Human Perception: Visual Heuristics in the Perception of Glossiness

New insights into the perception of surface glossiness embody a conceptual change in perception research. Instead of estimating the physical properties of objects, the brain exploits ‘invariants’ — even though these sometimes make us get the answer wrong.

Roland W. Fleming

Working out how the brain estimates the material properties of surfaces is one of the most exciting and rapidly developing areas of visual neuroscience. The perception of glossiness is of particular interest, partly because very small changes in the image, such as the addition of a small highlight, can have radical effects on how the brain interprets whole surfaces. One recurring theme in research on this topic is that factors other than the physical glossiness of a surface can have large and unpredictable effects on its perceived glossiness [1–7]. Changing the shape or lighting conditions can often make a huge difference to how glossy a surface appears; for example, the apple in Figure 1A appears less glossy than it does in Figure 1B because the lighting is more diffuse in A than in B. So far there is little consensus on when and why these different effects occur. Sometimes changing the shape makes surfaces appear more glossy, sometimes less, in a seemingly erratic way. Why does this occur? Is there some unifying principle that can explain the inconsistent effects of shape and lighting on perceived glossiness? A recent study by Marlow and colleagues [8], reported in this issue of Current Biology, suggests the explanation lies in the brain’s use of a number of simple — but imperfect — heuristics.

The main argument of Marlow et al. [8] is that the visual system doesn’t actually estimate glossiness per se; instead, they argue, the brain measures a set of simple ‘proximal stimulus properties’ that approximately correlate with glossiness in typical circumstances, but which may get the answer wrong under other conditions. In other words, rather than estimating the physical properties of surfaces, the brain measures whatever it can from the image itself, and identifies useful statistical patterns or cues among the measurements. The authors show that such cues can predict the failures as well as the successes of human gloss perception.

Marlow et al. [8] identified several different properties of highlights that represent their clarity and salience in the image — size, contrast, sharpness and binocular separation from the surface (see Figure 1C–E). Computing these quantities directly from the image is not trivial, so instead, the authors showed all the images to new participants and asked them to judge each property independently. The subjects were not asked anything about the glossiness of the surface. They just had to report how large, or high contrast the highlights appeared.

As expected, the participants’ responses also vary systematically with lighting and surface relief. Importantly, the authors found that a weighted combination of the ratings for the individual cues could account for 94% of the variance in the glossiness ratings from the other participants. In other words, the simple cues can account for almost all of the seemingly inconsistent effects of lighting and shape on the perception of glossiness. When asked to judge glossiness, subjects actually report the extent to which a surface
manifolds highlights — this is the definition of ‘glossiness’ for the human brain.

As this was a correlational study, Marlow et al. [8] cannot be certain that it is exactly these specific cues that determine glossiness perception. There are potentially many other ways of capturing the intuition of bigger, more salient highlights. However, this is not the important point of the study. The important point is that instead of representing glossiness in physical terms, the brain uses a set of imperfect cues or heuristics. These heuristics predict the errors as well as the successes of gloss perception in the authors’ experiments.

The idea that the brain exploits heuristics is almost as old as the study of perception itself. However, Marlow et al. [8] have captured a change in current thinking about the goals of perception, which applies to much more than just gloss ratings. Many problems in perception, from visual stereopsis to auditory pitch perception can be posed as a process of estimating physical parameters — distances, frequencies, and so on. But this may be the wrong way of thinking about the biological problems that the brain solves. Rather than estimating physical properties of the world, it may be more important — and easier — to compute systematic relations between internal states. In other words, it may be better to identify quantities that can guide decisions consistently in the face of changes of irrelevant variables, whether or not they map cleanly onto physical properties. For example, ‘color constancy’ should not be posed, as it usually is, as a problem of estimating the spectral reflectance of surfaces under varying illuminants; rather, color constancy is the process of identifying image measurements that are as close as possible to invariant across transformations caused by other scene factors, such as lighting, shape or viewpoint. Only with this alternative goal in mind does it make sense that purples and reds are subjectively similar to one another — that is, close to one another in the hue circle — even though they lie at opposite ends of the visible spectrum.

Other perceptual tasks make this idea even more explicit. For example, most people can easily recognize familiar linguistic accents across a wide range of differences between speakers (age, gender, and so on). In other words, we have good ‘accent constancy’. At the same time, it probably doesn’t make sense to pose accent constancy as the estimation of physical parameters of the speaker’s vocal tract. Instead, it is more likely that the brain identifies auditory quantities that are diagnostic of accent but relatively stable across age, gender and other vocal attributes. After all, the hidden Markov models used by computer speech recognition systems identify predictive features in the input stream, rather than model the physics of speech production. Thus, more generally, if we want to understand how the brain perceives, we should change the way we pose the aims of perception.

References

Department of Psychology, University of Giessen, Otto-Behaghel-Str. 10/F, 35394, Giessen, Germany.
E-mail: roland.w.fleming@psychol.uni-giessen.de
http://dx.doi.org/10.1016/j.cub.2012.08.030