ANNUAL REVIEWS

Annual Review of Vision Science

Lightness Perception in Complex Scenes

Richard F. Murray

Department of Psychology and Centre for Vision Research, York University, Toronto M3J 1P3, Canada; email: rfm@yorku.ca

Annu. Rev. Vis. Sci. 2021. 7:417–36

First published as a Review in Advance on July 1, 2021

The *Annual Review of Vision Science* is online at vision.annualreviews.org

https://doi.org/10.1146/annurev-vision-093019-115159

Keywords

human vision, psychophysics, modeling, lightness, brightness

Abstract

Lightness perception is the perception of achromatic surface colors: black, white, and shades of grey. Lightness has long been a central research topic in experimental psychology, as perceiving surface color is an important visual task but also a difficult one due to the deep ambiguity of retinal images. In this article, I review psychophysical work on lightness perception in complex scenes over the past 20 years, with an emphasis on work that supports the development of computational models. I discuss Bayesian models, equivalent illumination models, multidimensional scaling, anchoring theory, spatial filtering models, natural scene statistics, and related work in computer vision. I review open topics in lightness perception that seem ready for progress, including the relationship between lightness and brightness, and developing more sophisticated computational models of lightness in complex scenes.

INTRODUCTION

Lightness: perceived reflectance

Luminance: a measure of light intensity that takes into account the sensitivity of human vision to different wavelengths

Reflectance: the proportion of incident light reflected by a surface; a synonym is albedo

Lightness perception is the perception of achromatic surface colors: black, white, and gray reflective surfaces, which we usually perceive more or less accurately regardless of the amount of light shining on them or the context in which they appear.[1] Lightness has been a central research problem since the beginning of experimental psychology, for at least two reasons. One is that lightness is part of the more general problem of perceiving achromatic and chromatic surface colors, which is fundamental to many important visual judgements, such as recognizing objects and materials. The second is that lightness has turned out to be a surprisingly difficult phenomenon to understand, largely because of the deep ambiguity of retinal images. One of the key insights of modern vision research is that images provide much less information about the scenes that they depict than we might think, and the important role of ambiguity in lightness perception gives this research area connections to other important topics in vision science, including perception of shape, depth, luminance, and illumination (Brascamp & Shevell 2021).

One idea that motivates much work on visual perception, although it is not always explicitly stated, is that vision can be understood as a process of estimating properties of scenes and objects in the external world. Lightness perception, for example, can be understood as a process of estimating achromatic surface color, where by color we mean some physical property of real surfaces. More precisely, we can describe lightness as a perceptual estimate of reflectance, which is the proportion of incident light reflected by a surface. Without the idea that a percept is an estimate of some independently existing property, such as surface color, many common statements become meaningless: If a percept is not an estimate of something, how can it be accurate, or biased, or illusory? However, the idea that lightness is an estimate of surface color introduces a new problem: How can the visual system estimate the color of a surface in the external world given the limited information available at the retina?

A simple example illustrates the problem. When we view a surface, the light intensity (L) at a point on the retina is the product[2] of the illumination on the surface (I) and the surface reflectance (R): $L = IR$. If the retina records only the light intensity L, how can we unmultiply this product and recover the surface reflectance R? Clearly, this is only possible if the visual system exploits additional information, such as the spatial or temporal context of the point where reflectance is estimated.

The generalized bas-relief ambiguity gives another, more substantial illustration of why estimating reflectance from retinal images is a challenging problem (Belhumeur et al. 1999). Consider an arbitrary 3D scene of matte objects, as well as a 2D image of the scene. Belhumeur et al. (1999) showed that, if we put the scene through a shear transformation along the line of sight, then we can always adjust the positions of lights in the scene so that the locations of shadows in the image are unchanged. If we also adjust surface reflectance to counteract any resulting changes in image luminance, then even the pixel-by-pixel luminance values in the original image will be unchanged. Furthermore, this is only a lower bound on image ambiguity; besides shear transformations, we can adjust the shapes of objects in more or less arbitrary ways, and as long as we adjust lighting and surface color to counteract the resulting changes in the image, the original image will be unaffected. Thus, an image may appear to depict objects of particular shapes and

[1] Of course, there is a profound ontological difference between the subjective appearance of an achromatic color and the physical properties of the surface being perceived. As a result, it may take some effort to articulate what makes a percept accurate or inaccurate. I gloss over difficulties such as this when they are not crucial to the topic at hand.

[2] For simplicity, I omit constants relating to how units are defined. McCluney (1994) gives a clear treatment of the sometimes nuanced definitions of units for measuring light.

colors while in fact it is an equally good depiction of a wide range of other objects with very different shapes and colors. Any complete theory of lightness must explain why we see specific shades of black, white, or gray in an image, instead of the many valid alternatives.

Brightness: perceived luminance

In this article, I review selected psychophysical work from the past 20 years of research on lightness perception, with an emphasis on understanding complex scenes with cues to lighting, shape, and depth. During this time, new theoretical frameworks and experimental methods have emerged that have led to a renewed emphasis on perception of complex scenes, and a review should be useful for setting future research directions. I mostly discuss research areas that we might consider to be core topics in lightness, such as perception of matte surfaces under realistic lighting, and I do not address more exotic phenomena such as transparency. I also emphasize work that takes steps toward developing computational models of lightness. I do not cover work that is primarily about brightness, usually defined as perceived luminance, even though lightness and brightness are closely related; I briefly discuss this relationship in the section titled Open Problems. I take the approach of covering fewer topics in greater detail, so there are many interesting recent experiments that I do not touch on. Reviews by Adelson (2000), Gilchrist (2006), Brainard & Maloney (2011), Kingdom (2008, 2011), and Schirillo (2013) are valuable sources of further information.

SOME HISTORICAL CONTEXT

Gilchrist (2006) gives a comprehensive historical review of research on lightness. In this section, I note a few highlights that provide some context for the recent work discussed below.

The study of lightness began in earnest with the realization that percepts of surface color are more strongly correlated with surface reflectance, which is a distal property of surfaces themselves, than with the proximal light stimulus at the retina. This was understood by Ibn al-Haytham [1989 (1083)], and later by von Helmholtz [1924 (1910)] and Hering [1964 (1905)]. Of course, all perception must begin with the proximal stimulus, which depends on both surface reflectance and lighting, so if we are to perceive reflectance, then we must have some way of factoring out the effect of lighting. Von Helmholtz suggested that this was accomplished by using an estimate of illumination (possibly based on mean luminance) in an unconscious cognitive inference that results in an estimate of surface reflectance, whereas Hering argued that physiological mechanisms such as lateral inhibition and pupil dilation were mostly sufficient. (Unsurprisingly, this summary does not exhaust their views; e.g., von Helmholtz also considered physiological factors, and Hering allowed a role for memory.) The work reviewed below shows that this axis is still useful for describing theories of lightness: from theories of rational inference based on models of the physical world at one end, through various degrees of approximation and simplification, to theories of low-level image processing at the other end.

One long-standing approach to lightness has been to thoroughly study simple stimuli, such as edges and center-surround patterns, with the goal of using the visual system's response to these features to understand lightness in more complex scenes. Hess & Pretori [1970 (1894)], for example, reported parametric studies of simultaneous contrast effects, and similar studies have been influential throughout the history of lightness research (e.g., Heinemann 1955, Wallach 1948). Such effects, along with the discovery of lateral inhibition in visual neurons, suggested that center-surround contrast models could explain a great deal about lightness (e.g., Cornsweet 1970).

Others argued, however, that the most important lightness phenomena are primarily found in more complex scenes; thus, another approach has been to examine lightness in scenes of objects, surfaces, and lights. Katz (1935) described the phenomenology of lightness and lighting in realistic scenes, and also showed that lightness constancy tends to be stronger in large, complex scenes with many distinct elements. Furthermore, Koffka (1935) and other gestalt psychologists showed that factors such as grouping and figure-ground assignment are influential in lightness perception in ways that seem unlikely to be fully explained by the visual system's response to simpler stimuli.

In this case, too, both of these tendencies—reductionist and antireductionist—have long histories and are still found in current work. In this review, I focus on work that examines lightness in complex, realistic scenes, but I cover work with both tendencies.

Reductionist: in theories of vision, the claim that percepts of complex scenes can be explained in terms of the visual system's response to simple patterns

Anchoring: the problem of mapping luminance values to absolute (not just relative) lightness values

CURRENT RESEARCH

Anchoring Theory

Gilchrist and colleagues (Gilchrist 2006, Gilchrist et al. 1999) developed an anchoring theory of lightness perception that gives rules for dividing an image into regions of uniform illumination, called frameworks, and assigning a perceived reflectance to each image patch within a framework. A simplified outline of anchoring theory is as follows. The visual system divides an image into frameworks using cues to illumination edges, such as fuzzy shadow boundaries and depth discontinuities. The highest luminance in each framework is then an anchor that is assigned a local perceived reflectance of 0.90 (i.e., white). Other patches are assigned local reflectances based on the ratio of their luminance to the anchoring luminance in their framework; e.g., a luminance that is half the anchoring luminance is assigned a reflectance of $0.50 \times 0.90 = 0.45$. Finally, the perceived reflectance of each image patch is a weighted average of this local reflectance and the global reflectance computed in a global framework that includes the entire image. The weights in this average depend on how strongly the local framework is perceptually segmented within the global image; the stronger is the segmentation, the greater is the weight on the local reflectance. This outline omits some refinements, such as modified rules for large luminance regions and luminance outliers, but it captures the core of the theory.

Anchoring theory is an ambitious, broad-strokes account that aims at a general understanding of the most important factors in lightness perception. Its greatest strength is that it provides a compact and systematic account of how lightness depends on features such as nearby luminances, shadows, and depth edges that we would expect to provide useful information about surface reflectance. The theory makes qualitatively correct predictions for a wide range of scenes, and often quantitatively correct predictions as well, and it provides a starting point for understanding additional phenomena such as glow perception (Bonato & Gilchrist 1994; see also the sidebar titled Glow Perception).

GLOW PERCEPTION

What do we see when an object's luminance is greater than the luminance that a white object would have in the same environment? Sometimes the object can appear to emit light—that is, to glow. Bonato & Gilchrist (1994) suggested that the perceptual lightness continuum extends from black to white, and then beyond white to self-luminous. Recently, Murray (2020) found that a Bayesian lightness model that incorporates this idea can account for unusual properties of glow that Bonato & Gilchrist found psychophysically: Small surfaces appear to glow at lower luminances than do large surfaces, and large glowing surfaces make other surfaces in the environment appear darker. My colleagues and I have also found that luminance outliers are not the only factor influencing glow perception. Kim et al. (2016) showed that glow percepts can be toggled on and off by inverting the perceived 3D shape of a surface. When bright parts of a 3D surface are convex, people see them as a matte, diffusely illuminated material. When the the left- and right-eye images of a stereo pair are switched, so that the luminance pattern is practically identical, but the bright regions become concavities where it would be difficult for light to reach, people perceive the concavities as glowing.

The main weakness of the theory is that it has no computational implementation. Identifying lighting boundaries and assigning weights to local and global frameworks are left as tasks for the modeler. As a result, there are many scenes where the predictions of anchoring theory are unclear (Zeiner & Maertens 2014), and experiments must be designed around images for which predictions can be made. The theory does have some quantitative components, and Economou et al. (2007) point out that it can also be difficult to make predictions from more computational competing theories. Nevertheless, it would be an important step forward to formulate anchoring theory computationally—even parts of it, or for a limited stimulus domain. Simple quantitative theories can make surprising predictions, and it is easier to evaluate the content and success of a theory when its predictions are completely independent of the modeler's choices.

For example, Gilchrist (2006) has described anchoring theory as a theory of errors in lightness perception, as several of the visual behaviours that it documents seem to be arbitrary. For example, why should perceived lightness be a weighted average of lightness computed in local and global frameworks? However, Murray (2013, 2020) has shown that several of the rules of anchoring theory follow from a probabilistic, normative model based on simple assumptions about lighting and reflectance. If this is correct, then facts about lightness that appear in anchoring theory as arbitrary rules may in fact result from a deeper process of rational inference guided by natural scene statistics.

Normative model: a model that implements or is motivated by an optimal strategy for solving a problem

Multidimensional scaling (MDS): a method of modeling the similarity of pairs of elements by arranging them in an n-dimensional space

Multidimensional Scaling

Why do we believe that lightness is a fundamental perceptual dimension at all? Lightness and surface reflectance are only weakly related to simple stimulus properties such as local image luminance. What is the evidence, then, that perceived reflectance, perceived lighting, and similar terms are appropriate for describing the elements of our perceptual world?

Logvinenko and colleagues (Logvinenko 2015, Logvinenko & Maloney 2006, Logvinenko et al. 2008) addressed this question in a creative series of studies that used multidimensional scaling (MDS) methods and novel variants to investigate the perceptual dimensions of achromatic color. An MDS analysis attempts to arrange stimuli in an n-dimensional space, such that stimuli that people perceive to be similar are close together in the space, and stimuli that they perceive to be very different are far apart. If we can find such an arrangement, then it may provide some insight into the perceptual representation of the stimuli. For example, if no such arrangement can be found in a one-dimensional space, but one can be found in a two-dimensional space, and if a three-dimensional space does not provide any further improvement in modeling perceived similarity, then this suggests that similarity judgements are based on two perceptual dimensions. Furthermore, if surface reflectance increases along one axis of the two-dimensional space, then this suggests that some perceptual representation of surface reflectance plays an important role in similarity judgements.

Logvinenko & Maloney (2006) showed observers simple geometric paper patterns (**Figure 1*a***), and on each trial, observers used a 30-point scale to rate the dissimilarity between two randomly selected test patches. The patches varied in reflectance and illumination, but the instructions did not mention these properties; observers were simply asked to rate dissimilarity. Logvinenko & Maloney modeled these difference ratings using classic MDS and a novel maximum-likelihood difference scaling (MLDS) method. Both analyses showed that observers' dissimilarity ratings could be accounted for by arranging the stimuli in a two-dimensional space, with dimensions roughly corresponding to reflectance and lighting intensity (**Figure 1*c***). An alternative outcome could have been, for example, that dissimilarity ratings were predicted by a one-dimensional continuum corresponding to image luminance. However, this was not found,

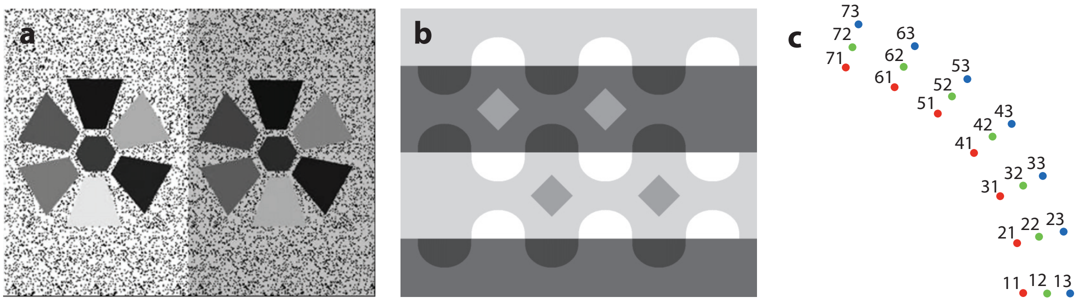

Figure 1

(*a*) Illuminated paper stimulus from Logvinenko & Maloney (2006). (*b*) The snake illusion (Adelson 2000). All four diamonds are the same shade of gray, but the two on top appear much lighter. (*c*) Results of Logvinenko & Maloney's (2006) scaling analysis. Points representing 21 stimuli (7 reflectances, 3 lighting intensities) are arranged in a two-dimensional space, such that the distance between any two points indicates the perceived dissimilarity of the two corresponding stimuli. In the two-digit labels, the first digit represents reflectance (1 = low, 7 = high), and the second represents lighting intensity (1 = low, 3 = high). Panels *a* and *c* adapted with permission from Logvineko & Maloney (2006), copyright The Association for Research in Vision and Ophthalmology. Panel *b* adapted with permission from Adelson (2000), copyright The MIT Press.

and in fact, stimulus pairs that had the same luminance but different reflectances and lighting intensities elicited some of the highest dissimilarity ratings. Thus, reflectance and illumination emerged as perceptual dimensions of achromatic color, and luminance did not, even in a task where observers were not explicitly asked to judge any of these properties—an intriguing finding. [For a complication to this story, the reader is referred to Madigan & Brainard (2014), who used methods that were similar, although not identical, to Logvinenko & Maloney's, and found that a one-dimensional space was adequate for modeling dissimilarity judgements.]

Logvinenko (2015) extended this work by varying not only reflectance and lighting intensity, but also the slant of test patches relative to the light source. In this case, MDS indicated a three-dimensional space for achromatic colors, and Logvinenko suggested that the third dimension represented the appearance of attached shadows at various slants. (Alternatively, did the third dimension simply represent 3D surface orientation, unrelated to achromatic color? This is probably not the case, as observers were specifically instructed to rate the dissimilarity of the achromatic color appearance of the test patches. However, the open-ended rating task, with no possibility of feedback, does raise a concern that observers in MDS studies may sometimes base their responses partly on stimulus properties that are not the target of the study.)

Logvinenko et al. (2008) used similar methods to examine perception of lightness and lighting in the snake illusion (Adelson 2000) (**Figure 1*b***). In this illusion, all of the diamonds have the same reflectance, but some look much lighter than others. A common explanation is that the figure appears to have horizontal strips of bright and dark lighting, perhaps due to shadows or filters, and that we discount these lighting conditions when estimating reflectance. As a result, diamonds in dimly lit regions appear lighter. But do we really see lighting boundaries in this flat, printed figure? Logvinenko et al. asked observers to rate the dissimilarity of pairs of diamonds in variants of the snake illusion, with different simulated lighting conditions and simulated reflectances, and tests based on the MLDS method (see above) showed that these ratings could be modeled on a one-dimensional continuum. That is, in these stimuli, simulated lighting and reflectance differences had qualitatively similar effects on appearance. This is in sharp contrast to Logvinenko & Maloney's (2006) findings with real papers and lights (see above), where scaling models produced a two-dimensional space, and lighting and reflectance could not be traded off against one another. Logvinenko et al.'s finding complicates the usual explanation of the snake illusion and similar figures: It is not simply that we see lighting boundaries, tout court, and that this explains the illusion. Instead, lightness mechanisms may be sensitive to pictorial cues to lighting boundaries (such as X-junctions), yet these cues may have little effect on the actual appearance of lighting conditions. [This echoes Brainard & Maloney's (2011) observation that the implicit lighting estimates that guide equivalent illumination models (see below) often differ from explicit judgements of lighting.] This finding also highlights the fact that MDS and MLDS can be seen as more systematic forms of the phenomenological method (Katz 1935), and that they share the strengths and weaknesses of that method. They can give a clear description of perceptual phenomena that we should seek to model but may provide little information about how those phenomena are computed by the visual system.

Logvinenko et al.'s (2008) findings with the snake illusion also suggest that we should be cautious about studying lightness using simplified pictures instead of real scenes with real lighting. This is a point that I return to below (see the section titled Lightness and Brightness).

Conditional random field (CRF): a probabilistic model in which each element has direct statistical dependencies only on neighboring elements; statistical dependencies between distant elements are mediated by dependencies between neighbors

Bayesian Intrinsic Image Models

The fundamental challenge of lightness perception is ambiguity: Any given image could have been generated by a wide range of combinations of reflectance, surface orientation, and lighting, and thus from the image alone it is impossible to infer a unique spatial distribution of reflectance. This inference requires further information. One possible source of information is knowledge about which patterns of reflectance, surface orientation, and lighting are most likely to occur in the limited range of scenes that we typically encounter. This suggests that a Bayesian approach to lightness may be productive. The idea that knowledge about typical scenes should be useful for lightness perception is a natural one, and several authors have taken this approach (e.g., Adelson & Pentland 1996, Land & McCann 1971), but progress in turning this insight into general-purpose, image-based computational models has been slow. There has been progress with Bayesian models of color constancy, but these typically exploit regularities in the spectra of illuminants and surfaces, instead of spatial statistics (e.g., Brainard et al. 2006). Lightness models do not model spectra, so they must rely on other sources of prior knowledge.

Allred & Brainard (2013) developed a Bayesian model of lightness and lighting perception on a 5×5 grid of gray-scale squares. Observers viewed grid stimuli (**Figure 2*a***) on a high-dynamic-range display and reported the lightness of the central square by choosing a matching paper patch from a separate palette (Allred et al. 2012). The driving assumptions of the model were that (*a*) the reflectance of each square is independently sampled from a probability distribution on the unit interval, and (*b*) illuminance values at neighboring squares are correlated. The authors fit the model by finding the reflectance and illuminance distributions that maximized the model's ability to predict human observers' lightness matches. The resulting model accounted for several features of the psychophysical data, some expected from previous studies and some novel. For example, the model predicted that increasing the luminance of the surrounding region causes the test patch to appear darker—the well-known simultaneous contrast effect [Hess & Pretori 1976 (1894)]. It also predicted a more subtle effect where increasing the luminance of the test patch has a greater effect on the patch's perceived lightness when the initial luminance of the patch is low than when it is high; the model suggested that this occurs because the test patch itself affects the observer's estimate of local lighting conditions.

Murray (2013, 2020) also developed a Bayesian model of lightness and lighting perception, with different methods and goals. This model described lightness perception on a 16×16 grid (**Figure 2*b***). The model was a conditional random field (CRF), so its statistical assumptions were formulated as potential functions on small regions (2×2 grid squares) of reflectance and

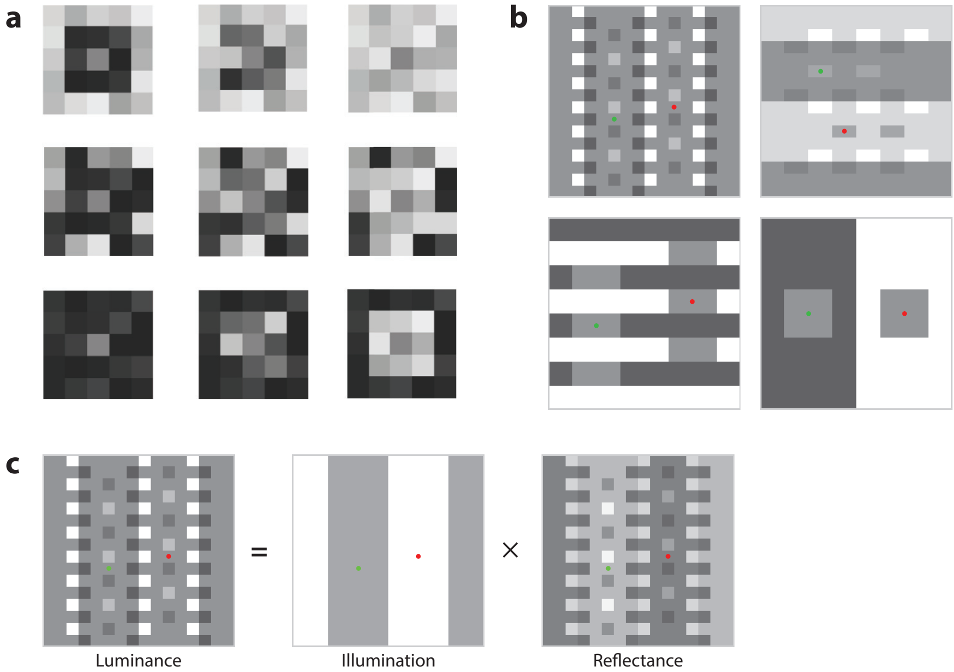

Figure 2

(*a*) Grid stimuli from Allred & Brainard (2013). The squares at the centers of the 5 × 5 patterns are all the same shade of gray but often appear different because of the surrounding context. Observers chose a match patch from a separate palette to match the perceived lightness of the central square. (*b*) Grid stimuli from Murray (2020). Observers reported whether each stimulus appeared lighter at the location of the red or green dot. In each case, the dots were in regions with the same physical shade of gray. (*c*) Murray's model decomposes a luminance stimulus, such as the variant of the argyle illusion shown, into illumination and reflectance images. Panel *a* adapted with permission from Allred & Brainard (2013), copyright The Association for Research in Vision and Ophthalmology. Panels *b* and *c* adapted with permission from Murray (2020), copyright The Association for Research in Vision and Ophthalmology.

illuminance, with high potentials assigned to patterns that are less probable. One assumption embedded in these potentials, for example, was that straight luminance edges are more likely than curved edges to be seen as shadow boundaries (Logvinenko et al. 2005). A belief propagation algorithm leveraged these local assumptions to make globally optimal assignments of reflectance and illuminance to a stimulus image. Human observers reported the relative lightness of test locations in grid versions of several well-known and difficult-to-model lightness illusions, such as the argyle illusion (**Figure 2*b***), and the model made qualitatively correct predictions of observers' judgements in many of these figures. Its main failure was that it could not account for assimilation effects such as White's illusion (White 1979), where patches that are surrounded mostly by high-luminance regions appear lighter than patches surrounded by lower-luminance regions; these effects are challenging for models based on discounting illumination, as it is unclear why increasing the luminance of the surround should decrease the estimate of local lighting intensity. The model was, however, also able to account qualitatively for several broad phenomena in lightness perception, such as the effect of perceived lighting boundaries on lightness, articulation effects, and glow perception.

These studies suggest that Bayesian approaches to lightness, despite their slow progress to date, are nevertheless promising. Bayesian models are theoretically well motivated, are easily interpretable, and have the potential to account for a wide range of phenomena using assumptions grounded in natural scene statistics. They may be able to synthesize two broad approaches to lightness described above (see the section titled Some Historical Context), in that they use simple, local image properties to explain perception of complex images, but these properties (e.g., smoothness, straightness) are chosen because they provide information about critical features of realistic scenes, such as shadows and reflectance edges. An obvious limitation of the two studies reviewed in this section is that they modeled coarse checkerboard patterns instead of realistic images. A further limitation is that, instead of incorporating measurements from naturally occurring scenes, both studies fitted parametric distributions of reflectance and lighting to make the model account for behavioral data. In the section titled Open Problems, I suggest that developments in computational modeling may provide tools for overcoming these limitations.

Equivalent Illumination Models

Perceptual errors can be highly informative for testing theories of vision, as all models of perception must agree when their predictions are correct, but each can fail in its own unique way [Gilchrist 2006, Tolstoy 2006 (1874)]. Can we develop a theory of lightness that accounts for the specific pattern of correct and incorrect percepts across a wide range of scenes?

Brainard & Maloney (2011) reviewed equivalent illumination models (EIMs) of surface color and lightness. These are normative models that describe vision as a fundamentally rational process of estimating reflectance by attempting to invert the physics of image formation. They do not necessarily predict that lightness percepts are always veridical, however, because they are also parametric models and assume that the visual system must estimate important parameters of a scene, such as lighting conditions, to infer lightness. If an observer's estimates of these parameters are inaccurate, then their lightness percepts will have characteristic biases. Thus, although they have a normative component, EIMs are descriptive models that aim to account for people's actual behavior.

Bloj et al. (2004) used an EIM to model how lightness changes as a function of the slant of a test patch in complex scenes that provide many cues to lighting conditions (**Figure 3*a***). In principle, observers could estimate the lighting conditions in such scenes and use this information, along with the luminance of a test patch, to infer the test patch reflectance. On each trial, observers viewed a test patch on a pedestal, rotated to some orientation relative to the light source, and from a 6 × 6 palette chose the patch that most closely matched the lightness of the test patch

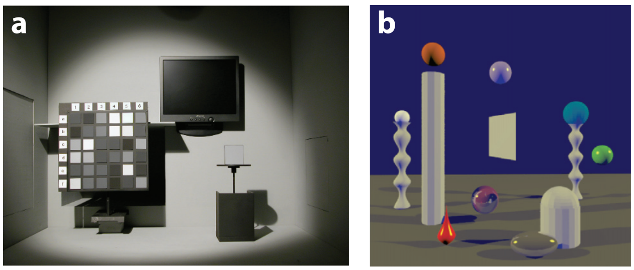

Figure 3

(*a*) A typical stimulus from Bloj et al. (2004). The grid at the left is the 6 × 6 palette of match patches, and the pedestal at the lower right supports the test patch. (*b*) A typical stimulus from Doerschner et al. (2007). The oblique central square is the test patch. The other objects in the scene provide cues to the lighting conditions, which consist of a diffuse blue sky and two yellow point light sources that are not in the observer's field of view. Panel *a* adapted with permission from Bloj et al. (2004), copyright The Association for Research in Vision and Ophthalmology. Panel *b* adapted with permission from Doerschner et al. (2007), copyright Elsevier.

(**Figure 3*a***). Bloj et al. found that observers matched reflectance correctly to some extent, but also that there were systematic departures from constancy. The strongest pattern of errors was that the lightness of the test patch decreased when it had a large slant relative to the light source, whereas if observers had perfect constancy, then lightness would not vary with surface orientation. When an EIM was fitted to these data, it was able to explain these biases by supposing that observers estimated the direction of the light source correctly but that they consistently overestimated its diffuseness.

It is clear why overestimating diffuseness would produce such a bias: As a test patch is slanted away from a light source, its luminance decreases, and if the lighting is highly collimated (i.e., not diffuse), then luminance decreases rapidly as slant increases. An observer who overestimates diffuseness expects that luminance should not decrease much as slant increases, so they mistakenly attribute the reduced luminance at high slants to a decrease in surface reflectance. Bloj et al. (2004) found that the EIM not only accounted for observers' biases qualitatively in this way, but also that it gave a good quantitative description over a wide range of slants, e.g., fitting the data better than a simple mixture model where lightness matches at all slants were a fixed weighted sum of the matches expected from full constancy (reflectance matching) and no constancy (luminance matching). Furthermore, the EIM was able to accommodate large individual differences by attributing different diffuseness estimates to different observers. Boyaci et al. (2003) independently reported similar experiments with similar conclusions, and Morgenstern et al. (2014) examined some differences between the two studies that may have emerged because Bloj et al. (2004) used physical stimuli, whereas Boyaci et al. (2003) used computer-generated scenes.

Doerschner et al. (2007) used an EIM to test whether observers could simultaneously discount two separate point light sources when estimating surface color. They showed that when two light sources are close together, a normative model treats them much like a single diffuse light source for the purpose of discounting illumination, but that when they are farther apart, the normative model has a qualitatively different pattern of discounting illumination. In their experiment, observers viewed a test patch at various orientations in a complex scene, illuminated by a uniform blue sky and two yellow point light sources (**Figure 3*b***). Observers adjusted the chromaticity of the test patch so that it appeared achromatic. Doerschner et al. found that observers made qualitatively different patterns of achromatic settings depending on whether the two point lights were close together or far apart; in both cases, their behavior was similar to the normative model.

According to EIMs, people perceive lightness as if a scene was illuminated by some not necessarily accurate estimate of lighting conditions. The value of such models depends on how far we can push this "as if." If idiosyncratic patterns in a wide range of lightness judgements are consistent with a rational observer who has a particular estimate of lighting conditions, then EIMs provide a powerful, compact, and interpretable description of lightness perception. The studies reviewed in this section suggest that EIMs fare well by this criterion. Bloj et al. (2004) and Boyaci et al. (2003) showed that an EIM accounted for how lightness estimates varied with slant significantly better than a simple mixture model, and Doerschner et al. (2007) found that, even in complex scenes with two light sources, lightness matches were just what we would expect from an observer who makes a reasonable but imperfect estimate of lighting conditions.

A Wholly Empirical Theory

Purves & Lotto (2010) developed a framework in which the visual system uses knowledge from past experience to interpret retinal images. Corney & Lotto (2007) applied this framework to lightness perception. In their stimulus set, reflectance images were dead-leaves patterns consisting of many randomly placed, randomly gray-colored, overlapping circles (Lee et al. 2001). Illumination images varied slowly and smoothly from pixel to pixel, and luminance images were created by multiplying reflectance and illumination images pointwise. Corney & Lotto trained artificial neural networks to take a luminance image as input and predict the center pixel of the corresponding reflectance image. After training, the networks had some ability to predict reflectance, as was expected. Significantly, though, they also exhibited several robust phenomena found in human lightness perception, including simultaneous contrast, articulation effects, and assimilation effects. Furthermore, they produced outputs consistent with lightness illusions perceived by human observers, such as the Vasarely illusion, Mach bands, the Hermann grid, and White's illusion.

The idea that visual perception exploits statistical regularities in natural images is not new (see the section titled Bayesian Intrinsic Image Models; see also Barlow 1961, Brunswik & Kamiya 1953, Geisler 2008, Knill & Richards 1996), but Purves, Lotto, and their colleagues' work is valuable in that it shows how some well-known perceptual biases may be adaptive byproducts of the visual system being tuned to unexpected regularities in natural scenes. It seems unlikely, though, that such an approach can ever be a wholly empirical theory of vision (as they call it), or even of lightness. The visual system uses learned regularities to infer object properties from retinal images—so far, so good. But which regularities does it use, and how does it use them? Simply knowing the marginal pointwise distribution of luminance, for example, will not be useful for difficult inferences such as lightness perception. Yet the visual system cannot use arbitrarily complex statistical properties of natural images either. (Consider the approximately seven-million-dimensional probability distribution of cone activations in each eye.) Between these two extremes lie countless models, and a probabilistic theory of vision is arrived at by building and testing such specific, limited models. Corney & Lotto argue, for example, that "since resolving stimulus ambiguity is a challenge faced by all visual systems, a corollary of these findings is that human illusions must be experienced by all visual animals regardless of their particular neural machinery" (p. 1790). However, surely different animals can (and do) make perceptual inferences using different statistical properties of images, with different levels of statistical sophistication, resulting in different illusions.

A Spatial Filtering Model

It has long been known that many retinal ganglion cells in mammalian visual systems have an ON-center, OFF-surround response: Each such cell is stimulated by light falling in a small disk-like region of the retina and inhibited by light in a surrounding ring-like region (Kuffler 1953). The responses of these cells are qualitatively consistent with some classic lightness and brightness illusions, such as simultaneous contrast effects where a medium-luminance square surrounded by a low-luminance region appears brighter than an identical square surrounded by a high-luminance region. There is a long history of spatial filtering models that use such center-surround mechanisms to explain how we perceive shades of light and dark (e.g., Cornsweet 1970), but these are almost always presented as models of brightness (perceived luminance), and not of lightness (perceived reflectance), when this distinction is made. However, Dakin & Bex (2003) developed a spatial filtering model of lightness that provides an opportunity to consider how the rich computational literature on brightness can contribute to modeling lightness as well.

Dakin & Bex's (2003) model is simple: To predict the lightness perceived in an image, it adjusts the Fourier spectrum of the image to match the $1/f$ spectrum that is typical of natural images (Field 1987). This operation can be described as an adaptive linear filter, or as band-pass contrast normalization. Dakin & Bex showed that if we band-pass filter an image that has uniform light and dark regions, then those regions may have the same luminance after filtering, but they still appear light and dark (**Figure 4**). This illusion only occurs, however, if there are residual low spatial frequencies in the filtered image. Dakin & Bex's model restores the low-spatial-frequency content that has been mostly filtered out and thereby restores the intensity difference between bright and dark regions. [There is clearly a tension to be explored between this model, which

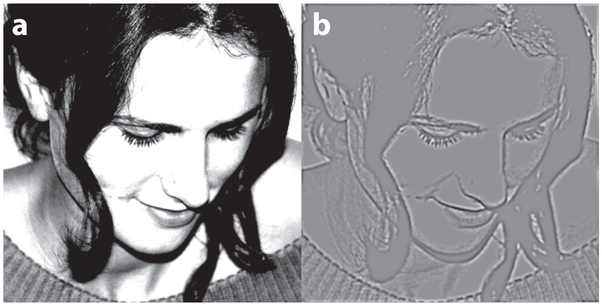

Figure 4

(*a*) A natural image from Dakin & Bex (2003). (*b*) The same image, high-pass filtered. The large uniform regions in the filtered image (e.g., forehead and hair) are physically the same shade of gray, but some appear lighter than others. Dakin & Bex's lightness model restores the weakened low-spatial-frequency components of this image and correctly predicts that the woman's face appears lighter than her hair. Figure adapted with permission from Dakin & Bex (2003), copyright The Royal Society, UK.

usually accomplishes filling-in by boosting low spatial frequencies, and Shapiro & Lu's (2011) model, which computes brightness by removing low spatial frequencies.] Interestingly, then, the model produces a filling-in effect without the node-to-node signal propagation found in other models of filling-in (e.g., Grossberg & Todorovic 1988).

A minor suggestion, which applies to most other filtering models as well, is that the model would probably be stronger if it was based on local filtering operations, rather than operations that depend on the whole image. Lightness and brightness depend on context but not usually on interactions between arbitrarily distant image features. Robinson et al. (2007) illustrate how to make a global filtering model [specifically, the oriented difference of Gaussians (ODOG) model] into a local model, and they also show that, with this modification, the model accounts for a wider range of phenomena.

A more substantial question is to what extent Dakin & Bex's (2003) model accounts for lightness percepts beyond filling-in effects. A simple argument suggests that there is a wide range of phenomena that it does not account for. The model does not make large adjustments to a typical natural image that already has a $1/f$ spectrum, so it seems to predict that, in such images, lightness is approximately proportional to image luminance. Lightness is not proportional to luminance, however, in images with strong lighting boundaries: A fixed luminance usually looks much lighter in a dimly illuminated region than in a bright region. If this argument is correct—it seems plausible, but it would need to be tested—then Dakin & Bex's model accounts for filling-in, but it may not be a strong theory of lightness more generally. (It is not clear whether it was meant to be. The paper's title suggests a filtering mechanism only for brightness filling-in, but the main text seems to aim at a broader theory of lightness and shows, for example, that the model also accounts for White's illusion.)

Betz et al. (2015) suggested that this failure to process boundaries correctly is a general problem with spatial filtering models. In band-pass masking experiments, they showed that the spatial frequencies that affect several filtering models' lightness or brightness estimates are inversely proportional to stimulus size, whereas the spatial frequencies that guide human observers' lightness judgements do not vary as strongly with stimulus size. They pointed out that lightness mechanisms based on edges, whose local spatial frequency content does not vary much with overall stimulus size, may be able to account for these findings. A related point is that small adjustments to an image can create or destroy the impression of lighting boundaries

(Adelson 1993, Gilchrist 1977), which can have a large effect on perceived lightness but a small effect on the response of filtering models (Murray 2020).

Betz et al.'s (2015) findings may not be fatal for all spatial filtering models, as such models can be designed to operate at a fixed scale instead of being scale invariant, and even adding an early spatial filter based on the contrast sensitivity function can make a model strongly scale dependent (Chung et al. 2002). Spatial filtering models of brightness can account for an impressively wide range of phenomena (Blakeslee & McCourt 2012), and it is tempting to think that, with suitable modifications, they could provide an implementation-level model of lightness phenomena that are described at a more abstract level by theories such as Bayesian intrinsic image models. There has been little work on computational models of lightness perception that can be applied to arbitrary images (unlike brightness perception, where such models are common), and spatial filtering models provide one paradigm for making progress on this front.

Lambertian: a material type that has the same luminance from all viewpoints; an approximate synonym is matte

Computer Vision

Reflectance estimation has long been an active topic in computer vision, where it is known as intrinsic image decomposition or inverse rendering (Barrow & Tenenbaum 1978, Ramamoorthi & Hanrahan 2001). That literature is mostly outside the scope of this review, but I mention two developments of interest.

Barron & Malik (2015) developed a Bayesian algorithm (called Shape, Illumination, and Reflectance from Shading, or SIRFS) for estimating reflectance, shape, and lighting from single images, guided by statistical assumptions about natural scenes. They measured priors for reflectance and shape from a database of 3D representations of natural objects, and priors for illumination from a set of low-dimensional spherical harmonic representations of natural lighting. SIRFS assigned reflectance, shape, and illumination to an image by finding the combination of these properties that (*a*) had the greatest prior probability and (*b*) exactly reproduced the image under a Lambertian shading model. From a single image, the model was able to make reasonable estimates of the reflectance, shape, and lighting of objects for which the ground truth was known. Naturally, the model had some limitations. For example, it required the boundaries of target objects to be traced out and assumed that the surface normals along these boundaries were all in the image plane. It also assumed that lighting conditions were constant across the object, whereas human vision is highly tolerant of lighting variations (Ostrovsky et al. 2005, Wilder et al. 2019). An important contribution of this work was to show the power of 3D scene statistics in addition to more widely studied 2D image statistics, and also to show that it may be easier to estimate lighting, reflectance, and shape simultaneously than individually. As discussed below (see the section titled Computational Tools), it also provides new computational methods for developing more realistic probabilistic models of human vision.

Yu & Smith (2019) used deep learning to train a convolutional neural network (InverseRenderNet) to solve the same problem addressed by SIRFS, namely estimating reflectance, shape, and lighting from a single image. They used a few simple priors (e.g., smooth reflectance) along with some clever heuristics to train the network using web-crawled images for which the ground truth of reflectance, shape, and lighting were unknown. For example, they used multiview stereo to find corresponding points in different images of the same scene and penalized the network for assigning different reflectances to the same physical point in different images. The trained network produced reflectance and surface normal maps from input images, and a closed-form calculation based on a Lambertian shading model used these maps to infer global lighting conditions. Interestingly, the reflectance and normal maps were computed first, without an explicit lighting estimate, and the lighting estimate was found afterward as a byproduct; this is the opposite of the order in many human vision models from von Helmholtz onward (e.g., equivalent illumination models), where lighting is estimated first and plays an essential role in inferring reflectance. (This network was a black box, however, so it may also have computed an implicit lighting estimate on the way to computing reflectance.) During training, the network was penalized for making reflectance, shape, and lighting estimates that did not account for the input image exactly, but this was a graded penalty and not a hard constraint. Rutherford & Brainard (2002) found that human observers often make reflectance and lighting estimates that are not jointly consistent with the observed image luminance; it would be interesting to know if the estimates from InverseRenderNet show a similar pattern of inconsistencies.

Until recently, it was common to convey the mystery of human vision by saying that no presently existing machine can emulate it. That is still true, but the gap is narrowing, and computer vision algorithms can provide novel, testable hypotheses for understanding biological vision (Yamins et al. 2014). It would be useful to know, for example, whether SIRFS or InverseRenderNet show characteristics of human lightness perception, such as contrast, assimilation, and articulation effects, that were not deliberately built into them, as this would suggest that these properties emerge robustly in systems that exploit natural scene statistics. More generally, computer vision models are often evaluated using summary statistics of performance on standard data sets, but evaluating their behavior parametrically along important stimulus dimensions, as is common in psychological experiments, would be more revealing when judging them as starting points for models of human vision.

Neither SIRFS nor InverseRenderNet seems likely to be able to explain several phenomena that have motivated recent work on human lightness perception. They both assume smooth, globally consistent lighting, so they probably cannot explain figures like the snake illusion that seem to rely on lighting boundaries, unless they accomplish this indirectly by assigning different surface orientations to perceived lighting regions. They also use a local Lambertian shading model, so they do not address specularities, interreflections, cast shadows, transparency, or material properties. There is nothing in their approaches, however, that would prevent future revisions from addressing these phenomena.

OPEN PROBLEMS

Lightness and Brightness

When we see two equal-reflectance gray patches, one in light and one in shadow, we can usually see that they are approximately the same shade of grey. This is lightness constancy. At the same time, we also recognize that there is a separate bright–dark continuum on which the patch in shadow appears darker than the one in light. This is brightness. Lightness is usually defined as perceived reflectance and brightness as perceived luminance (Arend 1993). These definitions can leave the impression that lightness and brightness are simply independent, largely unrelated perceptual dimensions, like hue and depth, for example, and that we can judge either of them at will. In fact, their relationship is more nuanced than this (Blakeslee et al. 2008), and there is still much that we do not understand about it. This is partly because the literatures on lightness and brightness have often proceeded independently, with little interaction. In this section, I suggest some avenues for investigation.

First, Logvinenko and colleagues' (Logvinenko 2015, Logvinenko & Maloney 2006, Logvinenko et al. 2008) scaling studies (reviewed above) did not reveal a role for perceived luminance in the appearance of achromatic surfaces. This would be surprising if brightness were a fundamental perceptual dimension. Does this mean that brightness is a weak and easily disregarded feature of realistic scenes, or instead, did observers perceive brightness but recognize that it was not an intrinsic property of the test patches that they were asked to judge? Lighting intensity, however, which we might also consider not to be an intrinsic surface property, did emerge as a dimension of achromatic color, so we are still left with the question of why some perceptual dimensions affect judgements of surface appearance and others do not.

Second, seeing a raw luminance distribution, for example when making a shaded drawing, is notoriously difficult (Perdreau & Cavanagh 2011). Furthermore, current brightness models produce outputs that bear little resemblance to simple image luminance. For example, ODOG does not respond to mean luminance, and its output is based on normalized, orientation-selective, band-pass channels (Blakeslee & McCourt 1999). If this is correct, then should we think of brightness as perceived luminance at all? If not, then what is it, and why does the visual system compute it? Is it a byproduct of early neural computations, with little adaptive value, or is it, in some sense that would need to be made more precise, one step on the way to lightness?

Third, how are lightness and brightness related in real and virtual scenes? There have been few studies that directly compared lightness and brightness in the same scene (e.g., Arend & Spehar 1993a,b; Blakeslee & McCourt 2012), and even fewer that did this using real objects and lights (e.g., Blakeslee et al. 2008, Jacobsen & Gilchrist 1988). Computer-generated images have obvious advantages for stimulus control, but if they are not completely realistic, then they probe lightness in an unusual type of pictorial space, rendered on pixellated, glowing surfaces. It is an open question how closely this mimics lightness in more realistic scenes. Some studies have found that real and virtual stimuli produce similar results (Blakeslee et al. 2008, Radonjic et al. 2016), but others have found discrepancies (Morgenstern et al. 2014, Patel et al. 2018), and until we understand the relevant stimulus factors, equivalence between specific real and virtual stimuli should always be an empirical conclusion, not an assumption. How much of our understanding of lightness and brightness has been shaped by the fact that some studies have used real objects and lights, whereas others (including most recent studies of brightness) have used virtual stimuli? How do lightness and brightness typically differ between real and virtual environments?

I do not have answers to these questions. I raise them to point out some ways in which the relationship between lightness and brightness is not well understood and to suggest that there is room for important new work on this fundamental issue.

Computational Tools

Research on computer vision and artificial intelligence has made remarkable progress on computational tools that are well suited to modeling human vision, and lightness in particular. Two examples of this are Markov random fields (Koller & Friedman 2009) and artificial neural networks (Goodfellow et al. 2016), which are powerful frameworks that have matured to a point where they can be used effectively by nonexperts. Flexible methods for learning complex statistical distributions that can be used as priors have also emerged (Kobyzev et al. 2020, Roth & Black 2005). The idea that human vision overcomes ambiguity by exploiting regularities in natural scenes is well established, but psychological models that build on this insight are often highly simplified and frequently do not do justice to the rich statistical structure of natural scenes. A notable feature of Barron & Malik's (2015) solution to estimating reflectance, shape, and lighting (reviewed above) is how familiar its approach is in broad outline (see Adelson & Pentland 1996) and yet how novel it is in making adroit use of computational tools to leverage natural scene statistics in this problem. When Fourier methods were adopted by vision researchers, they gave a new perspective on familiar problems that led to an explosion of research and fundamental advances (DeValois & DeValois 1988, Graham 1989). New computational modeling tools have similar potential, as they make it possible to go beyond simplified problems, to develop and test more sophisticated models of realistic scenes and tasks.

A ModelFest Proposal

There has been little work on comparing the wide range of theories of lightness on a common set of stimuli. For the most part, each investigator has tested their own model on their preferred stimuli, and as a result, it is often hard to be certain about the strengths and weaknesses of various theories. Model testing is most enlightening when we can show not only that a particular model explains a phenomenon, but also that other plausible models do not. ["It is not enough to succeed; others must fail" (Murdoch 1973, p. 98).] Use of standard test sets has been highly productive in other research areas, such as ModelFest for spatial vision (Carney et al. 1999) and ImageNet for machine learning (Deng et al. 2009), and research on lightness (and brightness) would benefit from such a project as well.

CONCLUSION

Several elements of potentially strong, computational theories of lightness have emerged during the time covered by this review. We have a partial understanding of the influence of lighting and depth boundaries on lightness (Gilchrist et al. 1999). We have some insight into the perceptual dimensions of achromatic color that should be modeled (Logvinenko 2015). We have some measurements of the statistics of reflectance, shape, and lighting in natural scenes (Adams et al. 2016, Barron & Malik 2015, Purves & Lotto 2010). We have computational tools that can turn hypotheses about the natural scene statistics that guide lightness perception into algorithms for inference (Goodfellow et al. 2016, Koller & Friedman 2009). We have studies that have taken steps in modeling lightness and lighting in simple stimuli (Allred & Brainard 2013, Murray 2020). A promising direction for future research is to integrate these elements into broader computational models of lightness that can be applied to arbitrary images. Clearly, there are many possible approaches to this goal, and it is unlikely to be reached in a single leap. There has been encouraging progress, however, in incorporating insights from psychophysical experiments into increasingly quantitative and computational models of lightness, and elements are available to support new advances in this direction.

SUMMARY POINTS

1. Human vision discounts complex lighting conditions when estimating surface reflectance.
2. Achromatic surface color is multidimensional.
3. Lightness mechanisms exploit statistical regularities in lighting and reflectance.
4. There has been substantial progress on quantitative and computational models of lightness.
5. New computational methods provide promising tools for modeling lightness.

FUTURE ISSUES

1. There has been substantial progress on quantitative and computational models of lightness. New computational methods provide promising tools for further advances.

2. The relationship between lightness and brightness is a fundamental issue in which there is room for important new work.
3. The idea that lightness perception is guided by natural scene statistics is long-standing. Recent work on measuring and modeling properties of natural scenes makes it possible to exploit this approach more fully.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

This work was supported by a Natural Sciences and Engineering Research Council of Canada Discovery Grant. I thank David Brainard, Alan Gilchrist, and Larry Maloney for helpful comments on the manuscript.

LITERATURE CITED

Adams WJ, Elder JH, Graf EW, Leyland J, Lugtigheid AJ, Muryy A. 2016. The Southampton-York natural scenes (SYNS) dataset: statistics of surface attitude. *Sci. Rep.* 6:35805

Adelson EH. 1993. Perceptual organization and the judgment of brightness. *Science* 262:2042–44

Adelson EH. 2000. Lightness perception and lightness illusions. In *The New Cognitive Neurosciences*, ed. M Gazzaniga, pp. 339–51. Cambridge, MA: MIT Press

Outlines creative and influential ideas on lightness perception, several of which have still not been thoroughly explored.

Adelson EH, Pentland AP. 1996. The perception of shading and reflectance. In *Perception as Bayesian Inference*, ed. D Knill, W Richards, pp. 409–12. Cambridge, UK: Cambridge Univ. Press

Allred SR, Brainard DH. 2013. A Bayesian model of lightness perception that incorporates spatial variation in the illumination. *J. Vis.* 13(7):18

Allred SR, Radonjic A, Gilchrist AL, Brainard DH. 2012. Lightness perception in high dynamic range images: local and remote luminance effects. *J. Vis.* 12(2):7

Arend LE. 1993. Mesopic lightness, brightness, and brightness contrast. *Percept. Psychophys.* 54:469–76

Arend LE, Spehar B. 1993a. Lightness, brightness, and brightness contrast: 1. Illuminance variation. *Percept. Psychophys.* 54:446–56

Arend LE, Spehar B. 1993b. Lightness, brightness, and brightness contrast: 2. Reflectance variation. *Percept. Psychophys.* 54:457–68

Barlow HB. 1961. Possible principles underlying the transformations of sensory messages. In *Sensory Communication*, ed. WA Rosenblith, pp. 216–34. Cambridge, MA: MIT Press

Barron JT, Malik J. 2015. Shape, illumination, and reflectance from shading. *IEEE Trans. Pattern Anal. Mach. Intel.* 37:1670–87

Barrow HG, Tenenbaum JM. 1978. Recovering intrinsic scene characteristics from images. In *Computer Vision Systems*, ed. A Hanson, E Riseman, pp. 3–26. New York: Academic

Belhumeur PN, Kriegman DJ, Yuille AL. 1999. The bas-relief ambiguity. *Int. J. Comput. Vis.* 35:33–44

Betz T, Shapley R, Wichmann FA, Maertens M. 2015. Noise masking of White's illusion exposes the weakness of current spatial filtering models of lightness perception. *J. Vis.* 15(14):1

Blakeslee B, McCourt ME. 1999. A multiscale spatial filtering account of the White effect, simultaneous brightness contrast and grating induction. *Vis. Res.* 39:4361–77

Presents ODOG, a highly successful spatial filtering model of brightness perception.

Blakeslee B, McCourt ME. 2012. When is spatial filtering enough? Investigation of brightness and lightness perception in stimuli containing a visible illumination component. *Vis. Res.* 60:40–50

Blakeslee B, Reetz D, McCourt ME. 2008. Coming to terms with lightness and brightness: effects of stimulus configuration and instructions on brightness and lightness judgments. *J. Vis.* 8(11):3

Bloj M, Ripamonti C, Mitha K, Hauck R, Greenwald S, Brainard DH. 2004. An equivalent illuminant model for the effect of surface slant on perceived lightness. *J. Vis.* 4(9):6

Bonato F, Gilchrist AL. 1994. The perception of luminosity on different backgrounds and in different illuminations. *Perception* 23:991–1006

Boyaci H, Maloney LT, Hersh S. 2003. The effect of perceived surface orientation on perceived surface albedo in binocularly viewed scenes. *J. Vis.* 3:541–53

Brainard DH, Longere P, Delahunt PB, Freeman WT, Kraft JM, Xiao B. 2006. Bayesian model of human color constancy. *J. Vis.* 6(11):10

Brainard DH, Maloney LT. 2011. Surface color perception and equivalent illumination models. *J. Vis.* 11(5):1

A clear introduction to equivalent illumination models of color and lightness.

Brascamp JW, Shevell SK. 2021. The certainty of ambiguity in visual neural representations. *Annu. Rev. Vis. Sci.* 7:465–86

Brunswik E, Kamiya J. 1953. Ecological cue-validity of "proximity" and of other gestalt factors. *Am. J. Psychol.* 66:20–32

Carney T, Klein SA, Tyler CW, Silverstein AD, Beutter B, et al. 1999. Development of an image/threshold database for designing and testing human vision models. In *Proceedings of SPIE*, Vol. 3644, *Human Vision and Electronic Imaging IV*, ed. BE Rogowitz, TN Pappas, pp. 542–51. Bellingham, WA: SPIE

Chung ST, Legge GE, Tjan BS. 2002. Spatial-frequency characteristics of letter identification in central and peripheral vision. *Vis. Res.* 42:2137–52

Corney D, Lotto RB. 2007. What are lightness illusions and why do we see them? *PLOS Comput. Biol.* 3:e180

Cornsweet TN. 1970. *Visual Perception*. Fort Worth, TX: Harcourt Coll.

Dakin SC, Bex PJ. 2003. Natural image statistics mediate brightness "filling in". *Proc. R. Soc. B* 270:2341–48

Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L. 2009. ImageNet: a large-scale hierarchical image database. In *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–55. Piscataway, NJ: IEEE

DeValois RL, DeValois KK. 1988. *Spatial Vision*. Oxford, UK: Oxford Univ. Press

Doerschner K, Boyaci H, Maloney L. 2007. Testing limits on matte surface color perception in three-dimensional scenes with complex light fields. *Vis. Res.* 47:3409–23

Economou E, Zdravkovic S, Gilchrist AL. 2007. Anchoring versus spatial filtering accounts of simultaneous lightness contrast. *J. Vis.* 7(12):2

A persuasive anchoring-theory account of simultaneous contrast, traditionally the domain of lateral inhibition models.

Field DJ. 1987. Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am. A* 4:2379–94

Geisler WS. 2008. Visual perception and the statistical properties of natural scenes. *Annu. Rev. Psychol.* 59:167–92

Gilchrist AL. 1977. Perceived lightness depends on perceived spatial arrangement. *Science* 195:185–87

Gilchrist AL. 2006. *Seeing Black and White*. Oxford, UK: Oxford Univ. Press

Gives an exceptionally thorough and informative review of the history of lightness research.

Gilchrist AL, Kossyfidis C, Bonato F, Agostini T, Cataliotti J, et al. 1999. An anchoring theory of lightness perception. *Psychol. Rev.* 106:795–834

Goodfellow I, Bengio Y, Courville A. 2016. *Deep Learning*. Cambridge, MA: MIT Press

Graham NVS. 1989. *Visual Pattern Analyzers*. Oxford, UK: Oxford Univ. Press

Grossberg S, Todorovic D. 1988. Neural dynamics of 1-d and 2-d brightness perception: a unified model of classical and recent phenomena. *Percept. Psychophys.* 43:241–77

Heinemann EG. 1955. Simultaneous brightness induction as a function of inducing- and test-field luminances. *J. Exp. Psychol.* 50:89–96

Hering E. 1964 (1905). *Outlines of a Theory of the Light Sense*. Cambridge, MA: Harvard Univ. Press

Hess C, Pretori H. 1970 (1894). Quantitative investigation of the lawfulness of simultaneous brightness contrast. *Percept. Motor Skills* 31:947–69

Ibn al-Haytham. 1989 (1083). *The Optics of Ibn al-Haytham*. London: Warburg Institute

Jacobsen A, Gilchrist A. 1988. The ratio principle holds over a million-to-one range of illumination. *Percept. Psychophys.* 43:1–6

Katz D. 1935. *The World of Colour*. London: Kegan Paul
Kim M, Wilcox LM, Murray RF. 2006. Perceived three-dimensional shape toggles perceived glow. *Curr. Biol.* 26(9):R350–51
Kingdom FA. 2008. Perceiving light versus material. *Vis. Res.* 48:2090–105
Kingdom FAA. 2011. Lightness, brightness and transparency: a quarter century of new ideas, captivating demonstrations and unrelenting controversy. *Vis. Res.* 51:652–73
Knill DC, Richards W, eds. 1996. *Perception as Bayesian Inference*. Cambridge, UK: Cambridge Univ. Press
Kobyzev I, Prince S, Brubaker M. 2020. Normalizing flows: an introduction and review of current methods. *IEEE Trans. Pattern Anal. Mach. Intel.* In press
Koffka K. 1935. *Principles of Gestalt Psychology*. New York: Harcourt Brace World
Koller D, Friedman N. 2009. *Probabilistic Graphical Models: Principles and Techniques*. Cambridge, MA: MIT Press

A detailed introduction to probabilistic graphical models.

Kuffler SW. 1953. Discharge patterns and functional organization of the mammalian retina. *J. Neurophysiol.* 16:37–68
Land EH, McCann JJ. 1971. Lightness and retinex theory. *J. Opt. Soc. Am.* 61:1–11
Lee AB, Mumford D, Huang J. 2001. Occlusion models for natural images: a statistical study of a scale-invariant dead leaves model. *Int. J. Comput. Vis.* 41:35–59
Logvinenko AD. 2015. The achromatic object-color manifold is three-dimensional. *Perception* 44:243–68
Logvinenko AD, Adelson EH, Ross DA, Somers D. 2005. Straightness as a cue for luminance edge interpretation. *Percept. Psychophys.* 67:120–28
Logvinenko AD, Maloney LT. 2006. The proximity structure of achromatic surface colors and the impossibility of asymmetric lightness matching. *Percept. Psychophys.* 68:76–83

The starting point for multidimensional scaling studies of achromatic color appearance.

Logvinenko AD, Petrini K, Maloney LT. 2008. A scaling analysis of the snake lightness illusion. *Percept. Psychophys.* 70:828–40
Madigan SC, Brainard DH. 2014. Scaling measurements of the effect of surface slant on perceived lightness. *i-Perception* 5:53–72
McCluney R. 1994. *Introduction to Radiometry and Photometry*. Norwood, MA: Artech House
Morgenstern Y, Geisler WS, Murray RF. 2014. Human vision is attuned to the diffuseness of natural light. *J. Vis.* 14(9):15
Murdoch I. 1973. *The Black Prince*. London: Chatto & Windus
Murray RF. 2013. Human lightness perception is guided by simple assumptions about reflectance and lighting. In *Proceedings of SPIE*, Vol. 8651, *Human Vision and Electronic Imaging XVIII*, ed. BE Rogowitz, TN Pappas, H de Ridder, art. 865106. Bellingham, WA: SPIE
Murray RF. 2020. A model of lightness perception guided by probabilistic assumptions about lighting and reflectance. *J. Vis.* 20(7):28
Ostrovsky Y, Cavanagh P, Sinha P. 2005. Perceiving illumination inconsistencies in scenes. *Perception* 34:1301–14
Patel KY, Munasinghe AP, Murray RF. 2018. Lightness matching and perceptual similarity. *J. Vis.* 18(5):1
Perdreau F, Cavanagh P. 2011. Do artists see their retinas? *Front. Hum. Neurosci.* 5:171
Purves D, Lotto RB, eds. 2010. *Why We See What We Do Redux: A Wholly Empirical Theory of Vision*. Sunderland, MA: Sinauer Assoc.
Radonjic A, Pearce B, Aston S, Krieger A, Dubin H, et al. 2016. Illumination discrimination in real and simulated scenes. *J. Vis.* 16(11):2
Ramamoorthi R, Hanrahan P. 2001. A signal-processing framework for inverse rendering. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 117–28. New York: ACM
Robinson AE, Hammon PS, de Sa VR. 2007. Explaining brightness illusions using spatial filtering and local response normalization. *Vis. Res.* 47:1631–44
Roth S, Black M. 2005. Fields of experts: a framework for learning image priors. In *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, pp. 860–67. Piscataway, NJ: IEEE
Rutherford MD, Brainard DH. 2002. Lightness constancy: a direct test of the illumination-estimation hypothesis. *Psychol. Sci.* 13:142–49

Schirillo JA. 2013. We infer light in space. *Psychon. Bull. Rev.* 20:905–15
Shapiro A, Lu ZL. 2011. Relative brightness in images can be accounted for by removing blurry content. *Psychol. Sci.* 22:1452–59
Tolstoy L. 2006 (1874). *Anna Karenina*. London: Penguin Classics
von Helmholtz HLF. 1924 (1910). *Treatise on Physiological Optics*. Rochester, NY: Opt. Soc. Am.
Wallach H. 1948. Brightness constancy and the nature of achromatic colors. *J. Exp. Psychol.* 38:310–24
White M. 1979. A new effect of pattern on perceived lightness. *Perception* 8:413–16
Wilder JD, Adams WJ, Murray RF. 2019. Shape from shading under inconsistent illumination. *J. Vis.* 19(6):2
Yamins DLK, Hong H, Cadieu CF, Solomon EA, Seibert D, DiCarlo JJ. 2014. Performance-optimized hierarchical models predict neural responses in higher visual cortex. *PNAS* 111:8619–24
Yu Y, Smith WAP. 2019. InverseRenderNet: learning single image inverse rendering. In *Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3150–59. Piscataway, NJ: IEEE
Zeiner K, Maertens M. 2014. Linking luminance and lightness by global contrast normalization. *J. Vis.* 14(7):3

Annual Review of Vision Science

Volume 7, 2021

Contents

Errata

An online log of corrections to *Annual Review of Vision Science* articles may be found at http://www.annualreviews.org/errata/vision