Adelson, E.H. Lightness Perception and Lightness Illusions. In *The New Cognitive Neurosciences*, 2nd ed., M. Gazzaniga, ed. Cambridge, MA: MIT Press, pp. 339-351, (2000).

24 Lightness Perception and Lightness Illusions

EDWARD H. ADELSON

ABSTRACT: A gray surface in sunlight may have much higher luminance than it has in the shade, but it still looks gray. To achieve the task of "lightness constancy," the visual system must discount the illumination and other viewing conditions and estimate the reflectance. Many different physical situations, such as shadows, filters, or haze, can be combined to form a single, simple mapping from luminance to reflectance. The net effect of the viewing conditions, including additive and multiplicative effects, may be termed an "atmosphere." An "atmospheric transfer function" maps reflectance into luminance. To correctly estimate lightness, a visual system must determine a "lightness transfer function" that performs the inverse. Human lightness computation is imperfect, but performs well in most natural situations. Lightness illusions can reveal the inner workings of the estimation process, which may involve low-level, mid-level, and high-level mechanisms. Mid-level mechanisms, involving contours, junctions, and grouping, appear to be critical in explaining many lightness phenomena.

Every light is a shade, compared to the higher lights, till you come to the sun; and every shade is a light, compared to the deeper shades, till you come to the night.
—John Ruskin (1879).

The amount of light coming to the eye from an object depends on the amount of light striking the surface, and on the proportion of light that is reflected. If a visual system only made a single measurement of luminance, acting as a photometer, then there would be no way to distinguish a white surface in dim light from a black surface in bright light. Yet humans can usually do so, and this skill is known as *lightness constancy*.

The constancies are central to perception. An organism needs to know about meaningful world-properties, such as color, size, shape, etc. These properties are not explicitly available in the retinal image, and must be extracted by visual processing. The gray shade of a surface is one such property. To extract it, luminance information must be combined across space. Figure 24.1 shows the well-known *simultaneous contrast* effect, which demonstrates a spatial interaction in lightness perception. The two smaller squares are the same shade of gray. However, the square in the dark surround appears lighter than the square in the light surround. Illusions like these are sometimes viewed as quirky failures of perception, but they help reveal the inner workings of a system that functions remarkably well. Here we will consider how lightness illusions can inform us about lightness perception.

EDWARD H. ADELSON Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Mass.

Levels of processing

The visual system processes information at many levels of sophistication. At the retina, there is *low-level vision*, including light adaptation and the center-surround receptive fields of ganglion cells. At the other extreme is *high-level vision*, which includes cognitive processes that incorporate knowledge about objects, materials, and scenes. In between there is *mid-level vision*. Mid-level vision is simply an ill-defined region between low and high. The representations and the processing in the middle stages are commonly thought to involve surfaces, contours, grouping, and so on. Lightness perception seems to involve all three levels of processing.

The low-level approach to lightness is associated with Ewald Hering. He considered adaptation and local interactions, at a physiological level, as the crucial mechanisms. This approach has long enjoyed popularity because it offers an attractive connection between physiology and psychophysics. Figure 24.2(a) shows the receptive field of an idealized center-surround cell. The cell exhibits lateral inhibition: light in the center is excitatory while light in the surround is inhibitory. A cross-section of the receptive field is shown in figure 24.2(b). This cell performs a local comparison between a given luminance and the neighboring luminances, and thus offers machinery that can help explain the simultaneous contrast (SC) illusion. This idea was formalized by Mach, who proposed a Laplacian derivative operator as the mechanism.

One of Mach's inspirations was an illusion now known as the Mach band. When a spatial ramp in luminance abruptly changes slope, an illusory light or dark band appears. A variant of the Mach band has been used by the op artist Vasarely, as shown in figure 24.3(a). The image consists of a set of

FIGURE 24.1 The simultaneous contrast effect.

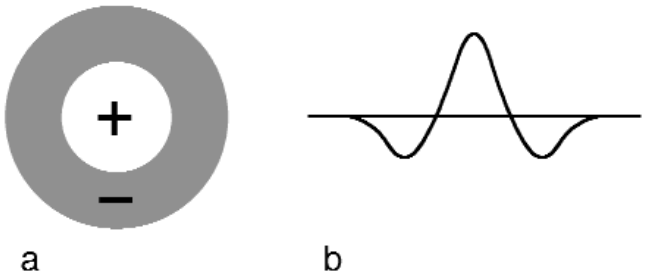

FIGURE 24.2 Center-surround inhibition.

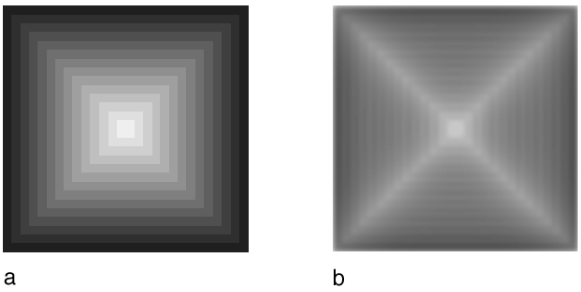

FIGURE 24.3 An illusion by Vasarely (a) and a bandpass filtered version (b)

nested squares. Each square is a constant luminance. The pattern gives the illusion of a glowing X along the diagonals, even though the corners of the squares are no brighter than the straight parts. When a center-surround filter is run over this pattern (i.e., is convolved with it) it produces the image shown in figure 24.3(b). The filter output makes the bright diagonals explicit.

A center-surround filter cannot explain a percept by itself: perception involves the whole brain. However, it is interesting that center-surround responses can go a long way to explaining certain illusions.

Derivative operators respond especially well to sharp intensity transitions such as edges. The importance of edges, and the lesser importance of slow gradients, is indicated by the *Craik-O'Brien-Cornsweet effect* (COCE) named after its several discoverers. Figure 24.4 shows one of several COCE variants. The figure appears to contain a dark square next to

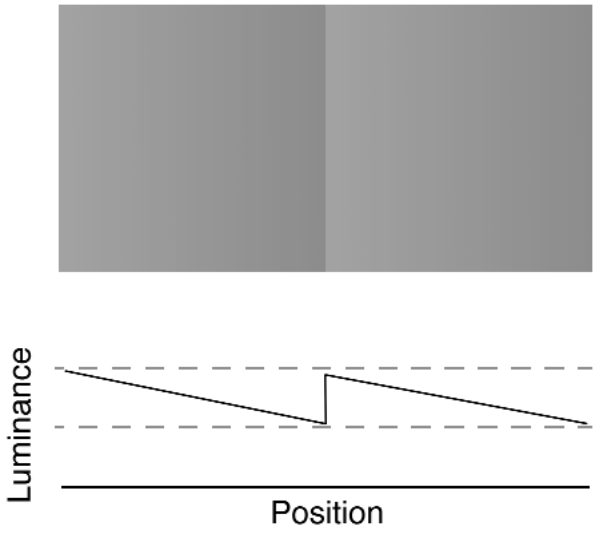

FIGURE 24.4 One version of the Craik-O'Brien-Cornsweet Effect

a light square. Actually, the two squares are ramps, and they are identical, as shown by the luminance profile underneath (the dashed lines show constant luminances). The response of a center-surround cell to this pattern will be almost the same as its response to a true step edge: there will be a big response at the edge, and a small response elsewhere. While this doesn't explain why the image looks as it does, it may help explain why one image looks similar to the other (Cornsweet, 1970).

Center-surround processing is presumably in place for a good reason. Land and McCann (1971) developed a model they called Retinex, which placed the processing in a meaningful computational context.

Land and McCann began by considering the nature of scenes and images. They argued that reflectance tends to be constant across space except for abrupt changes at the transitions between objects or pigments. Thus a reflectance change shows itself as step edge in an image, while illuminance will change only gradually over space. By this argument one can separate reflectance change from illuminance change by taking spatial derivatives: high derivatives are due to reflectance and low ones are due to illuminance.

The Retinex model applies a derivative operator to the image, and thresholds the output to remove illuminance variation. The algorithm then reintegrates edge information over space to reconstruct the reflectance image.

The Retinex model works well for stimuli that satisfy its assumptions, including the Craik-O'Brien-Cornsweet display, and the "Mondrians" that Land and McCann used. A Mondrian (so-called because of its loose resemblance to paintings by the artist Mondrian) is an array of randomly colored, randomly placed rectangles covering a plane surface, and illuminated non-uniformly.

The real world is more complex than the Mondrian world, of course, and the Retinex model has its limits. In its original form it cannot handle the configural effects to be described later in this chapter. However, the Land-McCann research program articulated some important principles. Vision is only possible because there are constraints in the world, i.e., images are not formed by arbitrary random processes. To function in this world, the visual system must exploit the ecology of images—it must "know" the likelihood of various things in the world, and the likelihood that a given image-property could be caused by one or another world-property. This world-knowledge may be hard-wired or learned, and may manifest itself at various levels of processing.

Limits on low-level processes

The high-level approach is historically associated with Helmholtz, who argued that perception is the product of *unconscious inference*. His dictum was this: what we perceive is our visual system's best guess as to what is in the world. The guess is based on the raw image data plus our prior experience. In the Helmholtz view, lightness constancy is achieved by inferring, and discounting, the illuminant. From this standpoint the details of low-level processing are not the issue. A lightness judgment involves the workings of the whole visual system, and that system is designed to interpret natural scenes. Simultaneous contrast and other illusions are the byproduct of such processing.

Hochberg and Beck (1954), and Gilchrist (1977), showed that 3D cues could greatly change the lightness perception in a scene, even when the retinal image remained essentially unchanged, in accord with Helmholtz's approach.

The importance of scene interpretation is also shown by a recent variant on the Craik-O'Brien-Cornsweet effect, devised by Knill and Kersten (1991). In figure 24.5(a), one sees two identical cylinders. In figure 24.5(b) one sees a brick painted with two shades of paint. Embedded within each image is a COCE pattern. The two ramps are interpreted as shading in figure 24.5(a), but as paint in figure 24.5(b).

The Gestalt approach

The Gestalt psychologists approached lightness perception, and perception generally, in a different manner than the Hering or the Helmholtz schools. They emphasized the importance of perceptual organization, much of it based on mechanisms that might be characterized as mid-level. The key concepts include grouping, belongingness, good continuation, proximity, and so on.

Koffka offered an example of how simultaneous contrast

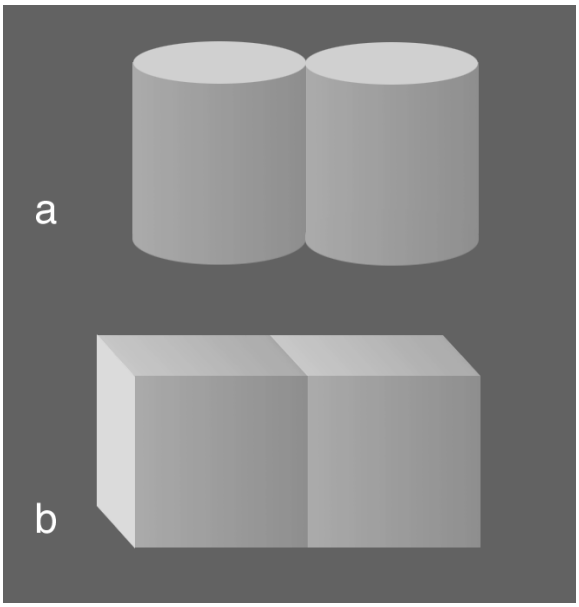

FIGURE 24.5 Knill and Kersten's illusion. Both figures contain the same COCE ramps, but the interpretations are quite different.

can be manipulated by changing spatial configuration. The ring in figure 24.6(a) appears almost uniform in color. When the stimulus is split in two, as shown in figure 24.6(b), the two half-rings appear to be different shades of gray. The two halves now have separate identities, and each is perceived within its own context.

An interesting variant that involves transparency is shown in figure 24.6(c). The left and right half-blocks are slid vertically, and the new configuration leads to a very different perceptual organization and a strong lightness illusion. We will return to this stimulus in our later discussion.

Some terminology

Having outlined some basic phenomena, we now return to the basic problems. First, we will clarify some terminology. More complete definitions can be found in books on photometry and colorimetry.

Luminance is the amount of visible light that comes to the eye from a surface.

Illuminance is the amount of light incident on a surface.

Reflectance is the proportion of incident light that is reflected from a surface.

Reflectance, also called *albedo*, varies from 0 to 1, or equivalently from 0% to 100%. 0% is ideal black; 100% is ideal white. In practice, typical black paint is about 5% and typical white paint about 85%. (To keep things simple, we only consider matte surfaces, for which a single reflectance value

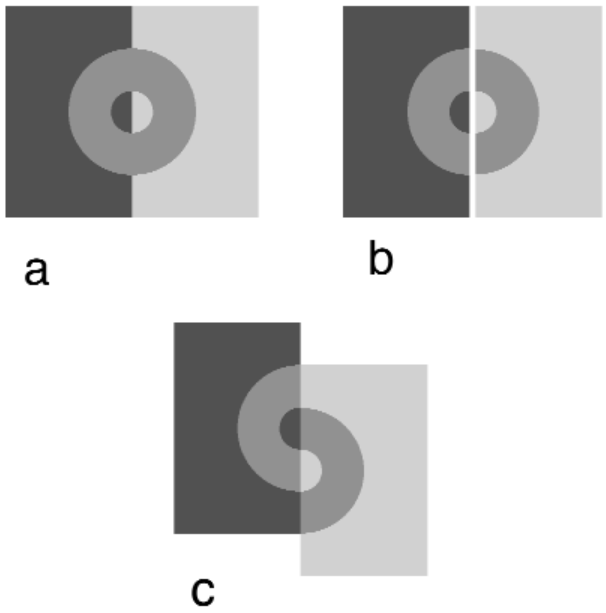

FIGURE 24.6 Variants on the Koffka ring. (a) The ring appears about uniform. (b) When split, the two half-rings appear distinctly different. (c) When shifted, the two half-rings appear quite different.

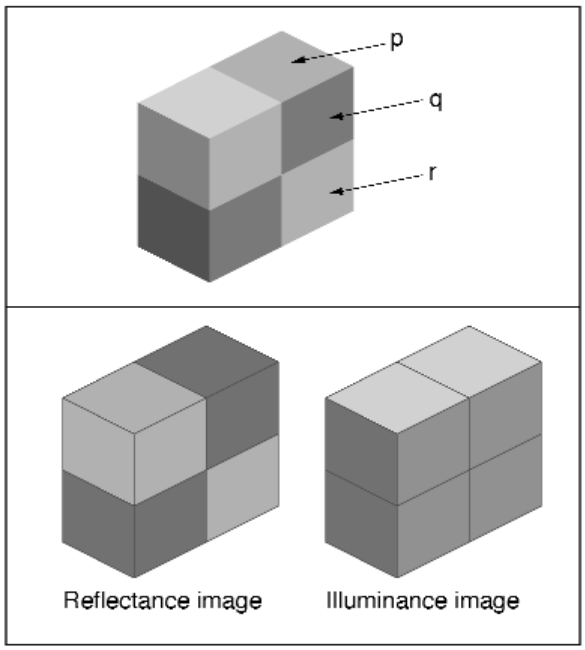

FIGURE 24.7 The "checker-block" and its analysis into two intrinsic images.

offers a complete description.)

Luminance, illuminance, and reflectance, are physical quantities that can be measured by physical devices. There are also two subjective variables that must be discussed.

Lightness is defined as the perceived reflectance of a surface. It represents the visual system's attempt to extract reflectance based on the luminances in the scene.

Brightness is defined as the perceived intensity of light coming from the image itself, rather than any property of the portrayed scene. Brightness is sometimes defined as perceived luminance.

These terms may be understood by reference to figure 24.7. The block is made of a 2x2 set of cubes, each colored either light or dark gray. We call this the "checker-block." Illumination comes from an oblique angle, lighting different faces differently. The luminance image can be considered to be the product of two other images: the reflectance image and the illuminance image, shown below. These underlying images are termed *intrinsic images* in machine vision (Barrow and Tenenbaum, 1978). Intrinsic image decompositions have been proposed for understanding lightness perception (Arend, 1994; Adelson and Pentland, 1996)

Patches *p* and *q* have the same reflectance, but different luminances. Patches *q* and *r* have different reflectances and different luminances; they share the same illuminance. Patches *p* and *r* happen to have the same luminance, because the lower reflectance of *p* is counterbalanced by its higher illuminance.

Faces *p* and *q* appear to be painted with the same gray, and thus they have the same lightness. However, it is clear that *p* has more luminance than *q* in the image, and so the patches differ in brightness. Patches *p* and *r* differ in both lightness and brightness.

The problem of lightness constancy

From a physical point of view, the problem of lightness constancy is as follows. An illuminance image, *E(x,y)*, and a reflectance image, *R(x,y)*, are multiplied to produce a luminance image, *L(x,y)*:

$$L(x,y) = E(x,y)R(x,y).$$

An observer is given *L* at each pixel, and attempts to determine the two numbers *E* and *R* that were multiplied to make it. Unfortunately, unmultiplying two numbers is impossible. If *E(x,y)* and *R(x,y)* are arbitrary functions, then for any *E(x,y)* there exists an *R(x,y)* that produces the observed image. The problem appears impossible, but humans do it pretty well. This must mean that illuminance and reflectance images are not arbitrary functions. They are constrained by statistical properties of the world, as proposed by Land and McCann.

Note that Land and McCann's constraints fail when applied to the checker-block image. Figure 24.8(a) shows two light-dark edges. They are exactly the same in the

image, and any local edge detector or filter will respond to them in the same way. Retinex will classify both as reflectance steps. Yet they have very different meanings. One is caused by illuminance (due to a change in surface normal); the other is caused by reflectance.

To interpret the edges, the visual system must consider them in a larger context. One good source of information is the junctions, such as those labeled in figure 24.8(b). A junction is a place where two or more contours come together. X, Y, L, T, and ψ, as shown, are some of the simple junction types. The configuration of a junction, as well as the gray levels forming the junction, can offer cues about the shading and reflectance of a surface.

Particularly strong constraints are imposed by a ψ-junction, like that in figure 24.8(b). The vertical spine appears to be a dihedral with different illuminance on the two sides. The angled arms appear to represent a reflectance edge that crosses the dihedral. The ratios of the gray levels, and the angles of the arms, are consistent with this interpretation.

The influence of a ψ-junction can propagate along the contours that meet at the ψ. A single light-dark edge, ambiguous by itself, can be pushed toward a particular interpretation by adjoining ψ's (Sinha and Adelson, 1993).

In figure 24.9, the dashed rectangle encloses a set of horizontal light and dark stripes. If one only considers the region within the dashed rectangle, it is impossible to determine the physical sources of the stripes. However, if one covers the right side of the figure and views the left side, it appears that the stripes are due to paint. If one covers the left side and views the right, it appears that the stripes are due the different lighting on the stairsteps. If one views both sides, the percept flip-flops according to where one looks.

The ψ-junctions seem to be in control here. If one follows a stripe to the left, it connects to a y with a vertical spine, and becomes an arm of that ψ. The junction configuration, along with the junction gray levels, suggest that the stripe is due to reflectance. When the same strip is followed to the right, it joins a ψ with a horizontal stem. Again, the configuration and gray levels suggest that illuminance is the cause.

Configurations involving ψ's can modulate brightness illusions. Figure 24.10 shows a stimulus we call the *corrugated plaid* (Adelson, 1993). In figure 24.10(a) the two marked patches are the same shade of gray. The upper patch appears slightly darker. Figure 24.10(b) shows another array of gray patches that have the same gray levels at the same positions as in figure 24.10(a), i.e., the same raster sequence of grays. Only the geometry has been changed, parallelograms having been substituted for squares and vice-versa. The illusion is much enhanced, the upper patch appearing much darker than the lower one. In the laboratory the apparent luminance difference is increased threefold.

A low-level filtering mechanism, or a mechanism based on local edge interactions, cannot explain the change in the illusion. We proposed (Adelson, 1993) a Helmholtzian explanation based on intrinsic images: the change in ψ-junctions causes a change in the perception of 3D surface orientation and shading. In figure 24.10(a) the two test patches appear to be in the same illumination, but in figure 24.10(b)

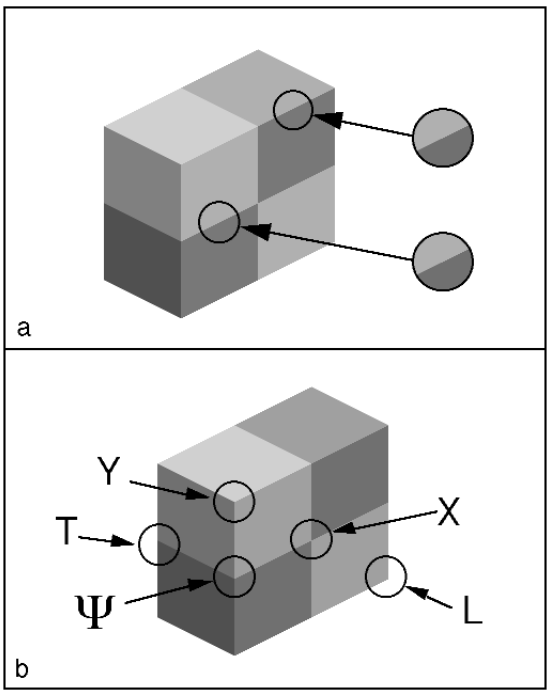

FIGURE 24.8 (a) The local ambiguity of edges. (b) A variety of junctions.

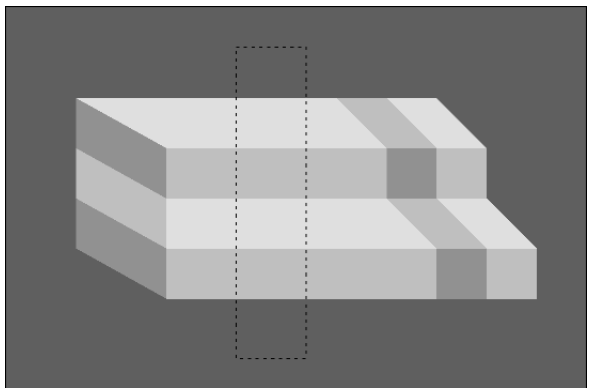

FIGURE 24.9 The impossible steps. On the left, the horizontal stripes appear to be due to paint; on the right, they appear to be due to shading.

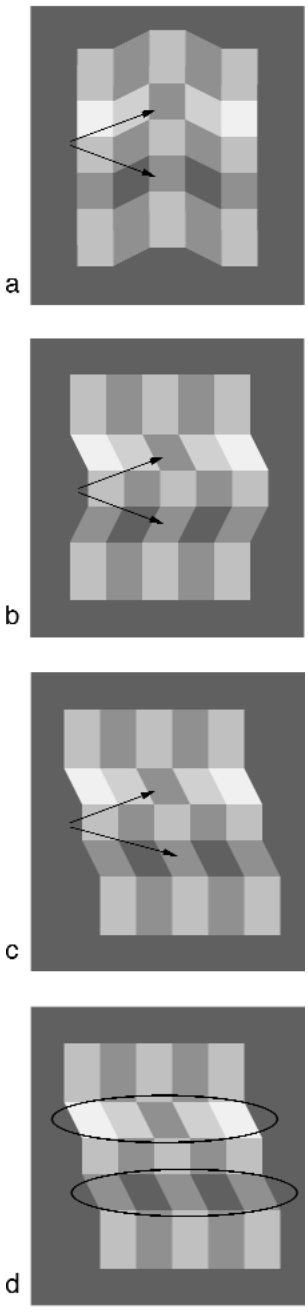

FIGURE 24.10 Variations on the corrugated plaid. (a) The two patches appear nearly the same. (b) The patches appear quite different. (c) The patches appear quite different, but there is no plausible shaded model. (d) Possible grouping induced by junctions.

they are differently illuminated. A brightly lit patch of dark gray looks quite different from a dimly lit patch of light gray. This lightness computation could have a strong influence on brightness judgments.

Thus a 3D shaded model can help explain the phenomenon, but is it necessary? Todorovic has devised a variant, shown in figure 24.10(c), that suggests not. The figure was made by mirror reversing the bottom two rows. The illusion remains strong—nearly as strong as figure 24.10(b)—for many subjects. However, there is no reasonable interpretation in terms of a 3D shaded model. The two strips containing the test patches appear to lie in parallel planes, and so they should be receiving similar illumination.

Perhaps the intrinsic image story can be saved by appealing to the notion of local consistency without global consistency, such as occurs in figure 24.9. However, it may be that the main effects are the result of simpler 2D computations. The ψ-junctions, taken as 2D configurations, could be used as grouping cues that define the context in which lightness is assessed, as indicated in figure 24.10(d). If this is correct, then the Helmholtzian approach is overkill.

A number of investigators have lately argued for models based on Gestalt-like grouping mechanisms (e.g., Ross and Pessoa, in press). Gilchrist, who in earlier years took a Helmholtzian stance (Gilchrist et al, 1983), has recently proposed a model of lightness that emphasizes 2D configuration and grouping mechanisms (Gilchrist et al., in press).

Anchoring and frameworks

Gilchrist's new model took shape in the course of his investigations into *anchoring*. The anchoring problem is this. Suppose an observer determines that patch *x* has four times the reflectance of patch *y*. This solves part of the lightness problem, but not all of it: the absolute reflectances remain unknown. An 80% near-white is four times a 20% gray, but a 20% gray is also four times a 5% black. For absolute judgments one must tie down the gray scale with an anchor, i.e., a luminance that is mapped to a standard reflectance such as midgray or white.

Land and McCann had encountered this problem with Retinex, and they proposed that the highest luminance should be anchored to white. All other grays could then be scaled relative to that white. This is known as *the highest luminance rule*.

Li and Gilchrist (in press) tested the highest-luminance rule using bipartite ganzfelds. They painted the inside of a large hemispherical dome with two shades of gray paint. When subjects put their heads inside, their entire visual fields were filled with only two luminances. A bipartite field painted black and gray appeared to be a bipartite field painted gray and white, as predicted by the highest luminance rule.

By manipulating the relative areas of the light and dark fields, Gilchrist and Cataliotti (1994) found evidence for a second, competing anchoring rule: the largest area tends to appear white. They argue that the actual anchor is a compromise between these rules.

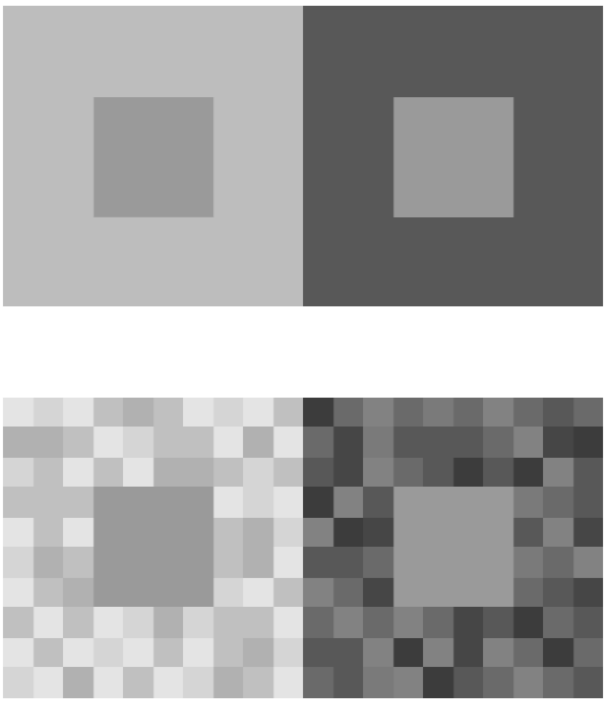

FIGURE 24.11 Simultaneous contrast is enhanced with articulated surrounds, as shown below.

Gilchrist also emphasizes the importance of *articulation* and *insulation* in anchoring. Articulation is a term used earlier by Katz (1935); it refers to the number of distinct surfaces or patches within a region. Katz observed that greater articulation leads to better lightness constancy, and Gilchrist proposes that it leads to better local anchoring. We can demonstrate the effect of articulation with a simultaneous contrast display, as in figure 24.11. Above is a standard display. Below is an articulated version. The surround mean luminances have not been changed, but the surrounds have been broken into many squares. The articulated version gives a stronger illusion. In our laboratory we find that the strength of the illusory contrast can be doubled. (As with all the demonstrations in this chapter, the effect may be weaker on the printed page due to the small image size and limitations in the printing process).

In Gilchrist's model, anchoring occurs within a *framework* which is a region containing patches that are grouped. Frameworks can be local or global. In figure 24.11, a local framework would be the patches surrounding the test square, and the global framework would be the entire page, and even the room in which the page is viewed.

If a local framework is well insulated, it has strong control over the anchoring. Insulation occurs when the local framework is strongly grouped as a separate entity from the global framework.

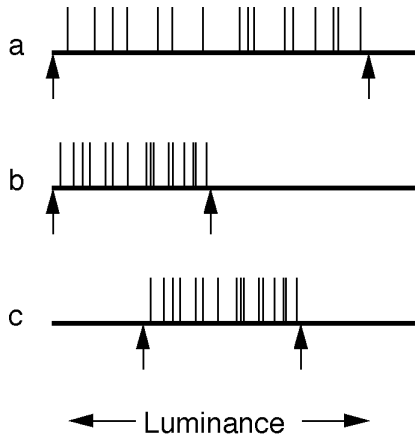

FIGURE 24.12 A collection of random gray surfaces will lead to a different luminance distribution in different viewing conditions.

Statistical Estimation

The various lightness principles might be thought of as heuristics that the visual system has arbitrarily adopted. These principles begin to make sense, however, if we consider the lightness problem from the standpoint of statistical estimation.

Suppose that the world consisted of a set of gray patches randomly drawn from some distribution. Then, under a given illuminance, one would observe a distribution of luminance samples such as that shown in figure 24.12(a). If the illumination were dimmed by half, then the luminances would follow suit, as shown in figure 24.12(b). The arrows bracketing the distributions represent the extremes of 0% and 100% reflectance, i.e., the luminances mapping to ideal black and ideal white.

The observed luminances can also be changed by an additive haze or glare, which slides the distribution upward, illustrated in figure 24.12(c). Again one can estimate which luminance corresponds to which reflectance, i.e., one can estimate the mapping between the observed luminance and the underlying reflectance. We use the term atmosphere to refer to the combined effects of a multiplicative process (e.g., illuminance) and an additive process (e.g., haze).

If one has prior knowledge about distributions of reflectances and atmospheres, then one can construct optimal estimates of the locations of various reflectances along the luminance axis. That is, one can estimate the mapping between luminance and reflectance, as is required for lightness constancy.

Estimating this mapping is a central task of lightness per-

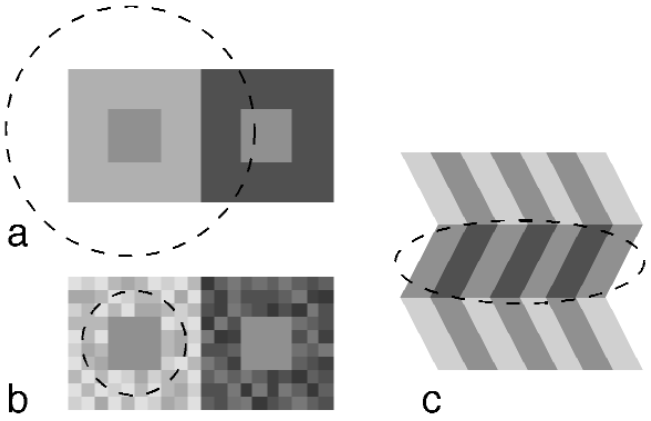

FIGURE 24.13 Lightness computations may employ adaptive windows.

ception. The image luminance is given and the perceived reflectance (lightness) must be derived. Anchoring is a way of describing part of this process. We will return to this problem when we discuss lightness transfer functions.

Adaptive windows

A larger number of samples will lead to better estimates of the lightness mapping. To increase *N*, the visual system can gather samples from a larger window. However, the atmosphere can vary from place to place, so there is a counterargument favoring small windows.

Suppose that the visual system uses an adaptive window to deal with this tradeoff. The window grows when there are too few samples, and shrinks when there are more than enough. Consider the examples shown in figure 24.13. In the classical SC display, figure 24.13(a), there are only a few large patches, so the window will tend to grow. In the articulated SC display, figure 24.13(b), the window can remain fairly small.

Lightness estimates are computed based on the statistics within the adaptive window. In the classic SC display, the window becomes so large that the statistics surrounding either of the test patches are rather similar. (In Gilchrist's terminology, the global framework dominates). In the articulated display, the windows can be small, so that they will not mix statistics from different atmospheres. This predicts the enhancement in the illusion.

It is reasonable to assume that the statistical window has soft edges. For example, it could be a 2D Gaussian hump centered at the location of interest. Since nearby patches are likely to share the same atmosphere, proximity should lead to high weights, with more distant patches getting lower weights (cf. Reid and Shapley 1988, Spehar et al., 1996). The dashed lines in figure 24.13 would indicate a level line of the Gaussian hump.

A further advantage occurs if the adaptive window can change shape. For example, in figure 24.13(c), it would be prudent to keep the statistical pooling within the horizontal region shown by the ellipse. This will avoid mixing luminances from the adjacent regions, which are in different lighting.

This reasoning might explain why ψ-junctions are effective at insulating one region from another. A set of ψ's along a contour (and with the appropriate gray levels) gives a strong cue that the contour is an atmospheric boundary. The statistical window should avoid crossing such a boundary in order to avoid mixing distributions. Thus the window should configure itself into a shape like that in figure 24.13(c).

Atmospheres

As noted above, illuminance is only one of the factors determining the luminance corresponding to a given reflectance. Other factors could include interposed filters (e.g., sunglasses), scattering, glare from a specular surface such as a windshield, and so on. It turns out that most physical effects will lead to linear transforms. Therefore the combined effects can be captured by a single linear transform (characterized by two parameters). This is what we call an atmosphere.

The equation we use is,

$$L = m\,R + e,$$

where L and R are luminance and reflectance, m is a multiplier on the reflectance, and e is an additive source of light. The value of m is determined by the amount of light falling on the surface, as well as the proportion of light absorbed by the intervening media between the surface and the eye.

The equation here is closely related to the linear equation underlying Metelli's episcotister model (Metelli, 1974) for transparency, except that there is no necessary coupling between the additive and multiplicative terms. The parameters m and e are free to take on any positive values.

An atmosphere may be thought of as a single transparent layer, except that it allows a larger range of parameters. It can be amplifying rather than attenuating, and it can have an arbitrarily large additive component.

In our usage, "atmosphere" simply refers to the mapping, i.e., the mathematical properties established by the viewing conditions without regard to the underlying physical processes. Putting on sunglasses or dimming the lights has the same effect on the luminances, and so leads to the same effect on atmosphere. To be more explicit about this meaning, we define the Atmospheric Transfer Function, or ATF, as the mapping between reflectance and luminance.

Figure 24.14 shows a set of random vertical lines viewed

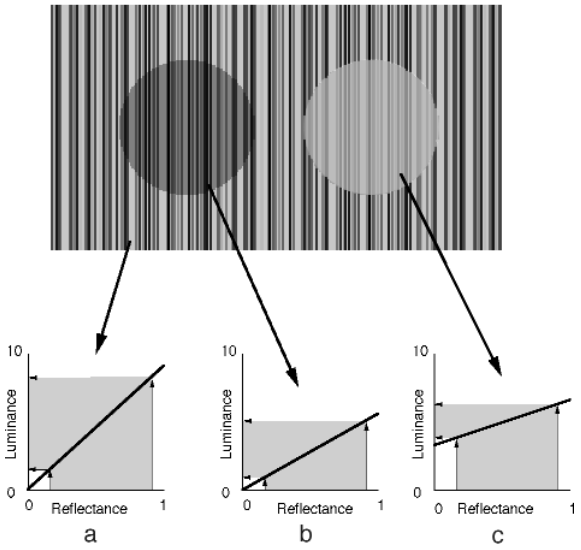

FIGURE 24.14 Lines of random gray, viewed under three different atmospheres. The ATF's, shown below, determine the mapping from reflectance to luminance.

in three different atmospheres. The large outer region is in some default atmosphere. The left disk is in an attenuating atmosphere (compared to the default). The right disk is in a hazy atmosphere.

The ATF for the main atmosphere is shown in figure 24.14(a). It passes through the origin, meaning that *e* is zero. The slope is specified by *m*. (*Note*: Since reflectance and luminance are in different units, there is also a scale constant that depends on the chosen units.) The small arrows in the panels show how the various reflectances are mapped to their corresponding luminances. The shaded area within the arrows shows how a typical range of reflectances will be mapped into the corresponding range of luminances.

Figure 24.14(b) shows the ATF for the dimmer atmosphere. The slope is reduced, and the intercept remains zero. On the right, in figure 24.14(c), is the ATF of the hazy atmosphere. The output luminance range is compressed by *m* and shifted up by *e*.

Note that there is no such thing as a "non-atmosphere." An observer cannot see the reflectances "directly," but rather requires an atmospheric transfer function to convert reflectances to luminances. The parameters of the ATF always have values.

Finally, note that the (*m,e*) parameterization has no privileged status. Any two numbers will do. For example, a useful alternative would be the white-point and the black-point. Since the atmosphere maps a reflectance to a luminance, the observer must implicitly reverse the mapping, turning a luminance into a perceived reflectance, as illustrated in figure 24.15. The inverting function, for a given observer in a given condition, may be called the *lightness transfer function* or *LTF*. The LTF is subjective; it need not be linear and need not be the correct inverse of the ATF. For a given observer it must be determined empirically.

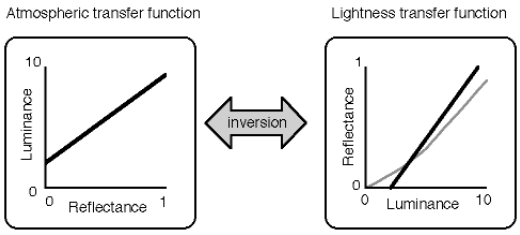

FIGURE 24.15 The inverse relation betweent he atmospheric transfer function and the ideal lightness transfer fuction.

Atmospheres and X-junctions

The connection between X-junctions and atmospheres is shown in figure 24.16. Different types of atmospheres lead to different ordinal categories of X-junctions (cf. Beck et al, 1984; Adelson and Anandan, 1990; Anderson, 1997).

Figure 24.16(a) shows a region with two shades of gray paint. The large light square has 75% reflectance and the small dark square in the corner has 25% reflectance. The two reflectances are marked with arrows on the abscissa of the corresponding ATF diagram, below. Figure 24.16(b) shows what happens when a new atmosphere is introduced in the central patch. The new ATF is shown in a dashed line in the ATF diagram; it might be produced by a dark filter or a shadow. The resulting luminances form an X-junction of the "sign-preserving" or "non-reversing" type (Adelson and Anandan, 1990), which is consistent with transparency.

Figure 24.16(c) shows a different category of *X*-junction: the single-reversing *X*. It gives the impression of a murky or hazy medium. For a single-reversing *X*, the new ATF must cross the original ATF at a point between the two reflectances. A crossover ATF can only arise from an additive process combined with an attenuative process, such as would occur with smoke or a dirty window. Another difference between single-reversing and sign-preserving *X*'s is that either edge of a sign-preserving *X* is potentially an atmospheric boundary, while only one edge of a single-reversing edge can be an atmospheric boundary. For this reason, the depth ordering of a single-reversing *X* is unambiguous.

Finally, figure 24.16(d) shows a double-reversing *X*, which does not look transparent. The ATF needed to produce this *X* would need a negative slope. This cannot occur in

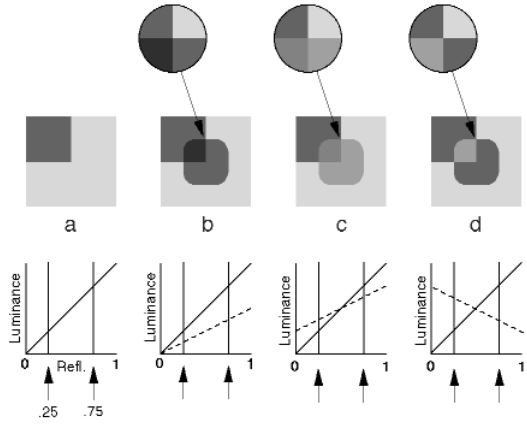

FIGURE 24.16 Transparency involves the imposition of a new atmosphere. The resulting X-junctions category depends on the atmospheric transfer function.

normal physical circumstances. Double-reversing *X*-junctions do not signal atmospheric boundaries to the visual system, and they typically look like paint rather than transparency. The junctions in a checkerboard are double-reversing *X*'s.

The ATF diagrams offer a simple graphical analysis of different *X*-junction types, and show how the *X*-junctions can be diagnostic of atmospheric boundaries.

Figure 24.17 shows an illusion using *X*-junctions to make atmospheres perceptible as such. The centers of the two diamond shaped regions are physically the same shade of light gray. However, the upper one seems to lie in haze, while the lower one seems to lie in clear air.

The single-reversing *X*'s surrounding the lower diamond indicate that it is a clearer region within a hazier region. The statistics of the upper region are elevated and compressed, indicating the presence of both attenuative and additive processes. Thus the statistical cues and the configural cues point in the same direction: the lower atmosphere is clear while the upper one is hazy.

The shifted Koffka rings

It is useful at this point to recall the modified Koffka display of figure 24.6(c). When the two halves are slid vertically, a set of sign-preserving *X*-junctions is created along the vertical contour. The junctions are consistent with transparency, and the contour becomes a strong atmospheric boundary between the left and right regions. The two semicircles are seen within different frameworks. The statistics on the two sides are different. In addition, grouping cues such as good continuation indicate that the left semicircle is connected to the light region on the right, and the right semicircle is connected to the dark region on the left. Thus there are several

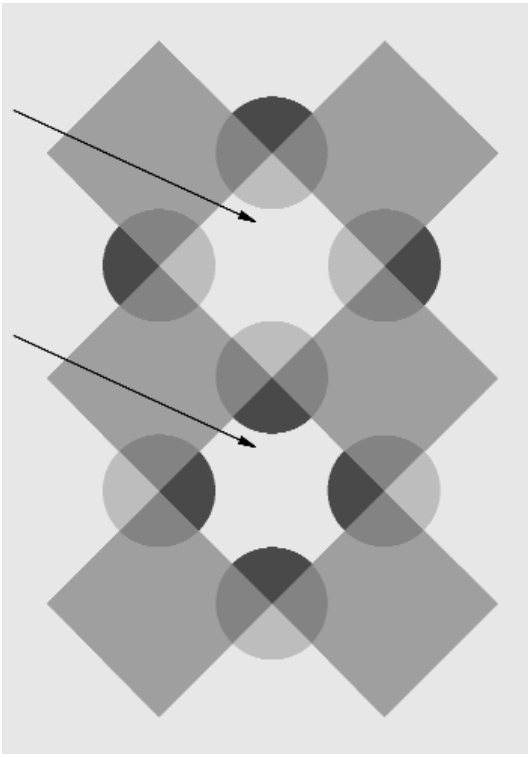

FIGURE 24.17 The haze illusion. The two marked regions are identical shades of gray. One appears clear and the other appears hazy.

cues that conspire to make the two semicircles look quite different.

T-junctions and White's illusion

White's illusion is shown in figure 24.18. The gray strips are the same. This is surprising: by local contrast, the left ones should look darker than the right ones. The left strips have a long border with white and a short border with black. The illusion is reversed from the usual direction. This effect has been interpreted in terms of the T-junctions (Todorovic, 1997, Gilchrist et al., in press). Patches straddling the stem of a T are grouped together for the lightness computation, and the cross-bar of the T serves as an atmospheric boundary. (cf. Anderson, 1997, for an alternative approach).

Zaidi et al. (1997) have shown that the action of T-junctions can be so strong that it overpowers traditional grouping cues such as coplanarity. Therefore the grouping rules for the lightness computation evidently differ from those underlying subjective belongingness.

Constructing a new illusion

One can intentionally combine statistical and configural cues

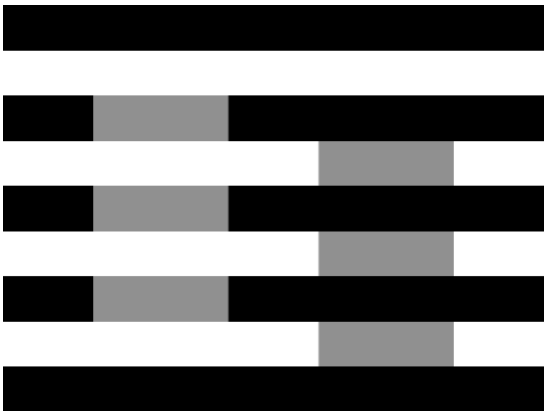

FIGURE 24.18 White's illusion. The gray rectangles are all the same shade of gray.

to produce large contrast illusions. In the "criss-cross" illusion of figure 24.19, the small tilted rectangles in the middle are all the same shade of gray. Many people find this hard to believe. The figure was constructed by the following principles: The multiple ψ-junctions along the vertical edges establish strong atmospheric boundaries. Within each vertical strip there are three luminances and multiple edges to establish articulation. The test patch is the maximum of the distribution within the dark vertical strips, and the minimum of the distribution within the light vertical strips. The combination of tricks leads to a strong illusion.

Each ψ-junction, by itself, would offer evidence of a 3D fold with shading. However, along a given vertical contour the ψ's point in opposite directions, which discourages the folded interpretation. Some subjects see the image in terms of transparent strips; others see it merely as a flat painting. However, all subjects see a strong illusion instantly. Thus, a 3D folded percept is not necessary: the illusion works even when "a ψ is just a ψ" (Hupfeld, 1931).

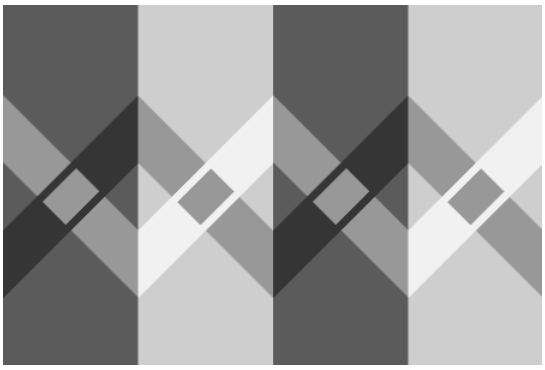

FIGURE 24.19 The crisscross illusion. The small tilted rectangles are all the same shade of gray.

The snake illusion

Similar principles can be used to construct a figure with *X*-junctions. Figure 24.20(a) shows an illusion we call the snake illusion (Somers and Adelson, 1997). The figure is a modification of the simultaneous contrast display shown at the right. The diamonds are the same shade of gray and they are seen against light or dark backgrounds. A set of half-ellipses have been added along the horizontal contours. The *X*-junctions aligned with the contour are consistent with transparency, and they establish atmospheric boundaries between strips. The statistics within a strip are chosen so that the diamonds are at extrema within the strip distributions. Note that the ellipses do not touch the diamonds, so the edge contrast between each diamond and its surround is unchanged.

Figure 24.20(b) shows a different modification in which the half-ellipses create a sinuous pattern with no junctions and no sense of transparency. The contrast illusion is weak; for most subjects it is almost gone. Thus, while figures 20(a) and (b) have the same diamonds against the same surrounds, the manipulations of the contour greatly change the lightness percept. In effect, we can turn the contrast effect up or down by remote control.

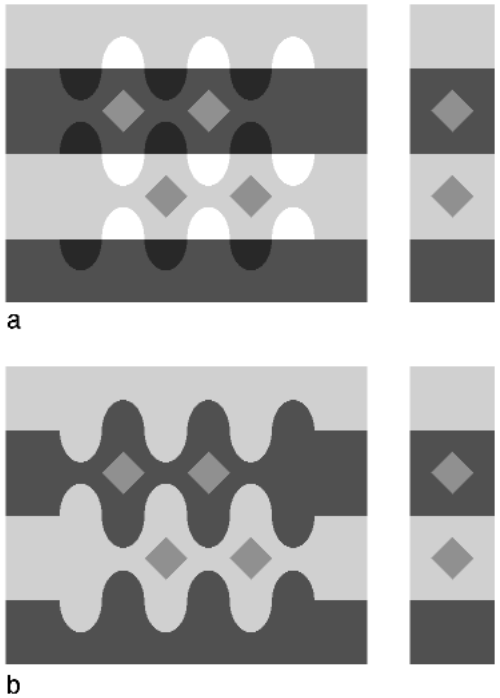

FIGURE 14.20 The snake illusion. All diamonds are the same shade of gray. (a) The regular snake: the diamonds appear quite different. (b) The "anti-snake": the diamonds appear nearly the same. The local contrast relations between diamonds and surrounds are the same in both (a) and (b).

Why should the illusion of figure 24.20(b) be weaker than in the standard SC? We have various observations suggesting that the best atmospheric boundaries are straight, and that curved contours tend to be interpreted as reflectance. The sinuous contours of figure 24.20(b) are not seen as atmospheric boundaries, and therefore the adaptive window is free to grow and to mix statistics from both light and dark strips.

Summary

Illusions of lightness and brightness can help reveal the nature of lightness computation in the human visual system. It appears that low-level, mid-level, and high-level factors can all be involved. In this chapter we have emphasized the phenomena related to mid-level processing.

Our evidence, along with the evidence of other researchers, supports the notion that statistical and configural information are combined to estimate the lightness mapping at a given image location. In outline, picture looks like this:

- At every point in an image, there exists an apparent atmospheric transfer function (ATF) mapping reflectance into luminance. To estimate reflectance given luminance, the visual system must invert the mapping, implicitly or explicitly. The inverting function at each point may be called the lightness transfer function (LTF).
- The lightness of a given patch is computed by comparing its luminance to a weighted distribution of neighboring luminances. The exact computation remains unknown.
- Classical mechanisms of perceptual grouping can influence the weights assigned to patches during the lightness computation. The mechanisms may include proximity, good continuation, similarity, and so on. However, the grouping used by the lightness system apparently differs from ordinary perceptual grouping.
- The luminance statistics are gathered within an adaptive window. When the samples are plentiful the window remains small, but when the samples are sparse the window expands. The window is soft-edged.
- The adaptive window can change shape and size in order to avoid mixing information from different atmospheres.
- Certain junction types offer evidence that a given contour is the result of a change in atmosphere. The contour then acts as an atmospheric boundary, preventing the information on one side from mixing with that on the other. A series of junctions aligned consistently along a contour produce a strong atmospheric boundary. Some evidence suggests that straight contours make better atmospheric boundaries than curved ones.

REFERENCES

ADELSON, E. H., 1993. Perceptual organization and the judgment of brightness. *Science* 262:2042-2044.

ADELSON, E. H., and P. ANANDAN, 1990. Ordinal characteristics of transparency. *Proceedings of the AAAI-90 Workshop on Qualitative Vision*, pp. 77-81.

ADELSON, E. H., and A.P. PENTLAND, 1996. The Perception of Shading and Reflectance. In *Perception as Bayesian Inference*, D. Knill and W. Richards, eds. New York: Cambridge University Press, pp. 409-423.

ANDERSON, B. L., 1997. A theory of illusory lightness and transparency in monocular and binocular images: the role of contour junctions. *Perception* 26:419-453.

AREND, L., 1994. Surface Colors, Illumination, and Surface Geometry: Intrinsic-Image Models of Human Color Perception. In *Lightness, Brightness, and Transparency*, A. Gilchrist, ed. Hillsdale, N.J.:Erlbaum, pp. 159-213.

BARROW, H. G., and J. TENENBAUM, 1978. Recovering intrinsic scene characteristics from images. In *Computer Vision Systems,* A. R. Hanson and E. M. Riseman, eds. New York: Academic Press, pp. 3-26.

BECK, J., K. PRAZDNY, and R. IVRY, 1984. The perception of transparency with achromatic colors. *Percept. Psychophysics* 35(5):407-422.

GILCHRIST, A. L., 1977. Perceived lightness depends on perceived spatial arrangement. *Science* 195(4274): 185-187.

GILCHRIST, A. L., and A. JACOBSEN, 1983. Lightness constancy through a veiling luminance. *J. Exp. Psych.* 9(6):936-944.

GILCHRIST, A. L., C. KOSSYFIDIS, F. BONATO, T. AGOSTINI, J. X. L. CATALIOTTI, B. SPEHAR, and J. SZURA, in press. A new theory of lightness perception. *Psych. Rev*.

HOCHBERG, J. E., and J. BECK, 1954. Apparent spatial arrangement and perceived brightness. *J. Exp. Psych.* 47:263-266.

HUPFELD, H, 1931. Lyrics to "As Time Goes By," Harms, Inc., ASCAP. See also *Casablanca*, Warner Brothers, 1942.

KATZ, D., 1935. *The World of Colour.* London: Kegan Paul.

KNILL, D., and D. KERSTEN, 1991. Apparent surface curvature affects lightness perception. *Nature* 351: 228-230.

LAND, E. H., and J. J. MCCANN, 1971. Lightness and retinex theory. *J. Opt. Soc. Amer*. 61:1-11.

LI, X., and A. GILCHRIST, in press. Relative area and relative luminance combine to anchor surfacelightness values. *Percept. Psychophysics.*

METELLI, F., 1974. Achromatic color conditions in the perception of transparency. In *Perception, Essays in Honor of J.J. Gibson* , R. B. McLeod and H. L. Pick, eds. Ithaca, NY: Cornell University Press, pp. 93-116.

REID, R. C., and R. SHAPLEY, 1988. Brightness induction by local contrast and the spatial dependence of assimilation. *Vision Res.* 28:115-132.

ROSS, W., and L. PESSOA, in press. Contrast/filling-in model of 3-D lightness perception. *Percept. Psychophysics.*

SINHA, P., and E. H. ADELSON, 1993. Recovering reflectance in a world of painted polyhedra. *Proceedings of Fourth*

International Conference on Computer Vision, Berlin; May 11-14, 1993; Los Alamitos, Calif: IEEE Computer Society Press, pp.156-163.

SOMERS, D. C., and ADELSON, E. H., 1997. Junctions, transparency, and brightness. *Invest. Ophthalmol. Vis. Sci. (Suppl.)* 38:S453.

SPEHAR, B., I. DEBONET, and Q. ZAIDI, 1996. Brightness induction from uniform and complex surrounds: A general model. *Vision Res.*, 36, 1893-1906.

TODOROVIC, D., 1997. Lightness and junctions. *Perception* 26(4):379-394.

ZAIDI, Q., B. SPEHAR, and M. SHY, 1997. Induced effects of backgrounds and foregrounds in three-dimensional configurations: The role of T-junctions. *Perception* 26(4):395-408.