Why we see what we do an empirical theory of vision.R Beau Lotto and Dale Purves. ***Optician*** (June 27, 2003): p2. (3294 words)

Full Text:

Beau Lotto and Dale Purves discuss the mechanisms by which we see what we do, and explain that we tend to see what a visual scene has typically signified in the past, rather than what it actually is in the present

For centuries, people have been fascinated by the fact that what we see - whether considered in terms of the brightness of objects, their colours, the arrangement of the objects in space, or their apparent motions - is often at odds with the underlying reality.

Insight into such discrepancies or 'illusions' of vision was provided by the Irish philosopher George Berkeley early in the 18th century. In his Essay Towards a New Theory of Vision, Berkeley pointed out the seemingly obvious, but deeply significant point, that we do not have direct access to the world around us.1

On the contrary, the only information our visual brains have to guide behaviour is the infinitely variable patterns of light that fall on the eye (namely, retinal images). This biological fact, which has long been appreciated by those in the visual sciences, creates the fundamental problem faced by humans, indeed any visual animal whose survival requires dealing with the 'sources' of retinal images.

The problem is that the pattern of light falling on the retina is meaningless as such, since all light stimuli conflate multiple aspects of the natural world (Figure 1). With respect to brightness or colour, for instance, the intensity and spectral distributions of light that activate retinal receptors are necessarily manifestations of the reflectances of objects being observed, their illumination, and the transmittance of the medium between the object and observer.

Given these considerations, it should be apparent that a particular retinal stimulus cannot directly specify the nature of its generative physical sources. As a result, identical objects under different conditions of illumination, at different distances and orientations from an observer can give rise to radically different stimuli, whereas radically different objects can generate the same stimulus (Figure 2). In short, the relationship between the world and our perception of it is, by its very nature, a fundamentally uncertain one.

If the retinal image cannot uniquely define the underlying reality to which observers must respond, how is it, then, that our visual systems generate behaviours that most of the time deal successfully with a world that cannot be apprehended directly?

The answer to this question is essential for an understanding of vision (and the brain) generally.

As we discuss here, a growing body of evidence indicates that the visual system of humans (and presumably other highly visual species) resolves this quandary by generating perceptions on a wholly empirical basis. That is, rather than elaborating sensations by an analysis of the components of the retinal image as such, what we see is evidently determined by feedback from the outcome of visually guided behaviour in the past that has gradually shaped the circuitry of the visual system to progressively improve performance.

What this means for vision is that what we see and how we see it must be understood, not in terms of the features of the retinal stimulus or the properties of the underlying objects as such, but according to the empirical significance of ambiguous stimuli in the experience of the species over the eons, and the experience over the lifetimes of individuals.2

The basis of brightness

Brightness is the perception elicited by the physical intensity of a light stimulus, and is arguably the most fundamental aspect of vision. Although a sensible expectation is that brightness should scale directly with the intensity of light (such that the more intense the light reaching the eye, the stronger the sensation of brightness), this is not always the case. In fact, two surfaces reflecting the same physical amount of light to the eyes typically look differently bright if the surfaces are observed in surroundings that are themselves returning different amounts of light. This phenomenon is called simultaneous brightness contrast, which is shown in Figure 3.

In the past, the explanation for this well known effect has been based on the fact that retinal neurons sending information from the eye to the visual part of the brain happen to respond more vigorously to a grey patch in a dark surround than the same grey patch on a light surround (for reasons that have to do with enhanced edge detection). If the apparent brightness of the patches were determined by the firing rate of retinal neurons, then the patch on a dark background would be expected to look brighter than the same patch on a lighter background.

The problem with this interpretation is that, among other things, patches embedded in scenes that have exactly the same surrounds can also be made to look differently bright. Indeed, as first shown by the 19th century physicist Wilhelm von Bezold, a target surrounded by a territory of predominantly higher luminance than the target, can, under the right circumstances, look brighter than the same target surrounded by territory of lower average luminance, which is just the opposite of the standard simultaneous brightness contrast effect (and the opposite of what the retinal firing rate explanation of brightness predicts - Figure 4a). An even more intriguing demonstration, called 'White's Illusion', also confounds this simple explanation (Figure 4b).3 How, then, can these puzzling facts about the relationship between the physical intensity of light and ensuing sensation of brightness be explained in terms of an empirical framework for vision?

Recall that the identical intensities of light arising from the two surface patches in question here are inherently ambiguous. They could signify similarly reflective surfaces under the same illuminant, or differently reflective target surfaces under different amounts of illumination. The only way to resolve this uncertainty is to take advantage of empirical information about what the source of the stimulus in question turned out to be in the past (determined by the success or failure of the related behaviour). Thus, to the extent that the context of the target patches in the preceding figures is consistent with past experience of similarly reflective target surfaces under the same illuminant, the targets will tend to appear similarly bright (because, to be behaviourally useful, things that are the same need to look the same). However, insofar as the overall stimulus is consistent with the experience of the target patches being differently reflective objects in different levels of illumination, the targets will tend to appear differently bright (because, by the same token, to be useful to the observer, things that are different need to look different).

Because the contextual information in the standard simultaneous brightness contrast stimulus is consistent with different surfaces under different illuminants and similar surfaces under similar illuminants, what the observer actually sees will reflect both possibilities (only in this way can vision take into account the full range of its experience). As a result, the targets look differentially bright because, in statistical terms, the stimulus is in some degree consistent with surfaces having different reflectances, and the visual system incorporates this probability into the resulting perception.4,5

Although this may seem a strange way to generate visual percepts, given the inevitable uncertainty of the information in the retinal image, this strategy may be the best - or even the only - way to resolve Berkeley's quandary.

A confirmation based on the Cornsweet effect

If this explanation of relatively simple brightness contrast stimuli is correct, then the same rationale should apply to the perceptual effects elicited by any stimulus in which target territories having the same luminance have typically turned out to be differently reflective objects in different amounts of light. Many more complex stimuli have been considered in these terms over the past few years. A good example is the perception of the so-called 'Cornsweet edge', a stimulus named after the psychologist who described it in the late 1960s (Figures 5 and 6).

In this effect, opposing luminance gradients that meet at an edge make physically identical adjoining regions look differently bright, the region adjacent to the lighter gradient appearing brighter than the region next to the darker gradient. Since this perceptual effect is opposite, the standard simultaneous brightness contrast effect, the Cornsweet stimulus provides yet another example of why explanations based on local contrast relationships do not work.

Despite its seemingly complicated structure, the Cornsweet edge effect can also be understood in the same empirical terms used to rationalise simpler effects.

The common denominator of the Cornsweet stimulus and conventional simultaneous brightness contrast stimuli is again that the percept can be understood in terms of the possible sources of physically identical target territories. As shown in Figure 6a, the equiluminant regions bordering the gradients that comprise the Cornsweet edge could have been generated by similarly reflective surfaces under the same illuminant (for example, painted gradients on the surface of a piece of paper on which light falls uniformly), or differently reflective surfaces under different intensities of illumination (for example, a cube with rounded edges one side of which is in light and the other in shadow).

Since both scenarios (and a host of others) are real possibilities, the percept elicited by the stimulus will, according to a wholly probabilistic framework of vision, take all the possible sources into account in proportion to their occurrence in the past. Since the identical targets in the Cornsweet stimulus will often have been generated by differently reflective surfaces in different illuminants, the target territories will look differently bright.7

If this statistical explanation based on past experience is correct, then the perceptual effect of the Cornsweet edge should be increased by altering the relative probabilities of the possible sources of the stimulus without changing the stimulus as such. As shown in Figure 6b, this is indeed the case.

Seeing colours

Given that these otherwise puzzling aspects of the sensations elicited by the intensity of light can be understood as a consequence of a wholly probabilistic strategy of vision, it is natural to ask whether the colour sensations elicited by different light spectra might also arise in this way. After all, the distribution of spectral power in a light stimulus (which is what gives rise to sensations of colour) are ambiguous for exactly the same reasons as the overall spectral intensity: illumination, reflectance and other factors that determine the characteristics of the light that reaches the eye are inevitably intertwined in the retinal image and cannot be disentangled.

A good starting point in thinking about colour sensations in these terms is simultaneous colour contrast (Figure 7a), a phenomenon similar to the brightness contrast effects already described. The standard stimulus for eliciting colour contrast is placing two targets with the same spectral composition on spectrally differently backgrounds. As in brightness contrast, the two targets look different, but now in terms of their respective colour qualities (namely, hue, saturation and colour brightness). In the past, most explanations of this phenomenon have been based on averaging across the entire stimulus (by adaptation to the average spectral content of the retinal image, for instance). As in brightness contrast, however, such schemes fail to account for the fact that colour contrast stimuli can be crafted in which the same average chromatic surrounds nonetheless elicit different colour percepts (as in Figure 7b).

An explanation of colour contrast (and colour constancy, which describes the related phenomenon in which the same object continues to appear similar in colour despite being under different illuminants) can, however, be given in wholly empirical terms.8 The sources of the target and surround in the standard colour contrast stimuli are, as all visual stimuli, profoundly uncertain: the same distributions of spectral power could have been generated by an infinite number of combinations of reflectances and illuminants (and other less critical factors). The visual system could, as for achromatic stimuli, resolve this dilemma by using feedback from the success or failure of the past behavioural responses to spectral stimuli. In this case, the percept elicited by a given stimulus would be determined by the relative frequencies of occurrence of the real-world combinations of reflectances and illuminants that had given rise to that distribution of spectral power in the past.

If perceptions of colour are indeed generated in this way, then the same spectral target on two differently chromatic backgrounds would be expected to give rise to different chromatic sensations. The reason is that, in addition to requiring behaviours appropriate to the same reflectances in the same illuminant, such stimuli would in other instances have required behaviours appropriate to targets that arise from different reflectances in different illuminants.

Consequently, a spectral stimulus should elicit a sensation that incorporates these possible (and indeed all possible) underlying sources in proportion to their past occurrence in human experience with spectral stimuli.4,9,10

To convey the merits of this way of understanding colour percepts, we devised a stimulus that looks something like a Rubik's Cube, in which physically identical tiles are placed at the centres of each side of the cube (Figure 8). By understanding the effects of spectral differences in this probabilistic framework, we could generate colour contrast and constancy effects that are much more dramatic than the usual textbook illustrations of these phenomena (the stimulus in Figure 7a, for example). Thus, when all the information in the scene with the cube was made consistent with either 'yellowish' illumination or 'bluish' illumination, the same identical tiles on the surface of the cube that both appeared to be the same shade of grey in a neutral context could be made to appear either blue or yellow (Figure 8). This manipulation provides an example of colour contrast made especially dramatic by empirical manipulation of the information in the scene.

Conversely, tiles that appeared differently coloured in a neutral setting could, by changing the probability of their possible sources, be made to look the same colour, thus providing an equally dramatic demonstration of colour constancy.

These demonstrations show not only that colour contrast and constancy are probabilistically determined, but that these seemingly opposite effects are both manifestations of the empirical strategy the visual system evidently uses to generate percepts.

Perception of scene geometry

Vision scientists noted long ago that the perception of lines doesn't always accord with the real-world geometry of the underlying objects.

For instance, the angles formed by lines making (or implying) an acute angle are seen as being a few degrees larger than they really are, whereas obtuse angles are seen as being a few degrees smaller.

Despite a great deal of speculation about this anomaly dating from the latter part of the 19th century, there has been no consensus about its origin. We therefore asked whether these geometrical 'misperceptions' might be explained in much the same empirical terms as the misperceptions of brightness and colour described in the preceding sections.

The stimuli giving rise to perceived angles, like luminances or spectral power distributions, are profoundly ambiguous. Thus an angle projected onto a surface (the retina, for example) can arise from objects having a variety of subtenses and arm lengths, arranged in infinitely many 3-D orientations (Figure 9). In interacting with the objects that give rise to a particular angle projection on the retina, observers through the ages would have experienced that the real-world angles giving rise to a given angle in the retinal projection vary greatly, and, as it turns out, systematically. In consequence, the perceptions elicited by different angles projected onto the retina would be expected to correspond to these frequency distributions.

To test this interpretation, we first determined the probability distribution of the all the possible 3-D sources of a projected angle. When these distributions for all possible angles are computed using the principles of projective geometry, it is apparent that the most frequently occurring sources of acute angle projections turn out to have solid angles larger than the subtense of the projected stimulus. Conversely, the sources of obtuse angle projections have typically been generated by sources that are somewhat smaller than the projected angle. Right angle projections and straight lines, are generated by sources that, on average, have the subtense of the object itself. If percepts are empirically determined, then the visual system should generate perceptions of angles that incorporate and reflect these statistical facts of projective geometry.11

To assess this prediction, subjects were asked to report their perceptions of different angular stimuli in a series of tests in which the adjustment of a test line indicated the angle subtense they were actually seeing. For example, if the subject perceived the angle to be bigger than its actual subtense, then the test line would be set in a position that revealed this discrepancy (ie, not quite parallel with the angle arm in the example shown here). The results derived from such tests tallied well with the probability distribution of the possible sources of the corresponding stimuli, indicating that the spatial arrangement observers see is neither the retinal projection nor its real-world source, but its empirical (past) significance.11

Taken together, this evidence drawn from the perception of brightness, colour and geometry formally describes and supports the idea that the problem first emphasised by Bishop Berkeley is resolved by generating visual percepts according to the probability distribution of the possible sources of the visual stimulus, whatever it may be. As a result, observers see what a visual scene has typically signified in the past, rather than what it actually is in the present. We see what we do, therefore, because the statistical relationship between the qualities of the stimulus and the objects that generated them in past experience are the sole basis by which the visual system can overcome the quandary posed by the inherent ambiguity of visual stimuli.

It is in this empirical/statistical context that one must continue (and in some instances begin) to understand the mechanisms by which we see what we do.

References

1 Berkeley G. (1709/1975) Philosophical works including works on vision (Ayers MR, Ed.). London: Everyman/J.M Dent.

2 Purves DP and Lotto, RB. (2002) Why we see what we do: A wholly probabilistic strategy of vision. Sinaur Associates INC. (Sunderland Massachusetts) and Macmillan Press (London, UK).

3 White M (1979) A new effect of pattern on perceived lightness. Perception, 8:413-416.

4 Lotto RB and Purves D. (1999) The effects of colour on brightness. Nature Neuroscience, 2:1010-1014.

5 Purves D, Lotto RB, Williams SM, Nundy S and Yang Z. (2001) Why we see things the way we do: Evidence for a wholly empirical strategy of vision. Philosophical Transaction of the Royal Society of London B, 356:285-297.

6 Cornsweet TN. (1970) Visual Perception. New York: Academic Press.

7 Purves D, Shimpi A and Lotto RB. (1999) An empirical explanation of the Cornsweet effect. Journal of Neuroscience, 19: 8542-8551.

8 Lotto RB and Purves D. (2002) A rationale for the structure of colour space. Trends Neuroscience, 25: 84-88.

9 Lotto RB and Purves D. (2000) An empirical explanation of colour contrast. Proceedings of the National Academy of Sciences USA 97: 12834-12839.

10 Lotto RB and Purves D. (2002) The empirical basis of colour perception. Consciousness and Cognition, 11:609-629.

11 Nundy S, Lotto RB, Coppola D, Shimpi A and Purves D. (2000) Why are angles misperceived? Proceedings of the National Academy of Sciences, USA 97(10): 5592-5597.

This article is based on the book Why We See What We Do: An Empirical Theory of Vision by R Beau Lotto and Dale Purves (265 pages). optician readers can obtain it at the special price of [pounds sterling]24.99 plus [pounds sterling]2.00 postage from Jana Bek, Brunel Road, Houndmills, Basingstoke, RG21 6XS.

R Beau Lotto is a lecturer funded by Fight for Sight at the Institute of Ophthalmology, University College London. Dale Purves is the George Barth Geller Professor for Research in Neurobiology, Professor of Psychological and Brain Sciences and chairman of the Department of Neurobiology at Duke University Medical Centre in the US

Source Citation: Lotto, R Beau, and Dale Purves. "Why we see what we do an empirical theory of vision." Optician (June 27, 2003): 2. General OneFile. Gale. Duke University Library - Perkins. 12 Feb. 2008
<http://find.galegroup.com/itx/start.do?prodId=ITOF>.

Gale Document Number:A104572007