

# Accepted Manuscript

Visual perception of materials and their properties

Roland W. Fleming

PII: S0042-6989(13)00278-2

DOI: <http://dx.doi.org/10.1016/j.visres.2013.11.004>

Reference: VR 6740

To appear in: *Vision Research*

Received Date: 16 September 2013

Revised Date: 12 November 2013



Please cite this article as: Fleming, R.W., Visual perception of materials and their properties, *Vision Research* (2013), doi: <http://dx.doi.org/10.1016/j.visres.2013.11.004>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## Visual perception of materials and their properties

Roland W. Fleming

*Experimental Psychology*

*Justus-Liebig-Universität Gießen*

### ABSTRACT

Misidentifying materials—such as mistaking soap for pâté, or *vice versa*—could lead to some pretty messy mishaps. Fortunately, we rarely suffer such indignities, thanks largely to our outstanding ability to recognize materials—and identify their properties—by sight. In everyday life, we encounter an enormous variety of materials, which we usually distinguish effortlessly and without error. However, despite its subjective ease, material perception poses the visual system with some unique and significant challenges, because a given material can take on many different appearances depending on the lighting, viewpoint and shape. Here, I use observations from recent research on material perception to outline a general theory of material perception, in which I suggest that the visual system does not actually estimate physical parameters of materials and objects. Instead—I argue—the brain is remarkably adept at building ‘statistical generative models’ that capture the natural degrees of variation in appearance between samples. For example, when determining perceived glossiness, the brain does not estimate parameters of the BRDF. Instead, it uses a constellation of low- and mid-level image measurements to characterize the extent to which the surface manifests specular reflections. I argue that these ‘statistical appearance models’ are both more expressive and easier to compute than physical parameters, and therefore represent a powerful middle way between a ‘bag of tricks’ and ‘inverse optics’.

KEYWORDS: materials, surface perception, computational models, theory.

## BACKGROUND

Different materials—such as soap, velvet and pâté—have distinct physical and functional properties, which determine how we can use them; for example, whether they are good for washing, wearing or eating, respectively. Being able to visually distinguish between materials and infer their properties by sight, is invaluable for many tasks. For example, when determining edibility, we can make subtle visual judgments of material properties to determine whether fruit is ripe, whether soup has been left to go cold or whether bread is going stale. When walking or climbing, the ability to judge whether a surface is slippery or fragile is critical for selecting foot- and handholds. Evidently, material perception is useful. One obvious question this raises is, are we any good at it?

Everyday experience, suggests that we are. We effortlessly distinguish numerous different categories of material: textiles, stones, liquids, foodstuffs, and so on, and can recognize many specific materials within each class such as silk, wool, cotton, etc. Indeed, it seems plausible that our capacity to categorize and recognize materials probably rivals our capacity to categorize and recognize objects—after all, every object is made out of some kind of materials, and we usually know which ones. Indeed, as Adelson (2001) points out, not everything that we can recognize is what we would normally call an ‘object’. Some ‘stuff’—like snow, sand or soil—is just ‘stuff’, without a clearly defined shape. In many cases such materials are not subject to key constraints—like cohesion and indivisibility—which we usually associate with ‘objecthood’. Despite this, we usually experience no problems recognizing such materials.

There is experimental evidence to support the intuition that human observers are good at recognizing and categorizing materials. For example, Sharan, Rosenholtz

& Adelson (2009) have shown that subjects can identify a wide range of materials from photographs even with brief presentations. Recently, Fleming, Wiebel and Gegenfurtner (2013) showed subjects photographs of materials from different categories and asked them to rate various subjective qualities, such as hardness, glossiness and prettiness. Even though subjects were not explicitly informed that the samples belonged to different classes, the subjective ratings of the individual samples were systematically clustered into categories, suggesting that subjects could theoretically classify materials through visual judgments of their properties.

At the same time, there is almost certainly more to material perception than our ability to categorize or recognize familiar materials. In general, without actually touching an object, we usually have a clear idea of what it would feel like were we to reach out and handle it: whether it would be hard or soft, rough or smooth, malleable or likely to crumble in response to force. Even with unfamiliar materials, we seem to be acutely aware of their specific visual and physical characteristics—is it sticky, runny, spongy, would it feel cold to the touch? We can usually answer such questions based on a material's visual appearance. In other words, in addition to recognizing and categorizing materials, we also form a vivid impression of their *material properties*.

In many cases, of course, physical and functional properties—such as density, thermal conductivity or toxicity—cannot be seen directly, so our impressions must presumably be learned associations. Nevertheless, many quite complex material properties do have a distinctive and vivid visual phenomenology: the frothy head of a freshly poured wheat-beer, for example, has a characteristic 'look', which is subjectively intimately associated with its physical properties. Because of this rich phenomenology, product designers go to great lengths in developing the visual 'look and feel' of consumer products, selecting and synthesising specific materials to elicit

a particular impression of the product as a whole. If we weren't highly sensitive to material appearance, it surely would not be profitable for companies to invest resources in perfecting complex paints and other surface finishing techniques. Indeed, material appearance plays a disproportionate role in the assignment of value to things. Precious metals and gemstones are not especially useful, yet they command high prices, largely because of their lustrous appearance. Again, humans appear to derive a compelling sense of material properties through vision.

There is a growing body of experimental evidence to back this up. For example, Sharan, Rosenholtz and Adelson (2008) tested how well subjects distinguish between photographs of 'real' and 'fake' materials—for example real fruit *vs.* realistic wax simulacra—in brief presentations. They found that even with presentation times of just 40 ms, subjects were able to make remarkably precise descriptions of the properties of materials and were above chance performance at distinguishing between real and fake materials. This is impressive because the image differences between real and fake materials are usually far from trivial to define. Real and fake materials have highly variable but overlapping appearances, which cannot easily be distinguished based on the overall colour distributions, intensities, contrasts or spatial attributes of the images. Clearly there is *something* about the 'look' of the real and fake materials that subjects rapidly identify, but what exactly comprises these—often subtle—appearance differences is not at all clear. Nevertheless, the empirical finding supports the intuition that we can make often quite subtle judgments of material attributes.

Other work has focussed on the visual estimation of specific properties of materials, such as glossiness, translucency or surface roughness (for recent reviews see Anderson, 2011; Thompson, Fleming, Creem-Regehr and Steffanucci, 2011 or Zaidi, 2011). For example, on the topic of glossiness, Nishida and Shinya (1998)

showed that subjects can judge the specular reflectance of computer simulated glossy surfaces and Fleming Dror and Adelson (2003), showed that this ability generalizes across differences in lighting, as long as the illumination has statistical structure that is typical of the natural environment. Motoyoshi and Matoba (2012) showed that varying the statistical characteristics of the illumination has systematic effects on perceived glossiness, which can be predicted from the low-level properties of the image. Judgments of specular reflectance are affected by both binocular disparity and motion information (Koenderink and van Doorn, 1980; Blake and Bühlhoff, 1990; Hurlbert, Cumming and Parker, 1991; Wendt, Faul and Mausfeld, 2008; Doerschner, Fleming, Yilmaz, Schrater, Hartung and Kersten, 2011; Murry, Welchman, Blake and Fleming, 2013), as well as the properties of highlights, including their brightness, position and orientation relative to diffuse shading on the surface (Beck and Prazdny, 1981; Fleming, Torralba and Adelson, 2004; Todd, Norman and Mingolla, 2004; Berzhanskaya, Swaminathan, Beck, and Mingolla, 2005; Kim, Marlow and Anderson, 2011; Marlow, Kim and Anderson, 2011, 2012). What cues does the visual system use to infer glossiness? Motoyoshi, Nishida, Sharan and Adelson (2007) found that glossy and matte stucco reliefs create different luminance (and sub-band) distributions, and suggested that the visual system could use the skewness of these histograms to distinguish between glossy and matte surfaces. They found that increasing the skewness of images of matte stucco reliefs made the surfaces appear glossy. However, others have noted that skewness is neither necessary nor sufficient to predict perceived glossiness, and have called into question the idea that such simple image statistics could account for surface perception more generally (Kim and Anderson, 2008; Anderson and Kim, 2009). Olkkonen and Brainard (2010, 2011) measured how perceived gloss varied as a function of illumination geometry, object shape and specular reflectance parameters, and also found that subjective matches

were poorly predicted by summary statistics (like skewness) derived from the intensity histogram.

On the topic of surface roughness, several authors have discussed how the visual system estimates and represents the characteristics of surface relief (e.g., Pont and Koenderink, 2005, 2008; Padilla, Drbohlav, Green, Spence and Chantler, 2008), although it remain unclear exactly which parameters of surface perturbations (e.g., scale, amplitude or profile) determine visual roughness, or indeed whether subjective roughness is a unitary quantity. Others have investigated how visual roughness relates to haptic impressions of roughness (Bergmann Tiest and Kappers, 2007), although it is still not clear how the brain compares or integrates the two. Ho, Landy and Maloney (2006) have shown that subjects' judgments of surface roughness are systematically biased by the illumination. They found that glancing illumination angles make surfaces appear rougher than frontal illumination.

Numerous other studies have investigated how we perceive the lightness, colour and opacity of thin transparent filters (Metelli, 1970; 1974a,b; Gerbino, 1994; D'Zmura, Colantoni, Knobacuh and Laget, 1997; Robilotto, Khang and Zaidi, 2002; Singh and Anderson, 2002a,b). By studying the structure of images created by transparent surfaces, a number of authors have identified photometric and geometric conditions that cause the visual system to separate single image intensity values into multiple causal layers—a process known as 'scission' (Beck, Prazdny and Ivry, 1984, Beck and Ivry, 1988; Adelson and Anandan, 1990; Anderson 1997, 2003). For example, thin transparent layers tend to create 'X-junctions' in the image, where the boundary of the transparent layer crosses over contours in the background layer. However, solid transparent and translucent objects—like an ice-cube or wax candle—behave quite differently from thin transparent filters, and appear subjectively to transmit light even when these photometric and geometric image

conditions are not met (Fleming and Bühlhoff, 2005). With solid translucent materials, light scatters within the body of the object, leading to a characteristic soft, glowing appearance. It is known that perceived translucency is affected by the thickness of the material, the direction of illumination, and colour properties of the image. However, how the visual system distinguishes shading gradients that are caused by opaque reflectance from those that are caused by sub-surface scattering remains unclear, although shadow regions are likely to play a role, as these are the portions of objects that are most affected by light that has passed through the object (Fleming and Bühlhoff, 2005). Motoyoshi (2010) notes that because translucency has much larger effects on shading than on specular highlights, relationships between shading and highlights provide important information about whether an object is translucent. He shows that varying the contrast (both magnitude and sign) and blur of the non-specular components of an object can dramatically alter its appearance from diffuse to translucent. Fleming, Jäkel and Maloney (2011) showed that subjects could match the refractive index of solid transparent materials, although, again their judgments were substantially biased by the thickness of the object and the distance to the background.

Taken together, these findings seem to support the general idea that—at least in certain circumstances—the human visual system can estimate the properties of materials from the retinal images. With this background in mind, it is interesting to ask *how* the visual system categorizes materials and infers their properties from the retinal images. Our theoretical understanding of material perception is very much in its infancy, but in the following section, we sketch out some of the key theoretical challenges posed by material perception.

## COMPUTATIONAL GOALS AND CHALLENGES OF MATERIAL PERCEPTION

Material perception can play many different roles in a wide variety of tasks, from judging surface friction when working out how to pick up an object, to choosing which scarf to wear. Depending on the particular goals, different levels of fidelity of material perception may be important. In some cases, it is sufficient to make a categorical judgment of some material attribute (e.g., wet or dry), in other cases, such as selecting between surface finishes for product design, extremely subtle distinctions are necessary. However, despite the wide range of high-level uses to which material perception can be put, I suggest that it is useful to group the underlying visual processing broadly into two kinds of computations: *categorization* and *estimation*.

It is important to point out that these are not mutually exclusive processes. On the contrary, they are likely to be highly interdependent. Estimating material properties (glossiness, roughness, colour, etc.) is likely to be a key stage in establishing the feature space within which samples can be categorized. Conversely, knowing which class a material belongs to presumably helps infer its properties (Fleming *et al.*, 2013). Thus, the two processes represent two ends of a continuum in terms of the fidelity of the internal representation of materials. Nevertheless, in order to specify the key computational challenges posed by material perception, it is useful to highlight the differences between categorization and estimation.

### **Material Categorization**

The main goal of categorization is to *assign a specific class label to a given visual sample of a material*, or put simply, to work out what kind of material it is. Note that this does not necessarily require a high-precision representation of the material's properties, as the end result is a simple label. The primary benefit of categorization

is that it provides access to stored knowledge about other members of the same class. This is especially useful for inferring characteristics, such as toxicity, which cannot easily be visually inferred for unfamiliar samples. Categorization also has other benefits, such as reducing information, by replacing a complex, high dimensional visual representation of the sample (e.g. in terms of 2D image features, or 3D shape and surface characteristics) with a much simpler, lower-dimensional label. Categories can be structured hierarchically from super-ordinate categories down to the recognition of individual exemplars (e.g. Textiles  $\supseteq$  Shirt cloth  $\supseteq$  Egyptian 2-ply cotton with blue pinstripe pattern), thus providing elements of a semantic system for understanding how the world of materials is organized.

From a computational point of view, material categorization is much like categorization of any other kind of entity (e.g., object categorization), in which the key computational challenge is establishing category boundaries through experience with exemplars. Theoretical approaches to categorization typically represent different exemplars as points in a high-dimensional 'feature space', and infer category boundaries in this space either using *unsupervised learning* techniques, such as clustering, which identify natural modes in the distribution of samples within the feature space, or through *supervised learning* techniques, such as support vector machines, which use explicit knowledge of class membership during training to determine boundaries. Having established category boundaries through learning, novel exemplars are classified simply by comparing their position in the feature space to these boundaries.

It seems reasonable that a theory of human material categorization could proceed more or less along these standard lines, although attempts to achieve human levels of categorization performance have not been very successful so far. For example, Liu *et al.* (2010) developed a Bayesian model for classifying photos of

different materials, based on a large number of low- and mid-level image features. Through training, the algorithm identifies and combines the most effective image features, but performance was only 44.6% correct with just ten material classes.

One reason for the relatively poor performance may be that materials can take on an enormous variety of different appearances, possibly even larger than scenes or objects, for example. Despite the fact that scenes—such as offices, beaches or forests—may contain many elements, there is a surprising degree of regularity in the overall spatial organization of typical exemplars, at least for typical views (Oliva and Torralba, 2001; Torralba and Oliva, 2003). This regularity enables quite reliable scene classification through simple image descriptions, like the ‘gist operator’ (Oliva and Torralba, 2006; Greene and Oliva, 2009). Object appearance can change dramatically depending on lighting, viewpoint and other factors, which makes successful object recognition and categorization challenging (see Rust and Stocker, 2010 for a recent discussion of the challenges). However, objects do at least typically have a well-defined shape. Even mutable objects, such as animals, tend to have a fixed topological structure, with distinctive shape features that are invariant across poses and shared by members of a common class.

By comparison, for a broad material class like ‘plastics’, the variation in possible appearance is huge (Figure 1): polythene bags, children’s toys and swimming goggles have widely diverging shapes and appearances. Even a given exemplar, such as a plastic bag, can take on many different shapes depending on the particular series of forces and processes to which it is subjected. Because plastic can be made into almost any shape and can have almost any colour, these features are much less diagnostic than for objects. Thus, material categorization presents the visual system with the significant challenge of enormous within-class variability.



FIGURE 1. Examples of different plastics with diverse visual appearances. Images from the MIT-Flickr materials database (Liu *et al*, 2010).

Recent trends in computer vision and computer graphics have emphasized the power of very large quantities of training data (e.g., Torralba *et al*, 2008; Hays and Efros, 2007). It is interesting to speculate whether one key to the effectiveness of human material categorization might simply be the massive quantity of materials that we have seen and remembered in our past. Perhaps training a system on comparably large quantities of training data would yield comparably good model performance.

### Estimation of Material Properties

In contrast to categorization, the main goal of estimation is to *identify specific characteristics of given sample of a material*, such as its specular reflectance properties or elasticity. In everyday language we use colour terms and words like ‘soft’, ‘lustrous’ or ‘sticky’ to describe the characteristics of materials. Visual estimation refers to the process of working out such properties. Estimation is useful for making subtle discriminations between similar materials, and allows us, for example, to predict how a given material would be likely to respond to external forces, irrespective of

whether we have seen such a sample before. Whereas categorization reduces information and tends to group materials together into nominal classes (at least in the limit), estimation tends to deal with metric differences between materials along continuous parameters. For example, if presented with a piece of glazed ceramic, we not only make the categorical judgment that the surface is glossy rather than matte, but also have a perceptual impression of the *degree of glossiness* along some continuous subjective scale. In this respect, estimation is computationally more challenging in the sense that it results in a higher fidelity of representation (i.e., higher information) than categorization. It is also worth noting that while categorization deals with the identity of materials independently of their specific characteristics (e.g., quartz and granite are both stones, but they differ in terms of their transparency), estimation deals with properties that may be common to materials of very different classes (e.g., both quartz and water are transparent and glossy). In this sense, categorization and estimation may be complementary to one another.

In many cases, physics has developed sophisticated descriptions of material properties. For example, the way light is reflected from a surface is completely described using the *bidirectional reflectance distribution function* (BRDF; Nicodemus, 1965; Nicodemus, Richmond, Hsai, Ginsberg and Limperis, 1977), which measures the proportion of light reflected in every direction as a function of the amount of light arriving from every direction in the hemisphere above the local surface tangent plane. Differences in appearance between surfaces like glossy plastics, brushed aluminium or matte paint are fully captured by the BRDF. In principle, if the visual system estimated the BRDFs of surfaces, it could represent the reflectance properties of any arbitrary material. However, even if we ignore wavelength variations, the BRDF is a function of four variables (two incoming angles and two outgoing angles),

making representing arbitrary BRDFs computationally costly. Fortunately, however, the BRDFs of real materials are highly constrained and represent only a tiny subspace within the set of all possible BRDFs. In fact, the reflectance properties of many materials can be quite well approximated by analytical BRDF models with just a handful of parameters (Torrance and Sparrow, 1967; Ward, 1992; Oren and Nayar, 1994; Ashikhmin and Shirley, 2000; Matusik, Pfister, Brand and McMillan, 2003), and even quite complex materials like multi-layered paints can often be modelled as the linear combination of a few such layers (e.g., Günther *et al.*, 2005).

Most analytical models separate reflectance into diffuse and specular components and have parameters controlling the relative weight of these terms (e.g., albedo and specular reflectance) and their angular distributions (e.g., the spread and anisotropy of the specular lobe). Varying these parameters has broadly intuitive perceptual consequences for the appearance of the surface (Figure 2). For example, increasing the albedo makes the surface appear a lighter shade of grey; increasing the specular reflectance makes the surface appear glossier (from left to right in the figure), and increasing the anisotropy of the specular lobe makes the highlights elongate as seen on varnished wood or brushed metal. Thus, it has become natural to pose the visual perception of surface reflectance as the process through which the visual system estimates these physical parameters. When we say that we see a surface as having a certain degree of glossiness, it is commonly assumed that this is because the visual system has estimated a specific value of specular reflectance for that material.

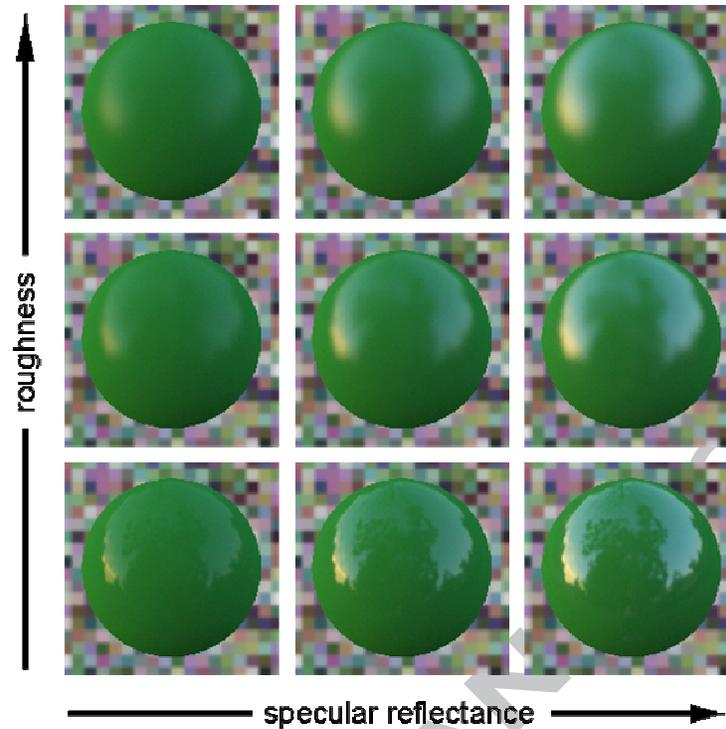


FIGURE 2. Varying the parameters of a BRDF model (Ward, 1992) leads to continuous changes in the appearance of the surface. It is common to pose visual reflectance estimation as the process of identifying the values of the parameters of a reflectance model.

With this in mind, it is typically assumed that as a computational process, estimation takes some image information (“cues”) as input and returns as output a visual estimate of various physical parameters of the material, perhaps along with a measure of the reliability or certainty of the estimates. For example, returning to the example of a piece of glazed ceramic, it seems quite natural to pose gloss perception as the process of making various measurements of the highlights visible on the surface to derive an estimate of the magnitude of specular reflectance of the sample. When posed this way, the two key scientific questions raised by the visual estimation of material properties are: (1) ‘what are the cues?’ – in other words, which image information does the visual system rely on to derive its estimate? – and (2) ‘how does the visual system compute the target material property (glossiness, translucency,

viscosity, etc.) from these cues?'. Most research on material estimation has focused on one or both of these questions. However, as I argue below, formulating both the goal and process of material perception in terms of *estimating physical properties* may be problematic. I suggest that by reformulating what we think 'material perception' means, we may stand a better chance of explaining some of the empirical curiosities that have emerged in the study of material perception (see below), and perhaps also unify categorization and estimation to some degree.

In the following sections, we first discuss the key computational challenges posed by material perception; then discuss the two main theoretical approaches that have emerged to account for how the visual system overcomes these. Using recent experiments that aim to shed light on the mapping between cues and material properties, I suggest that there may in fact be a third way of posing material perception.

### **What makes material perception difficult?**

Imagine being presented with two spheres: one is made of a highly polished chrome-like material, and the other of pearlescent plastic material, as shown in Figure 3. Because the two surfaces appear so different, it can be hard to appreciate what might be difficult about estimating reflectance properties from the retinal image. However, it is important to remember that the input to the visual system is highly ambiguous. The intensities in the image are a complex and unknown combination of many distinct physical processes, including the lighting, material properties and object geometry. In order to recover the intrinsic material properties of the surface—and identify which sphere is chrome, and which one is plastic—the visual system must somehow disentangle these various contributions from one another.

One reason this is difficult is because the image of a given material can change dramatically depending on the context. For example, the image of the chrome sphere consists of nothing more than a distorted reflection of the world surrounding it. Therefore, when it is moved from one context to another, the retinal image changes dramatically (see Figure 3). This means that the visual system cannot recognize materials by simply matching the image against a stored 'template'. Somehow, the visual system has to abstract what is common to the appearance of the sphere across these different contexts.



FIGURE 3: Photographs of two materials in two different contexts. We easily identify the spheres across changes in the illumination, and easily distinguish between the two spheres.

To make matters worse, a chrome sphere could, in principle, be made to take on any arbitrary appearance, simply by placing it in a carefully contrived context, so that it reflects certain intensities into the eye. For example, by placing the chrome

sphere in a carefully designed ‘smooth’ world, it could be made to produce exactly the same pattern of pixels as one of the pearlescent spheres. Because the images would be identical, the visual system would have no way to tell the difference. However, we do not have to go to such extremes to encounter problems. In figure 3, the image of the mirrored and pearlescent spheres on the left (same illumination) are actually more similar to one another on a pixel-by-pixel basis than the two images of the chrome surface in different contexts (top row). This occurs because the positions of the highlights and dark regions are the same when the illumination is the same.

This is the fundamental ambiguity facing the visual system: identical materials can create very different images, and very different materials can create surprisingly similar images. Under arbitrary viewing conditions, the image would be completely ambiguous and the visual system would have no way of knowing which aspects of the image are due to the material, and which are due to lighting, geometry or other effects.

## TWO THEORETICAL APPROACHES TO MATERIAL PERCEPTION

How then can the visual system overcome this ambiguity and estimate material properties? Broadly speaking, two general approaches have been suggested. The first is *inverse optics*, which is the idea that the visual system explicitly estimates and ‘discounts’ the contributions of illumination and geometry to the observed intensity values (Marr, 1982; Poggio and Koch, 1985; Poggio, Torre and Koch, 1985; Pizlo, 2001). According to this line of reasoning, the visual system ‘runs physics in reverse’ to accurately model the physical properties of the scene, reconstructing the positions of light sources, the surface geometry and the physical reflectance parameters of the surface from the image. For example, von Helmholtz

(1867/1962) famously conjectured that the visual system recovers albedo by estimating and actively discounting the contribution of the illuminant to the observed image intensity. Similar reasoning plays a role many more recent theories of colour constancy (e.g. Maloney and Wandell, 1986; Brainard, Kraft and Longère, 2003; Yang and Maloney, 2003; Maloney, Boyaci and Doerschner, 2005). In order to estimate the reflectance of the spheres in the figure, the visual system would model the scene surrounding the spheres, estimate that the surface is spherical, and use this information to factor out the contributions of lighting and geometry to the image. What is 'left over' once these other factors are removed would be the intrinsic reflectance properties of the object.

The main advantage of such an approach is that the visual system would theoretically end up with a *physical* model of the scene, much like a scene description in computer graphics. The main disadvantage is that the visual system is faced with a 'chicken and egg' problem: in order to estimate and discount the lighting, the visual system would need to estimate and discount the *reflectance* – but this is exactly what the brain is trying to work out in the first place. To get around this problem, inverse optics models often invoke various kinds of *a priori* assumptions about the properties of the world. For example, it is common to assume that the illumination comes from a single distant point source, or that the surface reflectance is uniform and Lambertian (*i.e.*, completely matte). This makes the problem tractable, but limits the range of viewing conditions and physical properties that can be recovered from the image. More recent computational approaches have shown that it is possible to successfully separate BRDF and illumination in a Bayesian framework that uses more realistic assumptions about the world (e.g., Romeiro and Zickler, 2010a,b). However it remains unclear to what extent these results could be adapted to model human visual processing.

An alternative approach to inverse optics would be to identify image measurements that are diagnostic of material properties, but which remain roughly invariant across changes in the illumination. That is, if there are certain image features that reliably correlate with a given material across a range of viewing conditions, then the visual system could use these measurements to recognize the material. This way, rather than explicitly estimating and discounting the effects of the illumination on the image, the visual system would try to ‘ignore’ them, and rather than explicitly estimating physical reflectance parameters, the visual system would recognize materials by representing their typical appearance in the image. This approach—which we can call the *image statistics* approach—has gained considerable traction in recent years (e.g., Nishida and Shinya, 1998; Fleming et al, 2003; Fleming and Bühlhoff, 2005; Motoyoshi et al, 2007). The logic underlying such an approach is as follows.

When we posed the problem facing the visual system we argued that under arbitrary viewing conditions the image is ambiguous. However, in the natural world, viewing conditions are *not* completely arbitrary. In the real world, illumination conditions are shaped by the environment, leading to certain statistical regularities that are generally well conserved from scene to scene. These statistical regularities in the world mean that a given material tends to present certain statistical regularities in the image. For example, although the precise positions of highlights and shadows can vary radically from scene to scene, certain features of the reflections (such as the average contrast or blurriness of the highlights) generally remain more constant. Thus, a given material will tend to produce certain tell-tale statistical ‘signatures’ in the image, which the visual system could use to recognize different materials. Detecting these signatures potentially allows the visual system to identify materials without having to accurately estimate all other parameters of

the scene. This approach has the disadvantage that the visual system could be fooled when the assumed statistics of the world are infringed. However, it has the advantage of being able to handle arbitrary material properties: as long as a material exhibits distinctive image features, the visual system can learn these to recognize the material.

We tested this idea by presenting subjects with images of glossy spheres rendered under different illuminations (Fleming, *et al.*, 2003). Their task was to adjust the reflectance parameters of one sphere until it appeared to be made of the same material as the other sphere while ignoring any differences in illumination between the two spheres. This allowed us to test the extent of ‘gloss constancy’, that is, the constancy of perceived gloss across changes in illumination conditions. We found that when the spheres were rendered under illuminations that were photographically captured from the real world (Debevec, 1998)—as shown in Figure 4a—subjects were quite good at performing the matches. By contrast, when the spheres were illuminated under unnatural illuminations, such as the one shown in Figure 4b, performance decreased significantly. This suggests that the visual system relies on characteristic signatures of specular reflection. When the tacit assumptions about the statistical structure of the environment are infringed, gloss perception breaks down.

(a) natural environment



(b) unnatural noise environment

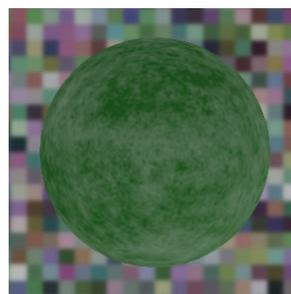


FIGURE 4. Computer simulations of two identical glossy spheres under (a) a real pattern of illumination, and (b) a pattern of random noise that has an unnatural intensity histogram. Observers reliably report that the sphere in (a) appears glossier than in (b) although the surfaces are identical.

#### AN ALTERNATIVE VIEW OF MATERIAL PERCEPTION AS 'STATISTICAL APPEARANCE MODELS'

One conclusion of our experiments on gloss constancy is that the visual system is far from perfect. We have 'partial' gloss constancy, as changing the illumination can also affect the perceived glossiness of surfaces. This in itself is not very surprising, given that numerous experiments have documented the limits of lightness and colour constancy (e.g., Land and McCann, 1971; Arend and Reeves, 1986; Bauml, 1993; Amano and Foster, 2004; Amano, Foster and Naschimento, 2006; Kraft and Brainard, 1999; Gerhard and Maloney, 2010; see Gilchrist et al, 1999 and Foster, 2011 for reviews). These findings show that perceived albedo and surface colour also vary to some extent with the lighting. Computationally it is also no great surprise. If the goal of the visual system is to estimate the physical properties of the surface, and these are confounded with the illumination in the image, then it makes sense that errors in separating the two sources could lead to mis-estimates of the surface reflectance properties.

However, more recent experiments have yielded some much less intuitive interactions between scene variable. For example, Ho, Landy and Maloney (2008) measured the perceived glossiness of surfaces with different reliefs. Their surfaces consisted of a conglomeration of ellipsoids, forming a smooth, bumpy surface, whose depth variations were varied across conditions, similar to the ones shown in

Figure 5. Subjects judged both the relief and glossiness of the surfaces. The somewhat unexpected result was that there was a significant ‘contamination’ between judgements of the two parameters. In other words, when asked to estimate the glossiness of the surface, the judgments varied significantly depending on the depth of the relief. This is a surprising finding: why should the visual system get confused between glossiness and surface relief?

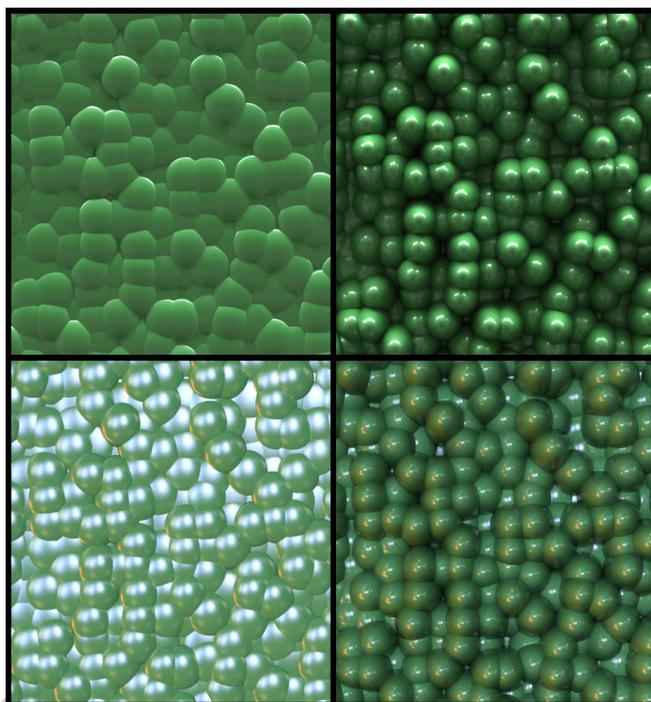


Figure 5: Images of glossy reliefs like those used by Marlow *et al.* (2012). All four surfaces have identical reflectance properties and yet the perceived glossiness varies depending on the interactions between surface relief (left: shallow *vs.* right: deep) and illumination direction (top: oblique *vs.* bottom: frontal). Image copyright 2012 Marlow, Kim and Anderson, reproduced with permission.

To gain further insights into these unexpected results, Marlow, Kim and Anderson (2012) recently extended the range of conditions tested. In addition to varying the relief and reflectance of the surface, they also varied the orientation of the illumination relative to the surface, so that it either arrived from head on, or from a more oblique angle. As can be seen in Figure 5, the way these different factors

interact has very significant effects on the highlights visible in the image. All four surface patches have identical surface reflectance properties, yet most observers perceive the patch with shallow relief and frontal illumination (bottom left) to appear significantly 'more glossy' than the others. In fact, Marlow and colleagues found that the interactions between the factors were not only large, they were in some cases non-monotonic, which suggests that the effect is not a simple 'contamination' of one quantity (perceived glossiness) by another (relief), as originally appeared to be the case in the Ho *et al.* study. For example, under oblique illumination, judgments of glossiness first increase and then decrease again as the depth of relief increases. What can account for these large and non-linear interactions between scene variables in the perception of gloss?

A key observation that casts light on this question is that the interplay between illumination and surface relief has substantial effects on how large and pronounced the surface's specular highlights appear in the image. Shallow relief illuminated frontally leads to large, high contrast highlights that dominate the image, whereas when the same surface is illuminated obliquely, the highlights appear much smaller and less pronounced by comparison.

This simple insight leads to a key hypothesis about gloss perception: when subjects are asked to report the apparent glossiness of a surface, it could be that their judgments reflect *the extent to which the surface manifests salient specular reflections*. Put another way, it could be that subjects use the characteristics of the reflections—their size, contrast, distinctness, *etc.*—as a 'proxy' for estimating the intrinsic physical surface parameters, such as specular reflectance. This makes intuitive sense as reflections and highlights are the defining visual characteristic of glossy surfaces. While surface reflectance properties are not visible directly, highlights and reflections—as the primary manifestations of specular reflection in the image—are

visible directly, and have properties that can be measured relatively easily by low- and mid-level visual hardware. Glossier surfaces manifest more salient specular reflections than less glossy ones, and thus it makes intuitive sense for the visual system to attach special import to the size, contrast and distinctness of specular highlights as a way of characterising surface gloss.

To test this hypothesis, Marlow and colleagues asked a different set of subjects to rate simple image properties of the highlights, such as their size, contrast and distinctness. These subjects were not asked to judge anything about the intrinsic properties of the surfaces themselves, they simply had to focus on the 2D image appearance of the highlight regions of the image. The results were quite striking. The authors found that a simple weighted combination of these latter judgments accounted for all the main trends in the glossiness judgments, including the non-monotonic effects of surface relief and the interactions between relief and illumination direction. That is, it is possible to predict gloss judgments just from the low-level image properties of the highlights. Thus, when asked to compare the glossiness of different images of surfaces, what subjects actually appear to do is to compare the relative salience of the highlights, based on their size, contrast, distinctness, and so on.

### **Statistical Appearance Models**

The idea that the visual system measures the properties of highlights instead of estimating the surface specularities perhaps sounds so obvious as to be almost trivial. However, it represents a subtle but profoundly significant shift in our understanding of what we mean by material perception. It implies that the goal of surface reflectance perception is *not* to estimate the BRDF or some parametric approximation of the intrinsic physical properties of the surface – whether through inverse optics or

image heuristics. Instead, the goal is to capture the typical ‘look’ of surfaces as they appear in the image, and to characterize how this appearance tends to vary from sample to sample. That is, the goal of material perception is to identify and measure *statistically informative appearance attributes*—like the size, contrast and distinctness of highlights—that capture how variations in material properties manifest themselves in the image<sup>1</sup>. Rather than estimating *physical* parameters of materials, the visual system somehow identifies key *image* parameters that vary between samples of related materials (e.g., surfaces with different degrees of gloss), and uses such measurements to represent the ‘typical appearance’ of glossy surfaces.

With this in mind, the central theoretical speculation of this article is the following:

**We suggest that the brain is highly adept at inferring a type of generative model of material appearance—a ‘*statistical appearance model*’—which captures the natural degrees of variation between samples in terms of easily measured appearance properties.**

We suggest that from even a relatively small number of samples of different materials, the visual system rapidly infers how to parse the image into appearance characteristics (like properties of the highlights) that vary parametrically between samples. Thus, such a generative model encapsulates the visual system’s

---

<sup>1</sup> The term ‘appearance’ has different uses in different fields. In psychology the term typically refers to the subjective phenomenological impression of surface characteristics. By contrast, in computer graphics—and to some extent computer vision—the term ‘appearance’ is routinely used to refer to how surface properties manifest themselves in the image; that is, certain physical aspects of the proximal stimulus associated with the material. In this context, we deliberately conflate the two meanings to make the point that the subjective ‘appearance’ of surfaces is intimately related to the goal of describing *regularities in the proximal stimulus* that they present to the visual system. We suggest that subjective ‘appearance’ is based on an internal model of ‘appearance’ (what surfaces look like in the image).

'knowledge' about the way samples typically behave, in terms of their changing appearance. Unlike a physical model, which represents materials in terms of pre-established intrinsic physical parameters, a statistical appearance model seeks to discover in what ways different material samples *look different* from one another, irrespective of the underlying physical basis for those differences (e.g., photon interference effects, microscopic roughness, and so on). Unlike simple image heuristics, which seek to approximate physical parameters by mapping crude statistics—such as the skewness of intensity or sub-band histograms—directly to surface properties, an internal model makes it possible for the visual system to predict what new samples might look like even from only a small number of exemplars. It is a generative model, which represents the dimensions along which samples tend to vary. It is also essentially a 'mid-level' theory of perception, in which perceptual organization principles—including geometrical constraints—play a central role in inferring and representing the characteristics of samples.

### **How might Statistical Appearance Models Be Computed?**

To give a concrete example of how such a model might be computed and what it represents, let us consider how the visual system could infer a statistical appearance model from images of glossy surfaces. Our goal here is not to provide a detailed computational model, but rather to adumbrate some of the key elements or processing stages that such a model might contain. We outline the approach in Figure 6.

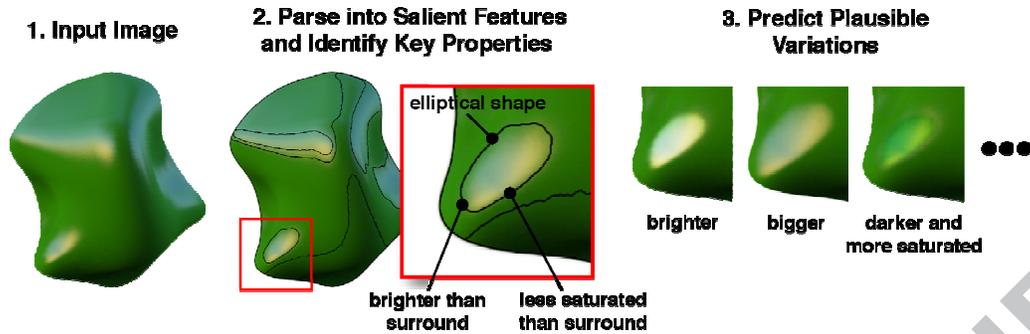


FIGURE 6. A cartoon schematic for inferring predicting plausible variations from a single exemplar image.

Suppose an observer is presented with a single image of a glossy surface. The visual system's goal is to work out in what ways the image would change if we were to change the properties of the surface. In theory, there is an infinite number of possible ways that an image could change, so how can the visual system work out what other samples, with different properties, might look like?

If the world were unconstrained, this would certainly be an impossible task. However, the real world is highly structured due to lawful generative processes, which have systematic effects on the image. For example, in the real world, moving a light source does not cause arbitrary changes in intensity independently at each image location. Instead, shading patterns undergo smooth and systematic transformations when light sources move. Thus, we suggest that the visual system can rely on the fact that variations in meaningful parameters of the world—such as the lighting or reflectance properties of the surface—generally lead to *systematic* changes in the image. The brain's goal is to characterize those changes. We suggest that starting with even a single exemplar of a material, the visual system may be able to cast initial hypotheses about which aspects of the image might be likely to change in which ways, and therefore develop an internal model of material appearance.

Specifically, we suggest that the visual system first uses general-purpose perceptual organization mechanisms to parse the image into salient regions or features, such as the highlights and shadows across the surface<sup>2</sup>. This segmentation provides some initial *candidate features* (e.g., highlights), which could plausibly vary in some measurable way with changes in the material properties. Again, we do not assume the visual system knows anything about the physical laws of reflection, or the properties of surfaces: it is trying to discover which aspects of the image change in lawful ways from material to material. The assumption is that in general, salient features are likely to relate in some systematic way to the underlying properties of the material. Put another way, salient features are likely to be *evidence of significant underlying causes* – and are therefore likely to vary from sample to sample. Of course, not all salient features turn out to vary in lawful ways, but they at least represent good initial candidates for building an internal model of material appearance.

Depending on the properties of the candidate features (as well as prior experience with other models) it should be possible to cast initial hypotheses about their likely degrees of variation. For example, as highlights differ from their surroundings in intensity, one plausible hypothesis is that the relative brightness of highlights might be an important feature: a feature that varies systematically from material to material. Along similar lines, highlight regions may also differ from sample to sample in terms of their size, shape, position or other measurable properties.

We suggest that this set of hypotheses about possible ways in which appearance features could vary from sample to sample represents an initial

---

<sup>2</sup> The segmentation processes may also involve ‘scission’ mechanisms (i.e., source separation), which parse the image into ‘causal layers’, as in transparency (cf. Anderson, 1997). Although we do not understand exactly how such mechanisms work, there is extensive empirical evidence that the visual system is adept at distinguishing multiple superimposed image contributions.

appearance model, which is then refined and corrected through experience with other samples. When new samples are encountered, the visual system can track how candidate features vary to improve its appearance model of glossy surfaces. For example, confirmation that different samples do have different highlight contrasts reinforces this element of the appearance model. In contrast, observing that highlight colour tends not to change very much reinforces that it is not a natural degree of image variation associated with glossy surfaces. The more evidence is provided, the more refined the appearance model becomes, but it remains, fundamentally a model of how image features tend to change, rather than an estimate of physical surface parameters. The key to the approach is that perceptual organization principles provide initial constraints on an otherwise limitless space of possible variations, and that accumulated experience with different materials allows the visual system to discover reliable dimensions along which samples tend to vary.

This is, of course, a highly speculative proposal, and raises many questions about how the putative segmentation processes work, how the visual system learns which hypotheses to cast in the first place, and how new evidence is incorporated into the model as the observer gains experience with different surfaces. It is important to point out that segmenting the image into diffuse and specular reflections, for example, is far from trivial, and a considerable amount of research still needs to be done to understand how such perceptual organization mechanisms work. It is likely that the segmentation processes, and resulting appearance model take into account multiple photo-geometric constraints, such as the consistency in image orientation between shading gradients and highlights (Beck and Prazdny, 1981; Fleming et al, 2004; Todd, Norman and Mingolla, 2004; Marlow, Kim and Anderson, 2011). It remains unclear at what level these constraints are imposed. Raw filter responses and intensity statistics are presumably not sufficient to measure

the relevant relationships between features, so additional grouping processes must be involved. At the same time, it may not be necessary for the visual system to determine consistency in world coordinates, using explicit estimates of unseen elements (e.g., light sources, rays, etc). For example, the visual system may not necessarily enforce consistency between shading and highlights in terms of consistency between estimates of the 3D surface structure and estimates of the light sources using ray geometry to enforce consistency. Instead, it may detect consistency in terms of image-level features, such as the directions of intensity gradients that are attributed to different causes. Thus, ‘mid-level’ visual processes (Adelson, 1999) may be sufficient to express the crucial photo-geometric constraints. By posing gloss perception this way, we believe we can make progress in understanding some of the otherwise confusing effects and interactions between different scene variables.

#### THEORETICAL BENEFITS OF STATISTICAL APPEARANCE MODELS

If the visual system cannot or does not estimate intrinsic physical parameters in the case of surface reflectance perception, there are scant grounds for thinking that it estimates intrinsic physical parameters for other material properties. Estimating complex real-world properties, like translucency, elasticity, viscosity or sponginess—whether through inverse physics or image heuristics—is surely at least as difficult as surface reflectance estimation. We therefore suggest that the general strategy of representing appearance differences likely applies to the perception of all kinds of material properties. Indeed, one of the major computational benefits of inferring an appearance model—as opposed to estimating physical parameters—is that it is highly flexible: it is unconstrained by pre-defined parameters of the physical model,

and can readily adapt to new materials (and new properties) that have never been seen before. Not only do we encounter new materials throughout our childhood, from time to time material science also creates new materials with completely novel appearances, such as complex paints and textiles with unusual colour characteristics, or holograms, which have a highly distinctive 'look', quite unlike most natural materials. A brain that represents material properties in terms of appearance features can learn these new materials by identifying statistical regularities in the patterns of sensory activity they evoke.

### **Representing arbitrarily complex material properties**

More importantly, by focussing on appearance, the visual system can capture arbitrarily complex physical processes, as long as they lead to systematic variations in features that can be easily measured. For example, consider the cylinders presented in Figure 7, from a recent computer graphics article (Narain, Pfaff and O'Brien, 2013), in which the authors model the crumpling behaviour of thin elastic and plastic sheets. The model has a number of physical parameters, which control the behaviour of the material in response to external forces (here, a compression of the cylinders). In the figure, one of the model parameters varies from left to right leading to a vivid subjective impression of differences in the properties of the material.



FIGURE 7: A series of simulated crumpled materials reprinted from Narain *et al.*, (2013). From left to right, one parameter of the physical model varies, leading to different crumpling behaviour and a concomitant change in appearance. Image copyright 2013, Narain, Pfaff, and O'Brien, reproduced with permission.

The cylinder on the left appears to be a thin, papery material, as seen on a Chinese lantern, whereas the one on the far right appears to be thicker and more rubbery material, which buckles elastically under pressure. The underlying physical processes are highly complex, and it seems quite implausible that the visual system has a sophisticated internal model that captures these physical processes and estimates the parameters of the model. However, despite the complexity, at the phenomenological level, the emergent effects of the crumpling process are relatively straightforward, and leave clearly identifiable signatures in the shape of the object. The crumples in the 'papery' cylinder are smaller, higher frequency and have sharper creases than the smooth, large scale undulations of the 'rubbery' cylinder. It seems much more plausible that the visual system identifies the key statistical differences between the materials, expressed in terms of these mid-level appearance characteristics, rather than estimating or approximating intrinsic parameters.

**Statistical generative models as categorization in high-dimensional feature spaces**

Another important advantage of statistical generative models is that they embody knowledge about the *relationships* between material samples, and are thus more expressive than simple heuristics. A ‘Bag of Tricks’ (Ramachandran, 1985) view of material perception—based on simple correlations between image features and physical surface parameters—does not capture a deeper understanding of the underlying generative processes that are responsible for the correlations. In their simplest form, heuristics represent a case-by-case mapping from sensory measurements to physical properties. In contrast, by capturing the typical behaviour of materials with an internal model, the visual system can also predict, to some extent, what plausible variations of exemplars might look like, and thus relate samples to one another in meaningful ways. An internal model adds ‘understanding’ to the heuristics by predicting the sensory consequences of changing samples and viewing conditions, much as internal models of limb movements are thought to predict the sensory consequences of actions (Wolpert, Miall and Kawato, 1998).

Indeed, by formulating material perception as the process of discovering statistical relationships between samples, we can to some extent unify *estimation* and *categorization* into a common theoretical framework, as depicted in Figure 8. We can think of individual samples of materials as points in a high-dimensional feature space (Figure 8a), where the features represent appearance characteristics, like the smoothness or extendedness of the liquid. Different materials which differ, for example, in terms of their viscosity, occur at different locations within the feature space, tracing out a sub-space or manifold within the space of all possible appearances. Material *estimation* is the process of establishing the true position of a given sample within the feature space (Figure 8b), and material *categorization* is the process of identifying the boundaries separating different classes of material. A

statistical appearance model facilitates both of these processes because it represents a hypothesis about the shape and internal parameterization of the sub-space occupied by related materials, and thus determines which appearance features are important for a given class of material. In other words, learning to estimate viscosity can be thought of as the process of working out how to parameterize the sub-space occupied by viscous materials, for example, through non-linear dimensionality reduction (Tenenbaum, de Silva & Langford, 2000; Roweis & Lawrence, 2000).

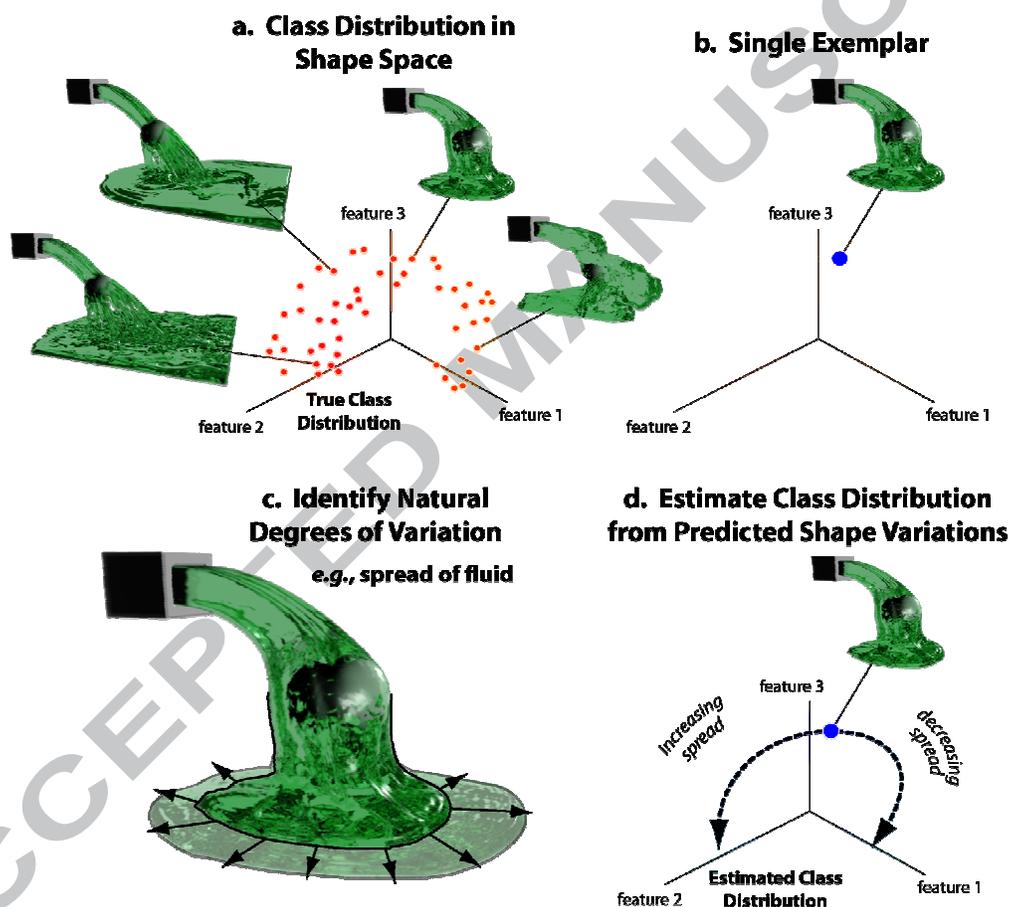


FIGURE 8: Relationship between statistical appearance models and material categorization, based on samples of viscous liquids.

A generative model that predicts natural degrees of variation between samples (Figure 8c) is equivalent to casting a hypothesis about the distribution of samples in the feature space (Figure 8d). As we experience more material samples, the accuracy

of the model improves, providing more accurate category boundaries as well as a more accurate representation of the natural dimensions of variation between samples within a class. We have speculated that the visual system derives initial hypotheses about the likely dimensions of variation between samples by parsing images into candidate features. This would potentially allow the visual system to infer material classes, and identify key dimensions of variation based on just a small number of exemplars. One of the most remarkable aspects of material perception is that when presented with just a small number of samples of related stimuli, we seem to be able to rapidly identify which features we should use for comparing them. One very interesting direction for future research is to test how subjects learn to recognize and distinguish novel materials (e.g., using simulations of BRDFs that are quite unlike those seen in natural materials) from small numbers of exemplars. If our hypothesis is correct, subjects should be able to identify key dimensions of variations from just a few samples, allowing them to predict the appearance of intermediate materials, for example.

This view of material perception is inspired by ideas from machine learning, which pose learning as a process of inferring underlying generative processes from data samples (e.g., Kemp and Tenenbaum, 2008; Schmidt and Lipson, 2009; Tenenbaum et al, 2011). Given only limited data, the expressiveness of the model is limited, and distinct physical processes may be conflated. For example, when a runny liquid is poured, the height of the splash is influenced by both the viscosity of the liquid and the height over which the liquid is poured. However, given only limited experience, the visual system may be unable to distinguish between these two factors and therefore conflate them in its representation of the liquid. This would show up as apparent errors in ‘estimates of viscosity’, with the visual system

unable to separate the effects of height from the effects of the intrinsic properties of the material.

Of course, a sufficiently detailed appearance model (i.e., one inferred from sufficient samples) should be able to separate distinct underlying causes as long as these create systematic and distinct effects on measured image features. By representing samples in a high-dimensional (over-complete) feature space, the visual system may be able to tease apart factors that have different physical origins. Statistical appearance models see this separation as a process of *discovering distinct dimensions of variation* between observed samples distributed in the feature space—rather than the application of pre-defined strategies to estimate specific physical quantities.

For example, when judging surface lightness, both reflectance and illuminance contribute to the observed image intensity of a surface patch. Given only isolated luminance measurements, there is no way in principle to separate the two factors to correctly infer surface albedo. Therefore, an appearance model based solely on luminance would confound illuminance and reflectance changes: increasing either factor would make the surface appear brighter. In traditional parlance, this would be a ‘failure of lightness constancy’, although of course, brightly illuminated surfaces *do* tend to appear subjectively brighter (although not lighter), presumably reflecting the fact the image intensity *is* an important dimension of variation between samples and thus serves as a useful low-order characterisation of appearance, even though it is not specifically diagnostic of reflectance. However, Gilchrist and Jacobson (1984) have shown that observers can distinguish a dark gray room under bright light from a light gray room under dim light, even when the average luminance is the same for the two rooms. This shows that the visual system does not rely *solely* on raw luminance values to represent the difference between different samples. In other

words, the appearance model for illuminated matte surfaces is not unidimensional. The input relies on more features (dimensions) than just luminance, and the resulting appearance model captures more natural degrees of variation between samples than just 'brightness'. The idea that lightness perception probably involves more than one subjective dimension has been discussed widely (Katz, 1935; Arend & Reeves, 1986; Whittle, 1992; Logvinenko and Maloney, 2006; Anderson and Winawer, 2008; Shapiro, 2008; Vladusich, 2012). According to the statistical appearance model idea, this reflects the fact that observers have internalized the statistical 'look' of inter-reflections, using higher order image features (contrasts, filter responses, etc.) in conjunction with luminance. It is these higher-order degrees of variation between samples that allow observers to distinguish the distinct contributions to observed luminance from reflectance and illuminance.

### **Predicting systematic estimation errors and other eccentricities of material perception**

Importantly, statistical appearance models also predict several perceptual phenomena that would otherwise be difficult to understand if the visual system's goal were to estimate physical properties of surfaces. There are many cases in which scene parameters have non-intuitive effects on material perception. We have already suggested that appearance-based explanations may account for some of these, such as the complex non-monotonic interactions between illumination, relief and glossiness in the studies by Ho *et al.* (2008) and Marlow *et al.* (2012).

They probably also account for the effects of other scene variables that we have observed in the perception of transparent materials (Fleming, Jäkel and Maloney, 2011). When a thick transparent object is placed in front of a patterned background, the patterns that are visible through the object appear spatially distorted due to

refraction. The degree of distortion depends on the *refractive index*, a physical parameter of the material that determines how light is ‘bent’ as it passes through the object. This means that the visual system could use estimates of the degree of distortion to infer the material properties of the object. However, the degree of distortion also varies with other scene variables, including the 3D thickness of the refracting object, as well as its distance from the background that is visible through it.

We asked subjects to adjust the refractive index of one simulated object to match the apparent material properties of another transparent object, which had a different thickness or distance to the background. We found that subjects’ matches were substantially biased by the thickness or distance of the object, even though these scene variables have nothing to do with the intrinsic properties of the material. However, these biases can be easily understood if instead of estimating the physical refractive index of the transparent objects, subjects simply match the degree of distortion observable in the image. Because the salient consequence of refraction is spatial distortion of the background, subjects use this distortion to capture the appearance of refractive materials.

It is important to note that observing ‘perceptual errors’ in material perception experiments does not on its own rule out the possibility that the visual system *is* trying to estimate physical surface parameters and failing (or only imperfectly succeeding). Indeed, to equate the estimation of physical properties with perfect performance would be a ‘straw man’ as it is well known that biological vision is far from perfect. The point is rather to account for the specific pattern of errors. It is difficult to address experimentally the teleological question of the true nature of the visual system’s ‘goal’. In the limit, a highly detailed and accurate appearance model may yield patterns of results that approximate what a full inverse optics

computation would achieve. Nevertheless, I suggest that where errors are large and systematic (i.e., not just noise), posing material perception as a process of representing appearance changes rather than the estimation of internal parameters provides a useful way to understand patterns of successes and failures. Perceptual theories based on analytical representations of the inverse optics problem do not readily predict which image information the visual system relies on, or why in some cases the visual system fails to compensate for the spurious affects of different scene factors on image measurements. By contrast 'statistical appearance models', or other data-driven approaches are directly connected to the image measurements. Thus, appearance models may offer a way to plug the explanatory gap between simple image measurements and higher-level goals, like representing the structure of the world in sufficient detail to support successful interactions.

The idea that the visual system does not care about estimating physical scene parameters may also explain the curious fact that hue is a circular dimension. It is widely assumed that the goal of early colour processing is to infer a low dimensional estimate of the distribution of wavelengths in the stimulus from cone excitations (Wandell, 1995). In spectral terms, narrow-band stimuli that evoke red and purple colour sensations lie at opposite ends of the visible spectrum. Despite this, in terms of subjective appearance, red and purple lie next to one another on the hue circle, which wraps the two ends of the spectrum close to one another. Purple appears more similar to red than green does, even though, in physical (spectral) terms green is closer to red than purple is.

If the goal of colour perception were to estimate the physical spectra of sources or surfaces, and to accurately represent the physical similarities or differences between stimuli, then this circularity makes no sense: why should hue be a circular dimension when wavelength is a linear dimension? In the case of pitch perception,

the physical properties of oscillating sound sources tend to create harmonic relationships between frequencies. This may explain why tones that are separated by an octave, for example, appear subjectively similar to one another: raising a harmonic signal by an octave leads to a physically similar signal (minus the fundamental). But in colour, there is no equivalence of harmonic relationships. Wavelengths are not coupled in the same way because colours are not created by standing wave oscillators like many sound sources are, and thus there is no ‘circularity’ in spectral relationships. One possible explanation is that circularity provides a way to represent similarity relationships between more complex spectra: for example placing spectra that are a combination of long and short wavelengths subjectively in between spectra that contain only short or long wavelengths. However, this could also be achieved using some Cartesian (rather than polar), organization of colour space. Another possibility is that the visual system does not care about representing the *physical* similarity between stimuli with different spectra. Instead, there may be some other organizational principles that benefit from representing hue on a circle. Future research—perhaps considering the effects of ripening, changes in daylight, or internal constraints, like the computation of iso-hue flow patterns—may provide an explanation of why hue is circular.

In summary, statistical appearance models are easier to compute than physical models, and more expressive than simple heuristics. Such models capture not only the key characteristics of individual samples, but also the relationships between samples, including novel samples that have not been seen before, by representing a material’s natural dimensions of variation. Appearance-based (rather than physics-based) explanations of material perception account for the otherwise baffling effects of irrelevant scene variable on material perception. Taken together, this suggests

that statistical appearance models represent a powerful and flexible way of thinking about how the visual system represents material properties.

### **Beyond Material Perception**

In this article, we have suggested that when we look at an object and experience a vivid subjective impression of its material properties, we are not actually perceiving its *physical* properties at all. Instead, we have learned a set of *appearance characteristics*—*i.e.*, properties of the way the material tends to appear in the image—that capture its distinctive ‘look’. The frothy head of the freshly poured wheat beer doesn’t look like surface tension and sub-surface scatter in action. It looks like a certain kind of whitish, softish, stuff that is different in important sensory ways from the whitish, softish stuff on the inside of a banana skin.

What consequences does this view of material perception have for the rest of perception? I suggest that the visual characterisation of material appearance is likely to be just a special case of a much more general perceptual and cognitive faculty for inferring statistical regularities related to the high-level attributes of things, scenes and events. The Gestalt psychologists referred to the ‘tertiary’ qualities of sensory experiences, such as the ‘mood’ of a room or the gracefulness of a ballet dancer. Much of the aesthetic pleasure of sensory experience seems to reside at this level of experience: the visual pleasure of seeing a curtain buoyed by a breeze, the poise of a well-crafted sculpture, or the melancholy air of favourite melody. How does the visual system represent these aspects of the world?

Presumably there are *some* physical regularities that underlie such experiences, as we can judge them consistently and dispute their relative merits with other observers. Yet, at the same time, it seems very difficult—perhaps even impossible—to describe what we are responding to in purely reductive physical terms. When we

hear someone speaking through a wall and can identify a familiar accent even though we do not understand the words, or when we note similarities in handwriting style between a father and son, it seems highly improbable that we are inferring properties of a physical generative model. What would this even mean? It would be almost comical to praise the low-level physical characteristics of a dancer's movements such as the ratio of tensions in particular muscle groups. Surely these experiences—the differences between good and bad dancers, *etc.*—are expressed along some other kinds of appearance dimensions. When we experience tertiary properties of objects, scenes and events, there is no concept of a physical model. Why, then, should the properties of materials be different?

## CONCLUSIONS

The subjective visual experience of materials and their properties is vivid and nuanced but poorly understood. Research over the last decade or so has started to make progress in this important area but each new finding seems to raise many new questions. Different senses make fundamentally different kinds of measurements about materials, so how are these different quantities compared and combined to yield a multi-sensory impression of material properties? To what extent and in what ways does semantic knowledge about materials influence perceptual processing? What limits our ability to generalize perceptual knowledge about specific materials—or material classes—to novel viewing situations? Future research must address questions such as these.

One theoretical idea that has been gaining traction is that the visual system may rely on a heuristic approach based on various image statistics that correlate with

material properties. Such an approach is appealing because it would not require the brain to perform sophisticated computations to arrive at estimates of material properties. However, as we have argued here, this approach is not without its problems. Given that there are often many possible image measurements that correlate to a greater or lesser extent with any given material property, there is a risk that the field will become satisfied simply to collect such correlations without seeking a deeper theoretical understanding of the origin of these cues. It is important that we test the ability of each hypothesized cue to predict not only the successes but also the failures of material perception. Methods must be developed for perturbing the putative image properties and measuring the consequences for perception, to establish their causal role in each judgment. Theory must be developed to model the processes through which the visual system selects which cues to use for any given material perception task. Ultimately, we must not only be able to answer the question “which cues does the visual system use?” but also the question “how does the visual system end up using this cue, rather than some other?” As this area of study matures, we must not allow material perception research to slip into a theory-blind process of collecting large numbers of weak correlations. This surely would not count as a deep understanding of how the brain works out how stuff looks.

## REFERENCES

- Adelson, E.H. (1999). Lightness perception and lightness illusions. In *The New Cognitive Neurosciences*, 2nd ed., M. Gazzaniga, ed. (Cambridge, MA: MIT Press), pp. 339-351.
- Adelson E. H. (2001). On seeing stuff: The perception of materials by humans and machines. In B. E. Rogowitz & T. N. Pappas (Eds.), *Proceedings SPIE Human Vision and Electronic Imaging VI*, (Vol. 4299, pp. 1-12).
- Adelson, E. H., & Anandan, P. (1990). Ordinal characteristics of transparency. *Proceedings of the AAAI-90 Workshop on Qualitative Vision*, 77- 81.
- Amano, K., & Foster, D.H. (2004). Colour constancy under simultaneous changes in surface position and illuminant. *Proceedings of the Royal Society of London B*, 271, 2319-2326.
- Amano, K., Foster, D.H., & Nascimento, S.M.C. (2006). Color constancy in natural scenes with and without an explicit illuminant cue. *Visual Neuroscience*, 23 (3-4), 351-356.
- Anderson B.L. (1997). A theory of illusory lightness and transparency in monocular and binocular images: the role of contour junctions. *Perception*. 26(4): 419-453.
- Anderson, B.L. (2003). The Role of Occlusion in the Perception of Depth, Lightness, and Opacity. *Psychological Review*, 110(4): 785-801.
- Anderson, B.L. (2011) Visual perception of materials and surfaces. *Current Biology*, 21 (24), R978-983.
- Anderson, B.L., & Kim, J. (2009) Image statistics do not explain the perception of gloss and lightness. *Journal of Vision*, 9(11):10, 1-17.
- Anderson BL, Winawer J. (2008). Layered Image Representations and the Computation of Surface Lightness. *Journal of Vision*. 8(7):18 1-22.

- Arend, L., & Reeves, A. (1986). Simultaneous color constancy. *Journal of the Optical Society of America A - Optics Image Science and Vision*, 3, 1743-1751.
- Ashikhmin, M. & P. Shirley (2000). An Anisotropic Phong BRDF Model, *Journal of Graphics Tools*.
- Bauml KH. Simultaneous color constancy: how surface color perception varies with the illuminant. *Vision Res.* 1999;39:1531-1550.
- Beck J, & Ivry R. (1988). On the role of figural organization in perceptual transparency. *Percept Psychophys.* 44(6): 585-594.
- Beck, J., Prazdny, S. (1981) Highlights and the perception of glossiness. *Perc. & Psychophysics*, 30(4), 407-410.
- Beck J, Prazdny K, Ivry R. (1984). The perception of transparency with achromatic colors. *Percept Psychophys.* 35(5): 407-422.
- Bergmann Tiest, WM and AM.L. Kappers (2007). Haptic and visual perception of roughness, *Acta Psychologica*, **124**(2): 177-189
- Berzhanskaya, J., Swaminathan, G., Beck, J., and Mingolla, E. (2005). Remote effects of highlights on gloss perception. *Perception* 34, 565-575.
- Blake, A., & Bülthoff H. H. (1990) Does the brain know the physics of specular reflection? *Nature* 343(6254) 165-168.
- Brainard, D. H., Kraft, J. M., and Longère, P. (2003). Color constancy: developing empirical tests of computational models. In *Colour Perception: Mind and the Physical World*, R. Mausfeld and D. Heyer (eds.), Oxford University Press, 307-334.
- Debevec, P. E. (1998). Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. *Proceedings of SIGGRAPH*, 1998, 189-198.

- Doerschner, K. Fleming, R.W., Yilmaz, O., Schrater, P.R., Hartung, B. and Kersten, D. (2011). Visual Motion and the Perception of Surface Material. *Current Biology*, 21(23): 1-7.
- D'Zmura, M., Colantoni, P., Knoblauch, K. & Laget, B. (1997). Color transparency. *Perception* 26, 471-492.
- Fleming, R. W. and Bühlhoff, H. H. (2005). Low-level image cues in the perception of translucent materials. *ACM Transactions on Applied Perception*, 2(3), 346-382.
- Fleming, R.W., Dror, R.O., and Adelson, E.H. (2003). Real-world illumination and the perception of surface reflectance properties. *J. Vis.* 3, 347-368.
- Fleming, R.W., Jäkel, F. and L. T. Maloney (2011). Visual Perception of Thick Transparent Materials. *Psychological Science*, 22(6): 812-820.
- Fleming, R. W., Torralba, A. and Adelson, E. H. (2004). Specular reflections and the perception of shape. *Journal of Vision*, 4(9), 798-820.
- Fleming, R.W., Wiebel C. and K Gegenfurtner (2013). Perceptual Qualities and Material Classes. *Journal of Vision*, 13(8): 9; doi:10.1167/13.8.9.
- Foster, D. H. (2011). Color constancy. *Vision Research*, 51, 674-700.
- Gerbino, W. (1994). Achromatic transparency. In A.L. Gilchrist, Ed. *Lightness, Brightness and Transparency*. (pp 215-255). Hove, England: Lawrence Erlbaum.
- Gilchrist, A. L., & Jacobsen, A. (1984). Perception of lightness and illumination in a world of one reflectance. *Perception*, 13, 5-19.
- Gilchrist A, Kossyfidis C, Bonato F, Agostini T, Cataliotti J, Li X, Spehar B, Annan V, Economou E. An anchoring theory of lightness perception. *Psychol. Rev.* 1999;106:795-834.

- Greene, M.R., & Oliva, A. (2009). Recognition of Natural Scenes from Global Properties: Seeing the Forest Without Representing the Trees. *Cognitive Psychology*, 58(2), 137-179.
- Günther, J., Chen, T., Goesele, M., Wald, I., and H-P Seidel (2005). Efficient Acquisition and Realistic Rendering of Car Paint. In *Proceedings of 10th International Fall Workshop on Vision, Modeling, and Visualization 2005 (VMV'05)*, G Greiner, J. Hornegger, H. Niemann and M Stamminger (Eds.) Erlangen, Germany, 487-494.
- Hays J. and AA. Efros. (2007) Scene Completion Using Millions of Photographs. *ACM Transactions on Graphics (SIGGRAPH 2007)*. August 2007, vol. 26, No. 3.
- von Helmholtz, H. (1867/1962). *Helmholtz's Treatise on Physiological Optics*. New York: Dover (edited by J. P. C. Southall).
- Ho, Y.-X., Landy, M. S. & Maloney, L. T. (2006), How illuminant direction affects perceived visual roughness, *Journal of Vision*, 6, 634-648
- Ho, Y.-X., Landy, M. S. & Maloney, L. T. (2008), Conjoint measurement of gloss and surface texture. *Psychological Science*, 19(2), 196-204.
- Hurlbert, AC, Cumming, BG and A.J. Parker (1991). Recognition and perceptual use of specular reflections. *Investigative Ophthalmology and Visual Science Supplement*, Vol. 32, No. 4.
- Katz, D. 1935. *The World of Colour*. Kegan Paul, London, UK.
- Kemp, C. and Tenenbaum, J. B. (2008). The discovery of structural form. *Proceedings of the National Academy of Sciences*. 105(31), 10687-10692.
- Kim, J., & Anderson, B. L. (2010). Image statistics and the perception of surface gloss and lightness. *Journal of Vision*, 10(9):3, 1-17.
- Kim, J., Marlow, P.J., & Anderson, B.L. (2011). The perception of gloss depends on highlight congruence with surface shading. *Journal of Vision*, 11(9):4, 1-19.

- Koenderink, J. J., & van Doorn, A. J. (1980). Photometric invariants related to solid shape. *Optica Acta*, 27(7), 981–996.
- Kraft, J.M. and D.H. Brainard. (1999). Mechanisms of color constancy under nearly natural viewing. *Proceedings of the National Academy of Sciences USA*. 96, 307-312
- Land, EH and JJ McCann (1971). Lightness and retinex theory. *J. Opt. Soc. Am.* 61:1-11.
- Logvinenko, A. D. & Maloney, L. T. (2006), The proximity structure of achromatic surface colors and the impossibility of asymmetric lightness matching. *Perception & Psychophysics*, 68. 76-83
- Liu, C., Sharan, L., Adelson, E. H. and R. Rosenholtz (2010). Exploring features in a Bayesian framework for material recognition. In CVPR, pages 239–246. IEEE, 2010.
- Maloney, LT and BA Wandell (1986). Color constancy: a method for recovering surface spectral reflectance. *JOSA A*, vol. 3 Issue 1 Page 29.
- Maloney, L. T., Boyaci, H. & Doerschner, K. (2005), Surface color perception as an inverse problem in biological vision. *Proceedings of the SPIE - IS & T Electronic Imaging*, 5674, 15-26.
- Marlow, P., Kim, J., and Anderson, B. L. (2011). The role of brightness and orientation congruence in the perception of surface gloss. *J. Vis.* 11, 1-12.
- Marlow, P.J., Kim, J., & Anderson, B.L. (2012) The perception and misperception of specular reflectance. *Current Biology*, 22, 1909-1913.
- Marr, D. (1982). *Vision*. San Francisco: Freeman.
- Matusik, W., Pfister, H., Brand, M., and L. McMillan (2003). A data-driven reflectance model. In ACM SIGGRAPH 2003 Papers (SIGGRAPH '03). ACM, New York, NY, USA, 759-769. DOI=10.1145/1201775.882343

- Metelli, F. (1970). An algebraic development of the theory of perceptual transparency. *Ergonomics*, 13(1): 59-66.
- Metelli, F. (1974a). Achromatic color conditions in the perception of transparency. In Eds. Macleod, R. B., & Pick, H. L. (Eds.) *Perception* (pp. 95–116). Ithaca, NY: Cornell University Press.
- Metelli F. (1974b). The perception of transparency. *Scientific American*, 230(4): 90-98.
- Motoyoshi, I. (2010). Highlight-shading relationship as a cue for the perception of translucent and transparent materials. *Journal of Vision* 10(9): 6, 1-11.
- Motoyoshi, I. and H. Matoba (2012). Variability in constancy of the perceived surface reflectance across different illumination statistics. *Vision Research* 53: 30-39
- Motoyoshi, I., Nishida, S., Sharan, L., and Adelson, E.H. (2007). Image statistics and the perception of surface qualities. *Nature* 447, 206-209.
- Muryy, A., Welchman, A.E. and Blake, A and R.W. Fleming (2013). Specular reflections and the estimation of shape from binocular disparity. *Proceedings of the National Academy of Sciences*, 110(6): 2413-2418.
- Narain R, Pfaff T, and JF O'Brien (2013). Folding and Crumpling Adaptive Sheets. *ACM Transactions on Graphics*, 32(4), 1-8.
- Nicodemus, F. (1965). "Directional reflectance and emissivity of an opaque surface" (abstract). *Applied Optics* 4(7): 767-775.
- Nicodemus, F. E., Richmond, J. C., Hsia, J. J., Ginsberg, I. W., & Limperis, T. (1977). Geometrical considerations and nomenclature for reflectance. *National Bureau of Standards Monograph*, 160.
- Nishida, S., and Shinya, M. (1998). Use of image-based information in judgments of surface-reflectance properties. *Journal of the Optical Society of America A*, 15, 2951-2965.

- Oliva, A., & Torralba, A. (2001). Modeling the Shape of the Scene: a Holistic Representation of the Spatial Envelope. *International Journal in Computer Vision*, 42, 145-175.
- Oliva, A. & Torralba, A. (2006). Building the Gist of a Scene: The Role of Global Image Features in Recognition. *Progress in Brain Research: Visual perception*, 155, 23-36.
- Olkkonen, M. & Brainard, D. H. (2010). Perceived glossiness and lightness under real-world illumination. *Journal of Vision*, 10(9:5), doi 10.1167/10.9.5.
- Olkkonen, 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, doi:10.1068/i0480.
- Oren, M. and S.K. Nayar (1994). Generalization of Lambert's Reflectance Model. In *Proceedings of the 21st annual conference on computer graphics and interactive techniques (SIGGRAPH '94)*, pp. 239-246.
- Padilla, S, Drbohlav, O, Green, P.R., Spence, A.D. and M.J. Chantler (2008). Perceived roughness of  $1/f^{\beta}$  noise surfaces. *Vision Research*, 48, 2008, 1791-1797.
- Pizlo, Z (2001), "Perception viewed as an inverse problem", *Vision Research* 41 (24): 3145-3161
- Poggio, T and C. Koch (1985). Ill-Posed Problems in Early Vision: From Computational Theory to Analogue Networks. *Proceedings of the Royal Society of London. Series B, Biological Sciences*: 226(1244), pp. 303-323.
- Poggio, T., V. Torre, and C. Koch. (1985). Computational Vision and Regularization Theory. *Nature*, Vol. 317, No. 6035, pp. 314-319.
- S.C. Pont, J.J. Koenderink (2005). Bidirectional Texture Contrast Function. *International Journal of Computer Vision*, 62(1/2), April/May 2005, special issue on Texture Synthesis and Analysis.

S.C. Pont, J.J. Koenderink (2008). Shape, surface roughness, and human perception.

In: *Handbook of Texture Analysis* (Majid Mirmehdi, Xianghua Xie, Jasjit Suri, Eds.), Imperial College Press, 197-222.

Ramachandran VS, (1985) The neurobiology of perception. *Perception* **14**:97-103

Robilotto R, Khang BG & Zaidi Q. (2002). Sensory and physical determinants of perceived achromatic transparency. *Journal of Vision*. 2(5): 388-403.

Romeiro, F. and T. Zickler (2010a), "Inferring Reflectance under Real-world Illumination", Technical Report TR-10-10, Harvard School of Engineering and Applied Sciences, 2010.

Romeiro, F and T. Zickler (2010b), "Blind Reflectometry." In K Daniilidis, P Maragos and N Paragios (Eds): *Computer Vision – ECCV 2010 (11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5-11, 2010, Proceedings, Part I)*. Springer Berlin Heidelberg. Lecture Notes in Computer Science Volume 6311, pp 45-58.

Roweis, S and S Lawrence (2000). Nonlinear dimensionality reduction by locally linear embedding. *Science*, **290**(5500): 2323-2326.

Rust NC, Stocker AA (2010) Ambiguity and invariance: two fundamental challenges for visual processing. *Current Opinion in Neurobiology* 20:382-388.

Schmidt, M and H Lipson (2009). Distilling Free-Form Natural Laws from Experimental Data. *Science* 324, 81; DOI: 10.1126/science.1165893.

Shapiro, A. G. (2008). "Separating color from color contrast." *Journal of Vision*, Vol. 8, No. 1:8, pp. 1-18. doi:10.1167/8.1.8.

Sharan, L, Rosenholtz, R. and EH Adelson (2008). Eye movements for material perception. *Journal of Vision*, **8**(6), 219a.

- Sharan, L., Rosenholtz, R. and Adelson, E. H. (2009). Material perception: What can you see in a brief glance? *Journal of Vision*, **9**(8):784, <http://www.journalofvision.org/content/9/8/784>, doi:10.1167/9.8.784.
- Singh M, Anderson BL. (2002a). Toward a perceptual theory of transparency. *Psychol Rev.* 109(3): 492-519.
- Singh M, Anderson BL. (2002b). Perceptual assignment of opacity to translucent surfaces: the role of image blur. *Perception*. 31(5): 531-52.
- Tenenbaum, JB, de Silva, V and JC Langford (2000). A Global Geometric Framework for Nonlinear Dimensionality Reduction. *Science* **290**(5500): 2319-2323.
- Tenenbaum, J.B., Kemp, C., Griffiths, T.L., and Goodman, N.D. (2011). How to Grow a Mind: Statistics, Structure, and Abstraction. *Science* **331**(6022), 1279-1285
- Thompson, W. B., Fleming, R.W. Creem-Regehr, S. and J. Stefanucci. (2011). *Visual Perception from a Computer Graphics Perspective*. CRC Press, Wellesley, MA, USA.
- Torralba A, Fergus R, Freeman WT. (2008). 80 million tiny images: a large data set for nonparametric object and scene recognition. *IEEE Trans Pattern Anal Mach Intell.* 2008 Nov;30(11):1958-70. doi: 10.1109/TPAMI.2008.128.
- Torralba, A., & Oliva, A. (2003). Statistics of Natural Images Categories.Network: Computation in Neural Systems, 14, 391-412.
- Torrance, K. E. and E.M. Sparrow (1967) Theory for off-specular reflection from roughened surfaces. *J. Opt. Soc. Am.* 57, 9:1105-1114
- Todd, J. T., Norman, J. F., & Mingolla, E. (2004) Lightness Constancy in the Presence of Specular Highlights. *Psychological Science*, 15:1, 33-39.
- T. Vladusich (2012), "Simultaneous contrast and gamut relativity in achromatic color perception." *Vision Research* 69, 49-63.
- Wandell, B.A. (1995). *Foundations of Vision*. Sunderland, MA: Sinauer.

- Ward, G.J. (1992). Measuring and modeling anisotropic reflection, ACM SIGGRAPH Computer Graphics, v.26 n.2, p.265-272.
- Wendt, G., Faul, F. & Mausfeld, R. Highlight disparity contributes to the authenticity and strength of perceived glossiness. *J. Vis.* 8, 14 (2008).
- Whittle, P. (1992a). Contrast-brightness and ordinary seeing. In A. Gilchrist, *Lightness, brightness and transparency*. Hillsdale, NJ: Erlbaum.
- Whittle, P. (1992b). The psychophysics of contrast-brightness. In A. Gilchrist, *Lightness, brightness and transparency*. Hillsdale, NJ: Erlbaum.
- Wolpert DM, Miall RC & Kawato M (1998) Internal models in the cerebellum. *Trends in Cognitive Sciences* 2:338-347.
- Zaidi, Q. (2011). Visual inferences of material changes: color as clue and distraction. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2 (6): 686-700.

# Visual perception of materials and their properties

Roland W. Fleming

*Experimental Psychology*

*Justus-Liebig-Universität Gießen*

## HIGHLIGHTS

In this review article, I suggest that the human visual system does not estimate the physical properties of materials and objects, but instead infers a kind of generative model—a ‘statistical appearance model’—that characterizes the natural degrees of variation between samples of different materials, expressed in terms of low- and mid-level image features. I discuss several curious findings from the literature and argue that such a model can account for these findings.

