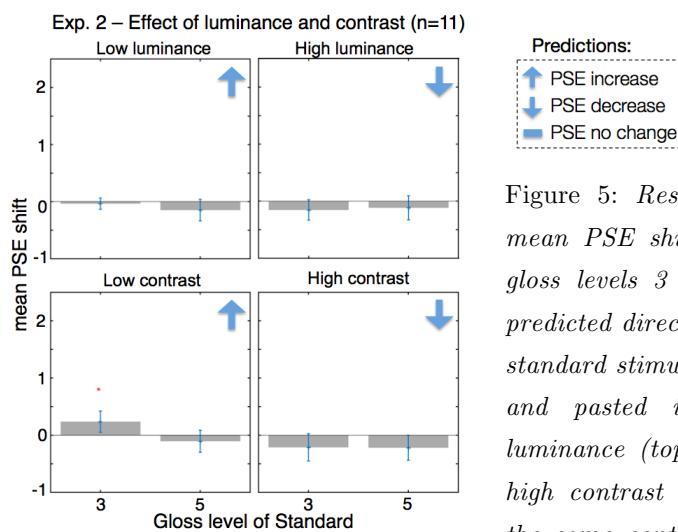


Figure 5: Results for Experiment 2. The plots show mean PSE shifts across all subjects ( $n=11$ ) for both gloss levels 3 and 5. The blue arrows represent the predicted direction of PSE shift. In this experiment the standard stimulus was rendered in sunny forest context and pasted into low luminance (top left), high luminance (top right), low contrast (bottom left) and high contrast (bottom right) manipulated versions of the same context. In all four conditions the standard stimulus was incongruent. The comparison stimulus was rendered in the sunny forest context and was presented in another similar sunny forest context.

## Discussion

The results of Experiment 2 suggest that simple image statistics such as luminance and contrast do not affect our gloss judgments. This means that subjects are not simply extracting these image statistics from the illumination contexts to infer surface gloss; they must be using some other information from the scene or a combination of many cues to make gloss judgments. Another source of information in the light field is the well-defined and structured shapes, lines and edges. If the light field has a lot of structure such as an indoor scene with clear lines, edges and shapes, this information would create a highlight field with similar properties. We are going to test the effect of light field structure by adding a fog effect to the sunny forest light field. Fog is a great way to reduce the amount of structure compared to phase scrambling because it spreads the reflected light and it still preserves the natural and realistic look of the light field. We are going to test the effect of adding fog in Experiment 3.

### 3.1.6 Experiment 3

#### Methods

In Experiment 3, we prepared the standard stimulus by rendering an object in the sunny forest light field and then cropping and pasting it into two light fields. The first was a foggy forest context, which was created by adding a fog effect to the original sunny forest light field. The fog effect is created in two steps. First, we combined the light field with the fog color (white) and then we added blur. Both the fog and the blur increased as a function of elevation because in natural outdoor scenes elevation is correlated with distance. Therefore, in the foggy light fields, the effect of fog increases with distance. The second light field was a different sunny forest with no fog. As a result, the object in the foggy context was incongruent and the object in the no-fog context was congruent. Comparison stimulus was rendered and shown against the same sunny forest light field and therefore, it was congruent.

#### Predictions

In Experiment 3, we pasted objects rendered in a high contrast light field into one that has been artificially blurred and reduced in contrast as if in fog. We expect that these pasted objects should appear glossier because their highlights are more distinct and higher in contrast than what is expected in a foggy context. Therefore, when the standard stimulus is presented in a foggy, incongruent context, we predicted that a

glossier comparison stimulus is required to match it. This will result in an increase in the PSEs.

## Results

For experiment 3, the prediction was a positive PSE shift when the incongruent context was foggy. The results matched this prediction when gloss level was equal to 3 ( $t(9)=2.83$ ,  $p<0.05$ ) but not for gloss level 5. As expected, the no-fog condition did not show any significant PSE shifts (Figure 6).

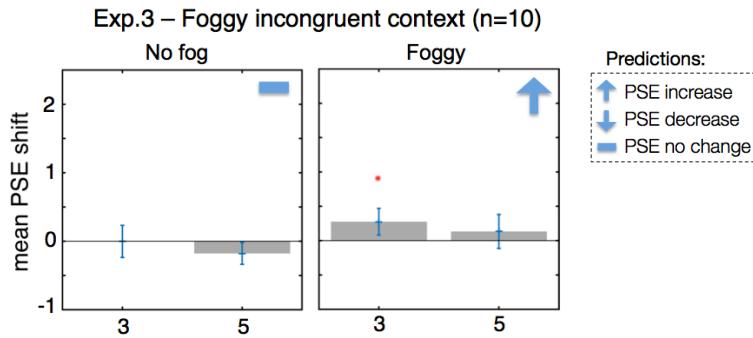


Figure 6: *Results of Experiment 3. The plots show mean PSE shifts across all subjects (n=11) for both gloss levels 3 and 5. The blue arrows represent the predicted direction of PSE shift. In this experiment standard stimulus was rendered in the sunny forest context and was presented in another sunny forest and a fog added forest. The comparison stimulus was presented in the congruent sunny forest.*

### 3.1.7 General Discussion

In Experiment 1 we found partial gloss constancy when objects were presented in incongruent contexts. Even though the effects we found did not match our predictions the fact that effects are present means that providing explicit information about the illumination modulates gloss judgments. However, these effects were not systematic. Certain contexts such as cloudy overcast illumination conditions affected gloss judgments more than others.

In Experiment 2 we tested whether observers are using simple image statistics like mean luminance or contrast of the illumination context to estimate the surface gloss. We manipulated the illuminations conditions so that they were high/low contrast and high/low luminance compared to the original illumination. The results showed no gloss constancy across changes in luminance and contrast. Thus suggests that observers must

be using information other than the simple image statistics of the context to estimate gloss.

In Experiment 3 we investigated the effects of reducing the structure of the context by making it foggy (blurry). This was a way to reduce the high-contrast information in the context but still keep the context look natural. Adding fog to the context did not cause significant effects on gloss constancy in both of the tested gloss levels.

### 3.1.8 Control Experiment

In previous work by Fleming, Dror and Adelson (2003) they assume that presenting an object with the congruent or incongruent light field context has negligible affect on the surface reflectance properties. However, this assumption has not been formally tested. We conducted a control experiment to test the difference in perceived gloss between conditions where the congruent context is visible and where the object is shown against a gray background.

Stimuli are blobby objects rendered in all four light fields. The standard stimuli for both with and without background conditions are rendered in cloudy coastal and sunny forest light fields and have fixed gloss levels 3 and 5. The comparison stimuli are rendered in all four light fields and have gloss levels ranging from 1 to 7 where 7 is the glossiest. For the with-background condition, stimuli are presented in the same, congruent context visible in the background. For the no background condition we cropped the objects from their congruent contexts and pasted them on a gray background.

In the condition where objects are presented on a gray background, the only information available to the subject is the image statistics available from the object's surface. This means, subjects will be matching image statistics from the highlight field to judge for gloss. We expect when both the comparison and standard stimulus are rendered in the same context, this matching should be more accurate. On the other hand, when subjects compare two objects rendered under different light fields, the objects will not provide reliable information about the illumination. As a result, subjects will try to compensate for the changes in surface gloss, which will lead to shifts in the PSEs.

In the condition where objects are presented with the congruent context visible in the background, the stimulus is going to provide reliable information about the illumination. Therefore, we expect subjects to be able to match gloss correctly and we do not expect to see any shifts in the PSEs.

The experimental task for both conditions was identical to the task in experiments 1-3. In the first condition all stimuli were presented on a gray background and subjects

picked whether the first or the second stimulus appeared glossier. In the second condition all stimuli were presented with their congruent context visible in the background and subjects repeated the same task. So far, we collected data from one subject (see Future Directions).

For the data analysis we were interested in the difference between no-background and with-background conditions. First, for both conditions we pooled the data for fixed gloss levels 3 and 5 separately, and calculated the proportion of choosing comparison stimulus as glossier to total number of trials for every condition. Next, we did psychometric fits using a cumulative normal function with the PALAMEDES toolbox for MATLAB (Prins & Kingdom, 2009). The settings of the fit were same as experiments 1-3. We calculated the PSE shift for with-background and no-background conditions for both gloss levels 3 and 5. Next, we subtracted no-background PSEs from with-background PSEs to see if there was an effect of background on the perceived gloss.

The results show a trend in the PSE shifts suggesting that the visible background has an effect on perceived gloss. We want to run the control experiment with more subjects after we generate new stimuli (see Future Directions).

### 3.1.9 Future Directions

There are several planned next steps for this project. The first plan of action is to make sure the objects are perceived to be a part of the context that they are presented in. To achieve this, we want to present objects sitting on a glass pedestal (Figure 7). The glass allows the observers to see the context through transparency, which makes it more believable that the objects are sitting in that particular environment. We will also show a larger portion of the context/light field. In the current status, the object is reflecting parts of the illumination field that are not visible in the stimulus. By increasing the visible portion of the light field we can make the stimuli look better integrated into the light field. After we render these new stimuli we also want to use them for the control experiment and collect data from more subjects.



Figure 7: *Stimuli shown in Experiments 1,2 and 3. The object looks pasted in the environment (left). New version of stimuli for future experiments is attached to a glass pedestal to make the object appear more embedded in the scene (right).*

We also want to test two other manipulations to the illumination. Fleming, Dror and Adelson (2003) suggest that power spectrum and luminance histogram characteristics of illumination can affect perceived gloss. We will manipulate (i) the spherical harmonic power spectrum of the illumination while preserving the luminance histogram and (ii) the luminance histogram while preserving the spherical harmonic power spectrum.

Proposed time for manuscript: Fall/Winter 2015

### 3.2 Are gloss and shape jointly estimated?

Previous work shows that shape and gloss interact; bumpier surfaces appear glossier and glossier surfaces appear bumpier (Ho et al., 2008). Even though this work provides us with an important interaction in perceived gloss, it does not explain how the visual system estimates gloss. The research in perceived gloss has been lacking in terms of matching physical dimensions of gloss and shape to perceptual dimensions. Most of the work in this area uses multidimensional scaling techniques to find the “perceptually meaningful” axes in gloss space (Obein et al., 2004; Pellacini et al., 2000). However, these methods have disadvantages as the resulting gloss space depends on the particular set of stimuli used for observer judgments. Work by Marlow and Anderson (2013) used image cues such as coverage, contrast and sharpness to explain how subjects estimate gloss. However, these cues were defined by the researchers and they were not based on what subjects might actually be using to estimate surface gloss. In this work, we want to

understand observers' direct estimates of gloss and shape (bumpiness). To do this, we generated physical gloss and bumpiness stimuli as a yardstick to measure an observer's direct gloss and shapes estimates.

### 3.2.1 Physical Stimuli

For this experiment we created both computer rendered and physical stimuli with different bump and gloss levels. We generated the computer rendered stimuli using the same method as Ho et al. (2008). An array of ellipsoids was created by distributing 400 ellipsoids pseudo-randomly in a 20x20cm grid where each ellipsoid had a radius of 1cm. The z radii of the ellipsoids were chosen randomly from a range  $[0, b_j]$  where  $b_j = [0.4, 0.8, 1.2, 1.6, \dots, 4.0, 4.4]$  in increments of 0.4cm. Different from the stimuli in Ho et al. (2008), we increased the range of z radii values and introduced smaller increments. We added random jitter in the range of  $\pm 0.2$ cm in the direction of x and y axes. The ellipsoids were allowed to overlap. The array was placed frontoparallel to the subject. We created the 3D models of the ellipsoid arrays using Blender software.

The physical stimuli for bumpiness were 3D printed versions of the ellipsoid arrays in a smaller, 5x5cm grid size with 25 ellipsoids. We used the same x, y and z radii as the computer rendered stimuli. We printed the arrays using the Mojo Desktop 3D printer. We painted the printed arrays to a mid-gloss gray color.

The physical stimuli for gloss are Ping-Pong balls painted with different gloss levels ranging from matte to very glossy. We first painted all our Ping-Pong balls with a flat gouache paint in gray. We created different gloss levels by mixing matte and glossy varnishes in different ratios. The gloss levels range from 0-100% in increments of 10%. For 0% gloss we left the object painted with flat gouache paint in gray color and did not apply any varnish. For 10-90% gloss levels we combined glossy and matte varnishes into a 10ml mixture and covered the object with a coat of this mixture. For example, for 10% gloss, we mixed 1ml glossy varnish with 9ml matte varnish. Similarly for 20% gloss, we mixed 2ml glossy varnish with 8ml matte varnish. The measurements were made with a 0.1ml sensitive syringe. We did this for all gloss levels 10-90%. For 100% gloss, we covered the object only with gloss varnish.

To match the physical gloss levels of our painted Ping-Pong balls to the rendered stimuli on the screen, we fit Ward bidirectional reflectance distribution function (BRDF) (Ward, 1992) parameters using the photos of the physical stimuli. We first photographed the Ping-Pong balls in a dark box environment. The Ping-Pong balls were placed in a cardboard box that was covered with black flock paper inside. The black paper created a dark room environment by not letting the light create interreflections. We cut a circular hole on the box perpendicular to the location of the Ping-Pong ball and fit the camera

lens into this opening tightly. On the top part of the left sidewall we cut another whole for the light source. The light source we used was a 50 W Bulbrite XP Aluminum Halogen Lamp. To convert the light source to a point light source we covered the front of the bulb with black paper and made a small hole to let the light through. The lamp was tightly fit into the opening on the left wall of the box to ensure no light is entering the box from outside.

With this setup we photographed all Ping-Pong balls. We used a Nikon D3200 camera with an 18-55mm lens. The ISO speed was 100, aperture was f/5.0, shutter speed was 1.0 sec and focal length was set to 18.0mm. We saved the photos in RAW format with an image size  $6080 \times 4012$ . We then converted the RAW images into PGM format using the [dcraw software](#).

### 3.2.2 BRDF Fitting

A bidirectional reflectance distribution function (BRDF) defines how light is reflected off of a surface. Ward BRDF (Ward, 1992) has four parameters to where  $\rho_d$  determines diffuse reflectance,  $\rho_s$  determines specular reflectance and  $\alpha_x$  and  $\alpha_y$  determine the spread or the roughness of highlights. As our objects have isotropic reflections we will set  $\alpha_x = \alpha_y$  and fit only one parameter  $\alpha$ .

The fitting is a 2-step grid search where in the first step we fix  $\alpha$  and fit only  $\rho_d$  and  $\rho_s$ . Once this grid search converges onto the best fit for this iteration, we move onto the second step where we fix the  $\rho_d$  and  $\rho_s$  to the best fit values from the previous step and fit only  $\alpha$ . We repeat these two steps until all three parameters converge. For every round during the grid search, we render a sphere using the three parameters  $\rho_d$ ,  $\rho_s$  and  $\alpha$ . This rendering is identical to the photograph of our physical stimulus and it matches the photographed stimulus in size, location and light source position. The renderings are done by RenderToolbox3 toolbox for MATLAB (Heasly, Cottaris, Lichtman, Xiao, & Brainard, 2014). The fit error is calculated by taking the pixelwise mean squared error between the rendered sphere and the photo of the physical stimulus. Parameters converge when all the parameter values are changing less than 1%.

We repeat the BRDF fitting for all 0-100% gloss level physical stimuli. As a result, for each gloss level we obtain a set of three parameters  $[\rho_d, \rho_s, \alpha]$  that match the physical gloss stimulus. Using these parameters we can generate matching computer rendered stimuli for gloss. The computer rendered stimuli will therefore match the physical stimuli in surface gloss.

### **3.2.3 Rendered Stimuli**

We generated the computer rendered stimuli with RenderToolbox3 (Heasly et al., 2014). The stimuli are 20x20cm arrays of ellipsoids with bump levels [1.6, 2.0, 2.4, 2.8, 3.2] cm and gloss levels if [30, 40, 50, 60, 70] percent. The reason we chose these values from our whole set of stimuli was to leave some range on both ends for subjects in case they over or under estimate bump and gloss levels. For each bump level, we rendered all 5 gloss level stimuli. The objects were rendered under a complex natural light field taken from the SYNS dataset (<https://syns.soton.ac.uk/>). We rendered the same stimuli in three different surface materials. The first is a Ward BRDF with a green albedo (CIE xyY: [0.305 0.478 23.4]), second is the grayscale version of the first and the third is a chrome material. We had 25 rendered images for each surface material.

### **3.2.4 Experimental Design**

The physical gloss stimuli will be ordered from matte to glossy and will be placed in a box with an open front. The gloss levels of the objects will be labeled. The subject will be able to look at them from different directions, but will not be able to touch the stimuli. The physical bump stimuli will be ordered as well but this time the subject will be able to pick up the objects and interact with them by touching.

In a given trial the observers will be asked to match either the bumpiness or the gloss level of computer-rendered stimulus with a physical stimulus. If the subjects are asked to match the gloss level of the stimulus on the screen, they will choose a physical gloss stimulus that has the perceived match for gloss. They will enter the gloss level of their choice to the computer. The response phase will not have a time limit and the stimulus will stay on the screen until the response is completed. If the subjects are asked to match the bump level of the stimulus on the screen, they will choose a physical bump stimulus that has the perceived match for bumpiness. The response procedure is as same as the gloss matching task.

Subjects will repeat the matching task for each stimulus for both glossiness and bumpiness. The order of computer-rendered stimuli will be randomized. In a session, subjects will complete 75 matching trials for each bump and gloss matching which will result in a total of 150 trials. We will repeat each session 4 times.

### **3.2.5 Project Goals**

One of the goals of this project is to understand the error in gloss and shape estimates. By asking subjects to match a rendered stimulus to a physical stimulus, we can estimate an observer's percept on an absolute scale, as we have the shape and gloss

measurements of our physical stimuli. Given these estimates we can also calculate the error in shape and gloss judgments. Then we will parameterize the highlight field (coverage, sharpness, contrast of highlights and skew of the luminance histogram) and do a regression to predict the subject's gloss measurements and compare it to ground-truth gloss. We also want to understand whether the results support the hypothesis that gloss and shape are jointly estimated or if there is a generalized linear model that can predict perceived gloss across all stimuli from image gloss cues such as coverage, sharpness and contrast (Marlow & Anderson, 2013).

Proposed time for manuscript: Winter/Spring 2016

### 3.3 Do additional cues influence material perception?

When we observe objects in daily life they do not usually appear stationary, we do not view them monocularly and we even interact with them by touch. Motion, surface texture, binocular disparity and touch provide us with additional information about an object's shape and material.

Motion information is known to contribute to better 3D shape estimation (Ullman, 1979). It also helps disambiguate glossy reflections from paint-like texture (Doerschner et al., 2011; Hartung & Kersten, 2002). Surface texture provides veridical information about shape and material especially when the object is in motion (Doerschner, Yilmaz, Kucukoglu, & Fleming, 2013). Binocular disparity is known to improve gloss constancy (Wendt, Faul, Ekroll, & Mausfeld, 2010) and can also be used to separate gloss from 3D-structure information (Marlow & Anderson, 2015). In addition to these ones, physical interaction with objects also improves tactile judgments for glossy surfaces compared to when visual information is used alone (Bergmann Tiest & Kappers, 2007).

The goal of this work is to enhance shape cues and surface material cues by adding haptic cues without modifying the highlight field. In a recently published paper by Marlow and Anderson (2015), they use binocular disparity as a cue to 3D shape without altering the image cues. They show that the underlying 3D shape information determines whether the same image gradient is seen as a matte surface or a glossy surface.

In this study we are going to use a haptic phantom setup to introduce haptic cues to shape and gloss. The subjects are going to make judgments about surface shape and surface gloss by while viewing visual stimulus and getting haptic feedback at the same time. With touch information we will introduce a cue without changing the visual gloss

information and we will be able to study the two-way interaction of haptic shape/gloss and visual shape/gloss.

### 3.3.1 Stimuli

The shape stimulus will be a single half ellipsoid with different bump levels. The bump level will be adjusted by setting the z-radius of the half ellipsoid. The z-radius values range from [0.4, 4.4] cm in increments of 0.4cm. The x and y radii of the ellipsoid are 1cm. These objects will be rendered with the physically based Mitsuba renderer using RenderToolbox3 (Heasly et al., 2014) under natural, complex light fields from the SYNS dataset (<https://syns.soton.ac.uk/>). The same shapes will be rendered for the haptic shape stimuli using OpenHaptics toolkit.

For every bump level we are going to render visual stimuli ranging from 0% glossy (matte) to 100% glossy (mirror). 50% gloss will be set as the baseline. The gloss is manipulated by varying the specular reflectance and the spread (roughness) parameters in the Ward BRDF. We are going to modulate the haptic gloss by varying the surface friction. It has been shown that objects are more likely to be classified as matte when they feel rubbery to touch and more likely to be classified as glossy when they feel slippery (cite: Adams et al., 2014). We will adjust friction by varying the compliance-related forces in the direction tangential to the surface. In terms of touch this is going to determine how easily you can slide your finger across the surface. When friction is high, surfaces are going to feel more rubbery and when the friction is low, surfaces are going to feel slippery to touch.

### 3.3.2 Experimental Design

We will be using a visual-haptic setup, which allows simultaneous presentation of visual and haptic stimulus. Haptic information will be presented using a PHANToM haptic feedback device. The visual stimulus will be presented via a monitor and it will be reflected off of a mirror to create the perception of touching and viewing the same object simultaneously (Figure 8). Observers will use an eye patch to remove binocular cues to depth and gloss.

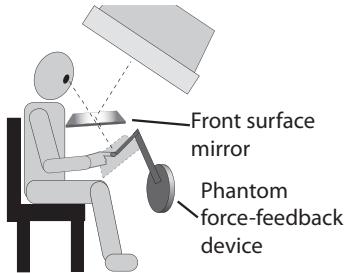


Figure 8: *The visual-haptic experiment setup. Subjects will be viewing a visual stimulus reflected from the monitor with a mirror. The touch information will be presented with a phantom haptic feedback device.*

In this experiment we are going to test the effects of haptic shape and haptic friction on perceived shape and perceived gloss. In some conditions we are going to present the visual stimulus with haptic shape information constant or haptic friction information constant while the other haptic modality is varying. We will also have a condition to test the perceived shape and gloss of visual stimulus when both haptic cues are present.

In the haptic shape condition we will present a visual stimuli with varying haptic shape information while haptic friction is kept constant. In the first condition, subjects will be doing a 2IFC task and they will be instructed to choose the bumpier stimulus between a pair of visual stimuli presented in two intervals. They will be getting haptic feedback about the shape during each trial. In the second condition, subjects are going to repeat the same task, but this time they will pick the glossier visual stimulus while receiving haptic shape feedback.

In the haptic gloss condition we will repeat the same task mentioned above but this time, we will provide only the haptic gloss information with the visual stimulus. The subjects are going to repeat both tasks for picking bumpier and picking glossier visual stimulus.

In the combined haptic condition, we will provide both haptic shape and friction information while presenting the visual stimulus. We are going to repeat both tasks for judging glossiness and bumpiness of the visual stimulus.

### 3.3.3 Project Goals

If shape and material are estimated jointly, we would expect to see more veridical shape perception to cause more veridical gloss perception. This means that, when we introduce haptic cues to 3D shape, we would expect people to have less error in their gloss judgments. For example, when we present only haptic gloss information with the visual stimulus, we expect that the gloss-bumpiness interaction (Ho et al., 2008) will

disappear as the haptic gloss will cause more veridical gloss perception. On the other hand, if the image cue approach holds, then as we are not going to modify the highlight field, we would expect to see no effect of additional cues on perceived gloss.

## Bibliography

- Anderson, B. L., & Kim, J. (2009). Image statistics do not explain the perception of gloss and lightness. *Journal of Vision*, 9(11), 10–10. <http://doi.org/10.1167/9.11.10>
- Bergmann Tiest, W. M., & Kappers, A. M. L. (2007). Haptic and visual perception of roughness. *Acta Psychologica*, 124(2), 177–189.  
<http://doi.org/10.1016/j.actpsy.2006.03.002>
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*.
- 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. <http://doi.org/10.1016/j.cub.2011.10.036>
- 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), 1–23.  
<http://doi.org/10.1167/13.11.8>
- Fleming, R. W. (2012). Human perception: Visual heuristics in the perception of glossiness. *Current Biology : CB*, 22(20), R865–6.  
<http://doi.org/10.1016/j.cub.2012.08.030>
- 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.  
<http://doi.org/10.1038/nature05724>
- Hartung, B., & Kersten, D. (2002). Distinguishing shiny from matte. *Journal of Vision*, 2(7), 551–551. <http://doi.org/10.1167/2.7.551>
- Heasly, B. S., Cottaris, N. P., Lichtman, D. P., Xiao, B., & Brainard, D. H. (2014). RenderToolbox3: MATLAB tools that facilitate physically based stimulus rendering for vision research. *Journal of Vision*, 14(2), 6. <http://doi.org/10.1167/14.2.6>
- Ho, Y.-X., Landy, M. S., & Maloney, L. T. (2008). Conjoint Measurement of Gloss and Surface Texture. *Psychological Science*, 19(2), 196–204.  
<http://doi.org/10.1111/j.1467-9280.2008.02067.x>
- Kleiner, M., Brainard, D., Pelli, D., & Ingling, A. (2007). What's new in Psychtoolbox-3.
- Marlow, P. J., & Anderson, B. L. (2013). Generative constraints on image cues for

perceived gloss. *Journal of Vision*, 13(14). <http://doi.org/10.1167/13.14.2>

Marlow, P. J., & Anderson, B. L. (2015). Material properties derived from three-dimensional shape representations. *Vision Research*, 1–10.  
<http://doi.org/10.1016/j.visres.2015.05.003>

Marlow, P. J., Kim, J., & Anderson, B. L. (2012). The Perception and Misperception of Specular Surface Reflectance. *Current Biology*, 22(20), 1909–1913.  
<http://doi.org/10.1016/j.cub.2012.08.009>

Marlow, P., Kim, J., & Anderson, B. L. (2011). The role of brightness and orientation congruence in the perception of surface gloss. *Journal of Vision*, 11(9), 16–16.  
<http://doi.org/10.1167/11.9.16>

Motoyoshi, I., Nishida, S., Sharan, L., & Adelson, E. H. (2007). Image statistics and the perception of surface qualities. *Nature*, 447(7141), 206–209.  
<http://doi.org/10.1038/nature05724>

Obein, G., Knoblauch, K., & Vieot, F. (2004). Difference scaling of gloss: Nonlinearity, binocularly, and constancy. *Journal of Vision*, 4(9), 4–4.  
<http://doi.org/10.1167/4.9.4>

Olkonen, M., & Brainard, D. H. (2010). Perceived glossiness and lightness under real-world illumination. *Journal of Vision*, 10(9), 5–5. <http://doi.org/10.1167/10.9.5>

Olkonen, M., & Brainard, D. H. (2011). Joint effects of illumination geometry and object shape in the perception of surface reflectance. *I-Perception*, 2(9), 1014–1034.  
<http://doi.org/10.1088/iop/2/9/014080>

Pellacini, F., Ferwerda, J. A., & Greenberg, D. P. (2000). Toward a psychophysically-based light reflection model for image synthesis. Presented at the Proceedings of the 27th ....

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

Ullman, S. (1979). The interpretation of visual motion. Oxford, England: Massachusetts Inst of Technology Pr.

Vangorp, P., Laurijssen, J., & Dutré, P. (2007). The influence of shape on the perception of material reflectance. *ACM Transactions on Graphics*, 26(99), 77–10.  
<http://doi.org/10.1145/1239451.1239528>

Ward, G. J. (1992). Measuring and modeling anisotropic reflection. *ACM SIGGRAPH Computer Graphics*, 26(2), 265–272. <http://doi.org/10.1145/133994.134078>

Wendt, G., Faul, F., Ekroll, V., & Mausfeld, R. (2010). Disparity, motion, and color information improve gloss constancy performance. *Journal of Vision*, 10(9), 7–7. <http://doi.org/10.1167/10.9.7>